

# An Adaptive Multiscale Information Fusion Approach for Feature Extraction and Classification of IKONOS Multispectral Imagery Over Urban Areas

Xin Huang, Liangpei Zhang, and Pingxiang Li

**Abstract**—An adaptive multiscale information fusion algorithm is proposed to extract the spatial features and classify IKONOS multispectral imagery. It is well known that combining spectral and spatial information can improve land use classification of very high resolution data. However, many spatial measures refer to the window size problem, and the success of the classification procedure using spatial features depends largely on the window size that was selected. In this letter, we first propose an optimal window selection method, based on the spectral and edge information in a local region, for choosing the suitable window size adaptively; second, the multiscale information is fused based on the selected optimal window size. In order to evaluate the effectiveness of the proposed multiscale feature fusion approach, the spatial features that were extracted by the gray-level cooccurrence matrix are utilized for multispectral IKONOS data. The results show that the proposed algorithm can select and fuse the multiscale features effectively and, at the same time, increase the classification accuracy.

**Index Terms**—IKONOS, multiscale information fusion, very high resolution (VHR), window size.

## I. INTRODUCTION

EARTH observation data are becoming available at increasingly higher resolutions. Very high resolution (VHR) satellite sensors such as QuickBird, IKONOS, and SPOT-5 provide a large amount of information, thus opening up avenues for new remote sensing applications. However, it seems evident that the new VHR images do create additional problems in terms of information extraction and automatic classification [1], [2]. The resulting high intra-class and low inter-class variabilities lead to a reduction in the statistical separability of the different land-cover classes in the spectral domain, and conventional spectral classification methods have proven to be inadequate for the VHR data [3]. The introduction of spatial features is an effective method of addressing this challenge, and it is well known that combining spatial and spectral information can improve land use classification for VHR satellite imagery [3]–[6].

Manuscript received January 18, 2007; revised May 21, 2007. This work was supported in part by the 973 Program of the People's Republic of China under Grant 2006CB701302 and in part by the National Natural Science Foundation of China under Grant 40471088 and Grant 40523005.

The authors are with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: zlp62@public.wh.hb.cn).

Digital Object Identifier 10.1109/LGRS.2007.905121

Therefore, many spatial feature extraction methods have been proposed and tested. Some commonly applied spatial procedures include the gray-level cooccurrence matrix (GLCM) [1], the wavelet-based spatial features [7], the texture features based on Markov random field models [8], and the technique of morphological profiles [9].

These spatial measures increase the classification performance of VHR images; however, usually, these features are extracted based on a moving window of a fixed size, and the success of the classification procedure using spatial measures depends largely on the window size that was selected. Different window sizes for wavelet features were tested and compared in [1], and the largest window was found to be most effective for regions of homogeneous texture. In [2], the effect of window size on classification accuracy was examined, and the results show that the window size should be selected according to homogeneity or heterogeneity in a local area. In [10], the optimal window size for texture extraction was identified by the resolution and the sizes of the objects in the image. In some literature, a small window of fixed size was employed since it can preserve detailed structures and avoid the influence of the texture features of adjacent different types [7].

Based on some approaches that were previously presented in the literature, in this letter, we propose an adaptive multiscale information fusion algorithm for addressing the issue of window size in spatial feature processing. The algorithm is made up of two steps: 1) A selection index is presented for choosing the optimal window size for each pixel, and 2) a multiwindow fusion approach is used based on the selected optimal window size in order to exploit multiscale information for the classification of VHR data.

## II. ADAPTIVE MULTIWINDOW INFORMATION FUSION

A geostatistical analysis indicated that there was no single window size that would adequately characterize the range of textural conditions that are present in remotely sensed images [2]. Therefore, it is not sound to use a fixed window size for feature extraction, particularly, for VHR images. In this letter, a multiscale feature fusion algorithm is proposed to utilize the information across different local windows adaptively.

### A. Adaptive Window (AW) Selection

It is preferable to characterize a texture using a large window, so that a sufficient amount of information is considered. Yet, a small window is necessary to accurately locate the boundaries between homogeneous regions. Therefore, in this letter, the window size is identified based on the spectral and edge information around the pixel under consideration. The proposed AW size selection algorithm can be described in four steps.

- 1) First, the edge features for different spectral bands can be defined as

$$\begin{cases} E_b(x, y) = 1, & \text{if pixel } (x, y) \text{ in band } b \text{ is an edge} \\ E_b(x, y) = 0, & \text{if pixel } (x, y) \text{ in band } b \text{ is a not an edge} \end{cases} \quad (1)$$

where edge information is obtained using the Canny filter [11], since it is an effective edge detector having a low probability of false or missing edges and a high accuracy of edge positioning.

- 2) After edge detection, a concept of edge density is defined for each pixel, i.e.,

$$ED_b^m(i, j) = \frac{\sum_{(x,y) \in [w_m(i,j)]^2} E_b(x, y)}{[w_m(i, j)]^2} \quad (2)$$

where  $w_m(i, j) \times w_m(i, j)$  denotes the local window around the central pixel  $(i, j)$  at scale  $m$  ( $1 \leq m \leq N$ ). The values of edge density indicate the heterogeneity around the central pixel for different window areas. For a pixel that is close to the edge regions, the edge density shows a larger value; for a pixel that is far from the edge pixels or in a homogeneous area, a smaller value is obtained. It should be noted that the range of edge density is  $[0, 1]$ .

- 3) Based on the edge density for different local areas, a window selection index (WSI) is defined for choosing the optimal window for the central pixel  $(i, j)$ , i.e.,

$$WSI_b^m(i, j) = ED_b^m(i, j) \times SD_b^m(i, j) \quad (3)$$

where  $SD_b^m(i, j)$  is the standard deviation in the  $w_m(i, j) \times w_m(i, j)$  window area. The standard deviation indicates the spectral heterogeneity in different local regions.  $ED_b^m(i, j)$  is used as the coefficient of  $SD_b^m(i, j)$ , considering that its range is  $[0, 1]$ , and at the same time, it is also an indicator of heterogeneity.  $WSI_b^m(i, j)$  combines the spectral and edge information around the central pixel. A larger value of  $WSI_b^m(i, j)$  shows that the information at scale  $m$  contains more detailed features, and hence, if this  $w_m(i, j) \times w_m(i, j)$  window is employed for pixel  $(i, j)$ , some important edge and structural information may be smoothed.

- 4) The optimal window size for a central pixel in band  $b$  can then be determined by

$$OW_b(i, j) = \arg \min_{1 \leq m \leq N} [WSI_b^m(i, j)]. \quad (4)$$

A low value of  $OW_b(i, j)$  shows that edge or geometrical detail is present around the central pixel, and hence, the information in a small window is more reliable for classification. A large  $OW_b(i, j)$  indicates that the central pixel lies in a homogeneous area, and a large window is used to improve the classification. It is worth noting that the optimal window size can be different in the different spectral bands since a given pixel could be an edge in a particular spectral band but not necessarily in others. Therefore, in (4), the optimal window is relative to the spectral band.

### B. Multiscale Feature Fusion

The calculation of spatial feature is relative to multiscale and multispectral information; accordingly, the spatial feature set can be defined as

$$\begin{aligned} \text{spatial feature set : } F &= \{F(1), \dots, F(b), \dots, F(B)\} \\ \text{band } b : F(b) &= \{F(b, 1), \dots, F(b, m), \dots, F(b, N)\} \\ &\text{with } 1 \leq b \leq B \text{ and } 1 \leq m \leq N \end{aligned} \quad (5)$$

where  $B$  and  $N$  represent the total number of multispectral bands and multiscale dimensionality, respectively.  $F(b, m)$  represents the feature set that was extracted using a  $w_m \times w_m$  window size in band  $b$ . In this letter, a detail-preservation scale-driven approach is utilized for the integration of multiwindow information [12]. Some new feature sets are computed by averaging all the sequential combinations for windows whose size is smaller than the optimal window size, i.e.,

$$\overline{F}_{(i,j)}(b) = \frac{1}{OW_b(i, j)} \sum_{x=1}^{OW_b(i,j)} F_{(i,j)}(b, x) \quad (6)$$

where  $\overline{F}_{(i,j)}(b)$  is the multiscale fusion feature value for pixel  $(i, j)$ , and optimal window  $OW_b(i, j)$  represents the largest window that is included in the average operation for pixel  $(i, j)$ . In (6), the features at windows that are smaller than the optimal one are averaged, as this can enhance the homogeneity and, at the same time, preserve the detailed and edge information. Therefore, the new spatial feature set that is obtained by adaptively fusing the multiscale information can be described as

$$\text{spatial feature set : } \overline{F} = \{\overline{F}(1), \dots, \overline{F}(b), \dots, \overline{F}(B)\}. \quad (7)$$

In order to validate the proposed AW approach, a commonly used multiscale method is compared. (In this letter, this method is called the multiple windows (MW) algorithm.) The MW algorithm extracts multiscale features based on a set of concentric windows around each pixel [13]. It exploits the information in all the windows, and a feature selection algorithm is then used to reduce the dimensionality of the multiscale feature set [9]. In this letter, an unsupervised feature selection approach, i.e., the feature similarity index (S-Index) [14], is used to select the optimal subset of  $F$ . The dimensionality after feature selection is set to the number of multispectral bands  $B$ . S-Index is

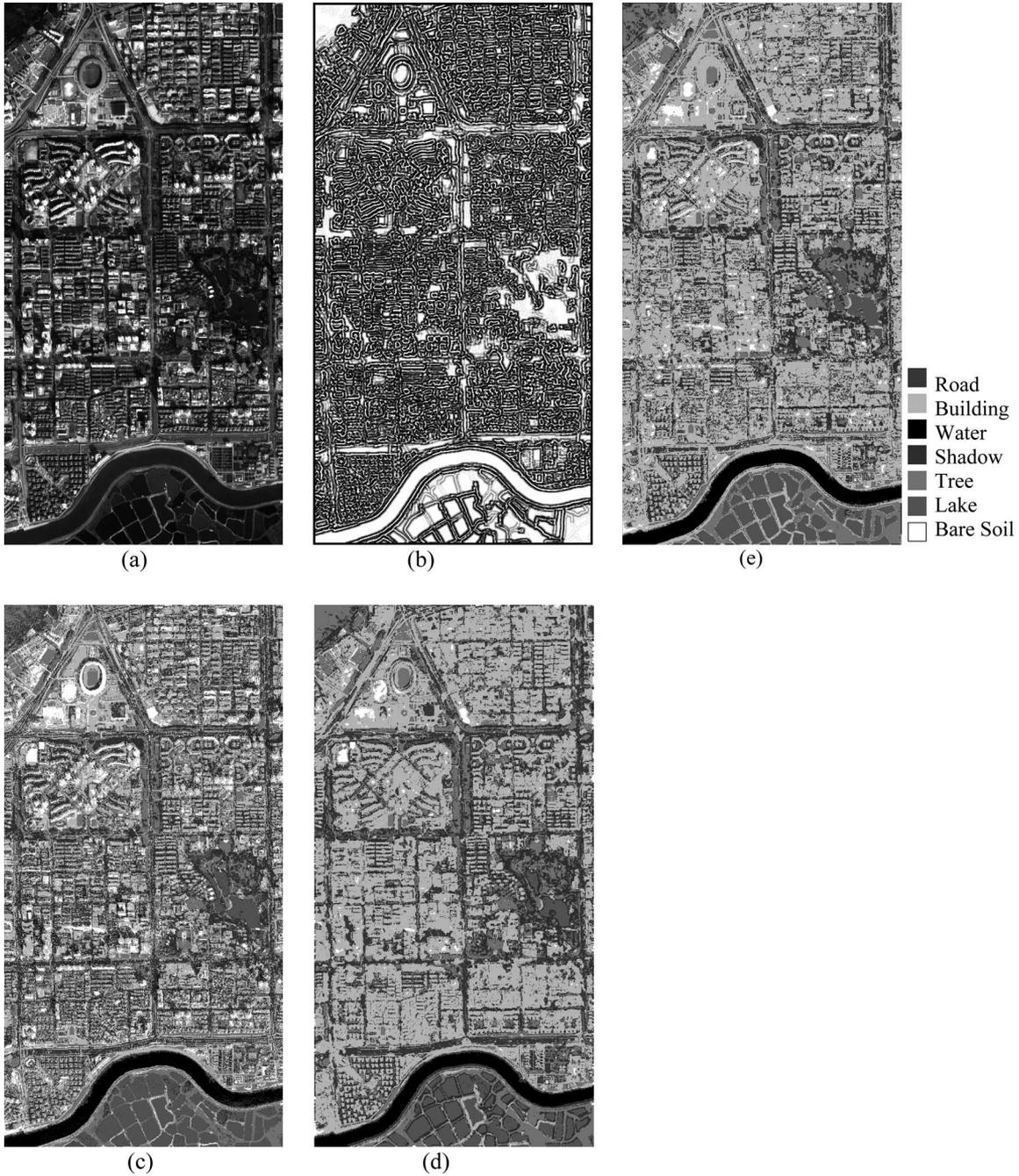


Fig. 1. (a) Test image of IKONOS for Shenzhen City in the south of China. (b) Results of adaptive scale selection. (c) Classification map for spectral classification with SVM classifier. (d) and (e) Results of adaptive information fusion algorithm for dissimilarity and mean measurement, respectively.

based on measuring the similarity between features whereby redundancy therein is removed. It does not need any searching and therefore is fast.

### III. EXPERIMENT

In the experiments, the adaptive multiscale fusion algorithm is implemented on the GLCM measures and multispectral IKONOS images. The imagery that was used for this experiment was acquired by the IKONOS commercial remote sensing satellite over Shenzhen City in the south of China and consists of four multispectral bands with 4-m resolution. The test image

comprised 998 lines and 509 columns, covering about  $4.0 \text{ km} \times 2.0 \text{ km}$ , and is shown in Fig. 1(a). This image presents a typical urbanized area in China, which includes rivers, lakes, long roads, collective and dense buildings, bare soil, and trees.

#### A. Experiment Approach

In the experiment, GLCM features are calculated on the IKONOS multispectral images with an interpixel distance of 1 and with four window sizes ( $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ , and  $9 \times 9$ ). The measures of mean and dissimilarity are used, and the directionality of GLCM is suppressed by averaging the extracted

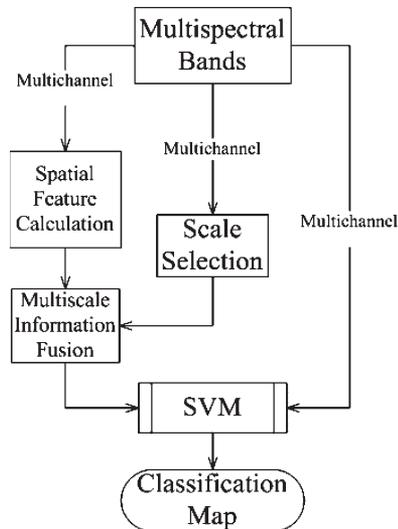


Fig. 2. Flow chart for the adaptive multiscale fusion and classification algorithm.

features over four directions [2]. In the experiment, the support vector machine (SVM) is adopted to classify all the different feature sets (including spectral classification), and the spatial features are added as additional bands to the multispectral images. The flowchart for the adaptive multiscale fusion algorithm is shown in Fig. 2.

### B. Results and Analysis

Fig. 1(b) is obtained by superimposing the optimal window maps of all the spectral bands, where cold colors indicate where smaller window sizes are used and hot colors indicate larger ones. It can be observed that smaller windows are set to edge and detailed pixels, and larger ones are employed for homogeneous regions. The classification accuracies for multiwindow feature sets of dissimilarity are shown in Table I, and the highest accuracies for different information classes are highlighted in gray. Three statistics, i.e., the Kappa coefficient, overall accuracy (OA) and average accuracy (AA), are used to evaluate the classification effectiveness. The AA represents how much the information class has been discriminated by the classifier, irrespective of the number of pixels that belong to each class in the ground truth. From Table I, we have two observations.

- 1) Comparison of the results for the four window sizes: Compared with the spectral classification ( $1 \times 1$  window size), the introduction of spatial information clearly enhances the results. Among the classification results for  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ , and  $9 \times 9$  window areas, the highest accuracy is obtained by the  $5 \times 5$  window. However, it is worth noting that, among the four window sizes, not all of the information classes reach their highest accuracies with the  $5 \times 5$  window. Larger windows, such as those with  $7 \times 7$  and  $9 \times 9$  sizes, obtain higher accuracies for road, building, river, bare soil, and lake; however, poor results are obtained for tree and shadow. It seems that the larger window size is reliable for homogeneous

and extensive objects. The  $5 \times 5$  window acquires the highest Kappa and OA because there is a tradeoff for the objects with different sizes and shape. It can be seen that a fixed window size is not the most effective for different information classes.

- 2) Multiscale information fusion: The proposed adaptive information fusion algorithm outperforms the classification with a  $5 \times 5$  window size. Comparing the information fusion approach to the results with fixed window sizes, the improvements in Kappa are 8.5%, 3.3%, 4.0%, and 5.4% for window sizes of  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ , and  $9 \times 9$ , respectively. The respective improvements in OA are 7.0%, 2.7%, 3.3%, and 4.5%. By observing the class-specific accuracies in Table I, it is found that the accuracies that were achieved by AW are comparable to the respective highest accuracies that were obtained by the different windows with fixed sizes. Therefore, the proposed algorithm is effective for multiscale objects, which can also be verified from the improvements of AAs: 4.0%, 2.7%, 3.7%, 5.1%, and 3.0% for window sizes of  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ , and  $9 \times 9$ , and the MW approach, respectively.

In this experiment, another measure of GLCM, namely, the mean, is implemented. The accuracy statistics for this measurement are provided in Table II. From the table, it can be seen that, among all the windows with fixed sizes, the highest Kappa and OA are acquired by the smallest ( $3 \times 3$ ) window. Compared to the fixed window methods, the improvements that were obtained by the AW approach are 2.3%, 4.1%, 5.9%, and 8.2% in Kappa for  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ , and  $9 \times 9$  window sizes, respectively. The respective improvements in OA are 2.0%, 3.6%, 5.1%, and 6.9%. The best classification rates for *building*, *river*, *tree*, *lake*, and *shadow* are achieved by the AW approach. As for the classification rates of *road* and *bare soil*, the results that were obtained by AW are similar and comparable to the respective highest rates. Meanwhile, the AW algorithm achieves the highest AA. It can be concluded that the proposed algorithm can increase the classification accuracies for objects with different scales and sizes and exploit the multiscale feature in different window areas effectively. The approach of adaptive scale selection and fusion is meaningful, considering that, conventionally, the optimal window size is determined by trials and may be different in different test areas.

It is worth noting that the AW approach outperformed MW for both dissimilarity and mean measures, mainly because of the optimal window selection and the multiscale information fusion in the feature level. The experimental results show that the AW and information fusion approaches are effective for multiscale feature extraction and classification. Fig. 1(c) shows the result of spectral classification; Fig. 1(d) and (e) shows the classification maps of adaptive multiscale fusion for dissimilarity and mean, respectively. Analyzing the results in Fig. 1(c)–(e), it can be observed that the pure spectral feature cannot effectively discriminate the spectrally similar objects, such as *river–shadow–lake* and *building–road*. The multiscale fusion approaches can improve the classification and reduce the pepper–salt effect and, at the same time, preserve the detailed and structural features in the image.

TABLE I  
CLASSIFICATION ACCURACIES (PERCENT) FOR THE DIFFERENT FEATURE SETS OF THE DISSIMILARITY MEASUREMENT

Window Size	Road	Building	River	Bare Soil	Shadow	Tree	Lake	Kappa	OA (%)	AA (%)
1	73.5	56.0	95.3	90.7	76.2	72.5	97.8	0.702	75.1	80.3
3	67.1	80.2	96.8	92.9	65.7	74.0	97.5	0.760	80.3	82.0
5	69.8	88.4	96.7	92.5	70.6	81.4	97.5	0.812	84.6	85.3
7	72.1	90.8	97.5	94.5	68.3	68.8	98.3	0.805	84.0	84.3
9	70.2	88.3	97.8	93.6	65.5	67.3	97.9	0.791	82.8	82.9
MW	67.0	94.3	93.7	92.9	61.0	89.8	96.3	0.820	85.3	85.0
AW	80.9	88.9	94.4	95.3	70.1	88.7	97.3	0.845	87.3	88.0

TABLE II  
CLASSIFICATION ACCURACIES (PERCENT) FOR THE DIFFERENT FEATURE SETS OF THE MEAN MEASUREMENT

Window Size	Road	Building	River	Bare Soil	Shadow	Tree	Lake	Kappa	OA (%)	AA (%)
3	72.1	70.6	95.7	89.5	76.0	70.2	98.0	0.742	78.7	81.7
5	73.0	64.1	96.8	91.9	74.5	69.6	98.1	0.724	77.1	81.1
7	75.3	62.0	98.0	88.8	65.9	66.9	97.3	0.706	75.6	79.2
9	74.6	64.0	96.3	86.8	64.3	66.2	96.5	0.683	73.8	78.4
MW	76.9	72.8	97.3	86.2	74.8	72.1	97.2	0.763	80.4	82.5
AW	73.2	73.6	99.8	88.0	81.7	74.1	99.5	0.765	80.7	84.3

#### IV. CONCLUSION

This letter concerns the problem of window size selection and multiscale information fusion. The size of the processing window is an important issue for spatial feature extraction and classification of VHR satellite imagery. In this letter, a selection index is presented for choosing the optimal window size adaptively based on the multispectral and edge information around the central pixel. This algorithm is sound, considering that large windows are used for homogeneous and nonedge areas, and small ones are utilized for edge regions. A multiscale information fusion approach is then employed for integrating the information within the windows that were smaller than the optimal one. This averaging processing can preserve the detail and structural features in the VHR imagery.

In the experiment on the IKONOS image, the proposed algorithm clearly improves the classification results of fixed window sizes. Two spatial features, i.e., dissimilarity and mean, are used for the evaluation, and similar results are obtained. The information fusion approach considers all the reliable multiwindow features, and hence, it is suitable for objects with different scales and sizes. It is worth noting that the proposed approach is independent of the spatial features and the classifiers.

#### REFERENCES

- [1] S. W. Myint, N. S. N. Lam, and J. M. Tylor, "Wavelets for urban spatial feature discrimination: Comparisons with fractal, spatial autocorrelation, and spatial co-occurrence approaches," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 7, pp. 803–812, Jul. 2004.
- [2] A. Puissant, J. Hirsch, and C. Weber, "The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery," *Int. J. Remote Sens.*, vol. 26, no. 4, pp. 733–745, Feb. 2005.
- [3] F. Dell'Acqua, P. Gamba, A. Ferari, J. A. Palmason, J. A. Benediktsson, and K. Arnason, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 322–326, Oct. 2004.
- [4] X. Huang, L. Zhang, and P. Li, "Classification and extraction of spatial features in urban areas using high-resolution multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 260–264, Apr. 2007.
- [5] J. Chanussot, J. A. Benediktsson, and M. Fauvel, "Classification of remote sensing images from urban areas using a fuzzy possibilistic model," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 40–44, Jan. 2007.
- [6] J. Palmason, J. A. Benediktsson, J. R. Sveinsson, and J. Chanussot, "Fusion of morphological and spectral information for classification of hyperspectral urban remote sensing data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2006, pp. 2506–2509.
- [7] M. Acharyya, R. K. De, and M. K. Kundu, "Segmentation of remotely sensed images using wavelet features and their evaluation in soft computing framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2900–2905, Dec. 2003.
- [8] Y. Dong, A. K. Milne, and B. C. Forster, "Segmentation and classification of vegetated areas using polarimetric SAR image data," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 321–329, Feb. 2001.
- [9] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.
- [10] Y. O. Ouma, T. G. Ngigi, and R. Tateishi, "On the optimization and selection of wavelet texture for feature extraction from high-resolution satellite imagery with application towards urban-tree delineation," *Int. J. Remote Sens.*, vol. 27, no. 1, pp. 73–104, Jan. 2006.
- [11] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
- [12] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2963–2972, Dec. 2005.
- [13] E. Binaghi, I. Gallo, and M. Pepe, "A cognitive pyramid for contextual classification of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2906–2922, Dec. 2003.
- [14] P. Mitra, C. A. Murthy, and S. K. Pal, "Unsupervised feature selection using feature similarity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 301–312, Mar. 2002.