# Classification and Extraction of Spatial Features in Urban Areas Using High-Resolution Multispectral Imagery

Xin Huang, Liangpei Zhang, and Pingxiang Li

*Abstract*—Classification and extraction of spatial features are investigated in urban areas from high spatial resolution multispectral imagery. The proposed approach consists of three steps. First, as an extension of our previous work [pixel shape index (PSI)], a structural feature set (SFS) is proposed to extract the statistical features of the direction-lines histogram. Second, some methods of dimension reduction, including independent component analysis, decision boundary feature extraction, and the similarity-index feature selection, are implemented for the proposed SFS to reduce information redundancy. Third, four classifiers, the maximum-likelihood classifier, backpropagation neural network, probability neural network based on expectation–maximization training, and support vector machine, are compared to assess SFS and other spatial feature sets. We evaluate the proposed approach on two QuickBird datasets, and the results show that the new set of reduced spatial features has better performance than the existing length–width extraction algorithm and PSI.

*Index Terms*—Feature extraction, feature selection, high spatial resolution multispectral (HSRM) imagery, spatial feature set.

## I. INTRODUCTION

CURRENTLY, commercially available high spatial resolution multispectral (HSRM) images, obtained from QuickBird, IKONOS, and SPOT-5, etc., can provide a large amount of detailed ground information in a timely manner. However, the availability of this type of data poses challenges to image classification. Due to the complex spatial arrangement and spectral heterogeneity even within the same class, conventional spectral classification methods are inadequate for HSRM imagery [1]. It is well known that combining spatial and spectral information can improve land use classification from HSRM data [2]. Therefore, many effective spatial features concerning the structure, shape, and geometric characteristics have been proposed. One commonly applied statistical procedure for interpreting texture is the gray level cooccurrence matrix (GLCM), which is a widely used texture and pattern recognition technique in the analysis of satellite data, and has been successful to a certain extent [1], [3], [4]. A method based on straight lines to assess land development in high-resolution satellite images is introduced in [5], where a set of statistical measures are extracted based on the regional line distribution. A Markov random-field-based method using both contextual information and a multiscale fuzzy line process for classifying HSRM imagery is investigated in [6]. In [7] and [8], the structural information is extracted by applying the extended morphological profiles with a multiscale approach. Some algorithms that focus on contextual spectral similarity have been proposed. In [9], a length–width extraction algorithm (LWEA) is developed to extract the length and width of spectrally similar connected groups of pixels from the HSRM imagery. These values of length and width are found by searching along a predetermined number of equally spaced lines radiating from the central pixel. In [10], a pixel shape index (PSI) is introduced, which sums the length of all the radiating lines to describe the structure around the central pixel.

In this letter, we first propose some new statistical measures to extract the structural features of direction lines, such as weighted mean ($w$-mean), length–width ratio, and standard deviation (SD), which overcome the inadequacy of the previous algorithms (PSI and LWEA). Second, some dimension reduction approaches, including feature extraction and feature selection, are tested and compared in order to reduce information redundancy. Third, different classifiers including maximum-likelihood classifier (MLC), backpropagation neural network (BPNN), probability neural network based on expectation–maximization training (EM-PNN), and support vector machine (SVM) are used to process the hybrid spectral-structural features after the steps of spatial feature extraction and dimension reduction.

## II. STRUCTURAL FEATURE SET (SFS)

PSI and LWEA provide some useful spatial measures including length, width, and sum of direction lines. In this letter, the methods of LWEA and PSI are extended with some new spatial measures of the direction lines.

### A. Direction Lines

Direction lines can be defined as a series of a predetermined number of equally spaced lines through the central pixel [9], [10]. The extension of direction lines is based on the neighboring gray level similarity and the lines radiating from the central pixel in different directions. In one direction, the spectral difference is measured between a pixel and its central pixel in

order to decide whether this pixel lies in the homogeneous area. The pixel homogeneity for a direction line is defined as

$$\mathrm{PH}(s, c) = |p(s) - p(c)| \tag{1}$$

where $\mathrm{PH}(s, c)$ represents the homogeneity measurement between the central pixel and its surrounding pixel. $p(c)$ and $p(s)$ are the spectral values of the central pixel and the surrounding pixel, respectively. Two parameters (T1 and T2) control the extension of each direction line as follows.

1) $\mathrm{PH}(s, c)$ is less than a predefined threshold T1.
2) The total number of pixels in one direction line is less than another predefined threshold T2.

The extension of one direction line will cease if either of the above conditions is not met. In that case, the extension of the $i$th direction line will be terminated, and the algorithm will skip to search the $(i + 1)$th direction line. After determining all the direction lines around one central pixel, their lengths are calculated as follows:

$$d_i = \left((m^{e1} - m^{e2})^2 + (n^{e1} - n^{e2})^2\right)^{\frac{1}{2}} \tag{2}$$

where $(m^{e1}, n^{e1})$ denote the row and column number of the pixel at one end of the direction line, and $(m^{e2}, n^{e2})$ denote the row and column number of the pixel at the other end. The parameters T1 and T2 indicate the maximum spectral and spatial distances, respectively, between the central pixel and its surrounding pixels. Theoretically, T1 will be set to between 2.5 to 4.0 times the average SD of the Euclidean distance of the training pixel data from the class means [9]. T2 is a scale factor in the algorithm and it is relevant to the image resolution and the size of the interesting objects. Given a larger T2 is proper for multiscale objects [10]; it is set as 0.5 to 0.6 times the number of rows or columns for the image.

Upon completion of the calculation of the lengths of all direction lines, the direction-line histogram of one central pixel can be defined as follows, where $I$ represents the whole image, and $c$ denotes a central pixel:

$$H(c) : \left\{c \in I \middle| [d_1(c), \ldots, d_i(c), \ldots, d_D(c)]\right\}.$$

*B. Statistical Features for the Direction Lines*

Some statistical measures will be employed to reduce the dimensionality and extract the features from the direction-lines histogram. Six statistical measures are proposed for the histogram, called the SFS.

1) Length: the maximum in the histogram.

$$\mathrm{length} = \max_{i=1}^{D} (d_i(c)) \tag{3}$$

2) Width: the minimum in the histogram.
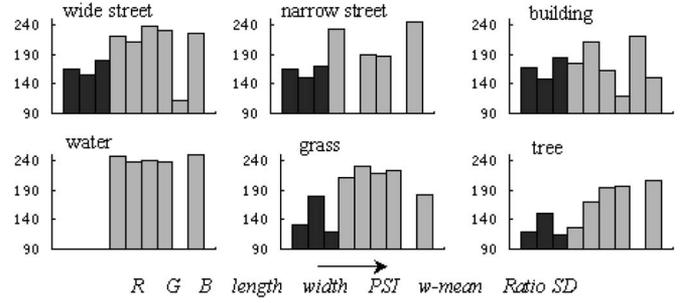
$$\mathrm{width} = \min_{i=1}^{D} (d_i(c)) \tag{4}$$



Fig. 1. Spectral and spatial features for different information classes.

3) PSI: the mean or sum of the histogram [10].

$$\mathrm{PSI} = \sum_{i=1}^{D} d_i(c) \middle/ D \tag{5}$$

4) $w$-mean:

$$w\text{-mean} = \sum_{i=1}^{D} \frac{a \cdot (k_i - 1)}{\mathrm{st}_i} d_i(c) \middle/ D \tag{6}$$

$k_i$ is the length of the $i$th direction line for a central pixel, $a$ is a constant that adjusts the value of $w$-mean, and $\mathrm{st}_i$ represents the SD of the spectral values in the $i$th direction line. A larger value of $\mathrm{st}_i$ shows that in this direction line the spectral values fluctuate a lot, and hence it is used as weight in order to reduce the length of unstable direction lines.

5) Ratio: It aims to describe the shape contour around the central pixel. Sorting the lengths of all the direction lines around one central pixel in sequence, the $n$th maximum and the $n$th minimum are represented as $\mathrm{sort}_{\max}^{n}(H(c))$ and $\mathrm{sort}_{\min}^{n}(H(c))$, respectively. Then, the ratio can be defined as (in this letter, $n = 5$, $D = 20$)

$$\mathrm{Ratio} = \arctan \frac{\sum_{j=1}^{n} \mathrm{sort}_{\min}^{j}(H(c))}{\sum_{j=1}^{n} \mathrm{sort}_{\max}^{j}(H(c))}. \tag{7}$$

6) SD: It is meaningful when one considers that a narrow but elongated rectilinear object (for instance a road) could lead to the same length or PSI values as a square feature (for instance grass or a building)

$$\mathrm{SD} = \frac{1}{D-1} \sqrt{\sum_{i=1}^{D} (d_i(c) - \mathrm{PSI})^2}. \tag{8}$$

Examples of spectral information and spatial measures for different classes are given in Fig. 1. All the features have been normalized to [0, 255] beforehand using the method in [10]. From Fig. 1, it can be clearly seen that different characteristics are achieved for the different classes. Therefore, the use of these measures should help in classification.

## III. DIMENSION REDUCTION FOR THE SPATIAL FEATURES

Conventionally, the first spectral principal component of multispectral data is extracted to build a set of spatial features [2], [8]. But in this letter, all the spectral bands are used to extract the spatial information, after which an unsupervised feature selection approach using the feature similarity index (S-Index) [11], which is a fast and effective algorithm, is employed to select the optimal subset.

S-Index is based on measuring similarity between features whereby redundancy therein is removed. This method does not need any searching and therefore is fast. A new feature similarity measure, called maximum information compression index (MICI), is introduced in [11]. Let $\Sigma$ be the covariance matrix of random variables $x$ and $y$. Define MICI as $\lambda =$ smallest eigenvalue of $\Sigma$, i.e.,

$$2\lambda(x,y) = \mathrm{var}(x) + \mathrm{var}(y)$$
$$- \sqrt{(\mathrm{var}(x) + \mathrm{var}(y))^2 - 4\mathrm{var}(x)\mathrm{var}(y)(1 - \rho(x,y)^2)} \quad (9)$$

and

$$\rho(x,y) = \mathrm{cov}(x,y) \Big/ \sqrt{\mathrm{var}(x)\mathrm{var}(y)} \quad (10)$$

where $\mathrm{var}(\cdot)$ denotes the variance of a variable and $\mathrm{cov}(\cdot)$ the covariance between two variables. The value of $\lambda$ is zero when the features are linearly dependent and increases as the amount of dependence decreases.

It has been proved that this index outperforms some conventional approaches to feature selection, such as branch and bound, sequential forward search, sequential floating forward search, and stepwise clustering, both in accuracy and CPU time [11]. See [11] for a detailed description of the S-Index. For the purpose of comparison, the methods of decision boundary feature extraction (DBFE) [8] and independent component analysis (ICA) [12] are used to achieve the spatial feature extraction for the HSRM data.

## IV. EXPERIMENTS

In this letter, the SVM is employed for the classification of spectral-structural features. At the same time, the classifiers of BPNN, MLC, and PNN based on EM training [13] will also be tested for the purpose of comparison. The one-against-one method is employed to resolve the multiclass problem of SVM and the leave-one-out algorithm is used to select the kernel parameters for the radial basis function [10]. The configuration for BPNN is: the number of input neurons equals the number of input features, the number of hidden layer neurons equals twice the number of input layer neurons, and the stopping rule of the training phase equals 1000 epochs. Two statistics, overall accuracy (OA) and Kappa coefficient based on confusion matrix, are utilized to evaluate the classification performances.

The QuickBird image used in experiment (1) consists of three multispectral bands (RGB) with 2.44-m resolution. The image is shown in Fig. 2(a), and a reference image is available
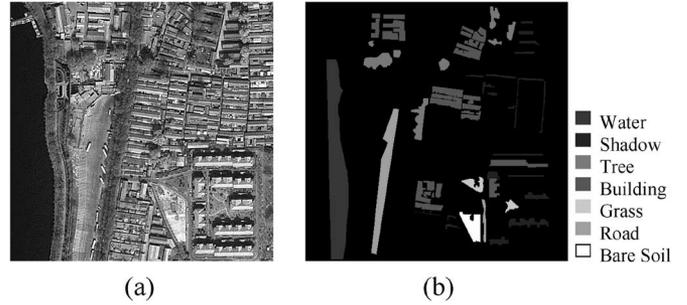


Fig. 2. Experimental area and reference image. (a) Test area of QuickBird image. (b) Reference image.
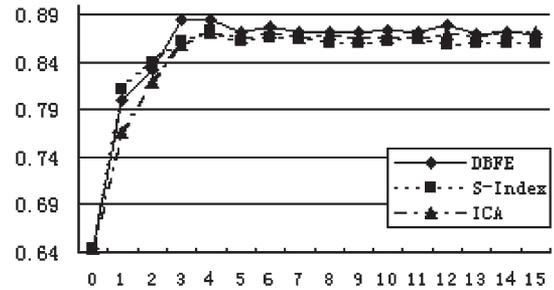


Fig. 3. Accuracies statistics for dimension reduction methods in experiment (1).

[Fig. 2(b)]. The numbers of training and testing samples were selected randomly from the reference data.

*1) Comparison of DBFE, ICA, and S-Index:* In this experiment, 6-D SFS measures and three bands produce 18-D spatial features, which will be preprocessed by a normalization step beforehand [10]. The dimensionally reduced structural features were integrated with spectral information to achieve the classification. The parameters of the direction lines were set as T1 = 110 and T2 = 175, and the SVM classifier was used. The OAs of classification for test data as a function of the number of dimensionally reduced features for the different approaches are shown in Fig. 3. From the figure, it can be seen that the tendency of the three curves are significantly similar, and the highest classification accuracies are achieved when three or four spatial features are included in the feature sets. When the number of spatial features increases from 0 to 3 (or 4), the three curves ascend obviously to their respective summits. But when 6-D SFS features are included, the overall accuracies do not improve much, which demonstrates that more spatial features do not necessarily produce higher accuracy. The highest accuracy of 0.885 is achieved by 3-D features using DBFE; the highest accuracies by S-Index and ICA are 0.873 and 0.871, respectively. Considering that the S-Index is just a method of feature selection instead of linear combination and its computational cost is essentially low, the S-Index is successfully applied for the SFS features.

*2) Comparison of GLCM, LWEA, PSI, and SFS:* The textural measures adopted on GLCM include: mean, variance, contrast, dissimilarity, entropy, and homogeneity [4]. These measures were calculated setting the interpixel distance equal to one in all the four directions and finally the directionality

TABLE I
CLASSIFICATION ACCURACIES FOR DIFFERENT SPATIAL FEATURES IN
EXPERIMENT 1), AND SVM WITH 3-D SPATIAL FEATURES
SELECTED VIA S-INDEX ARE USED

| Class | RGB | GLCM | LWEA | PSI | SFS |
|---|---|---|---|---|---|
| Water | 0.983 | 0.974 | 0.949 | 0.902 | 0.971 |
| Grass | 0.668 | 0.708 | 0.839 | 0.739 | 0.878 |
| Tree | 0.401 | 0.391 | 0.446 | 0.429 | 0.481 |
| Building | 0.709 | 0.682 | 0.678 | 0.770 | 0.866 |
| Road | 0.430 | 0.710 | 0.701 | 0.675 | 0.760 |
| Bare-Soil | 0.896 | 0.884 | 0.853 | 0.772 | 0.862 |
| Shadow | 0 | 0.005 | 0.976 | 0.985 | 0.971 |
| OA | 0.643 | 0.660 | 0.796 | 0.801 | 0.863 |
| Kappa | 0.540 | 0.567 | 0.745 | 0.750 | 0.836 |



Fig. 5. Accuracy statistics for different classifiers with different input features in experiment (1).
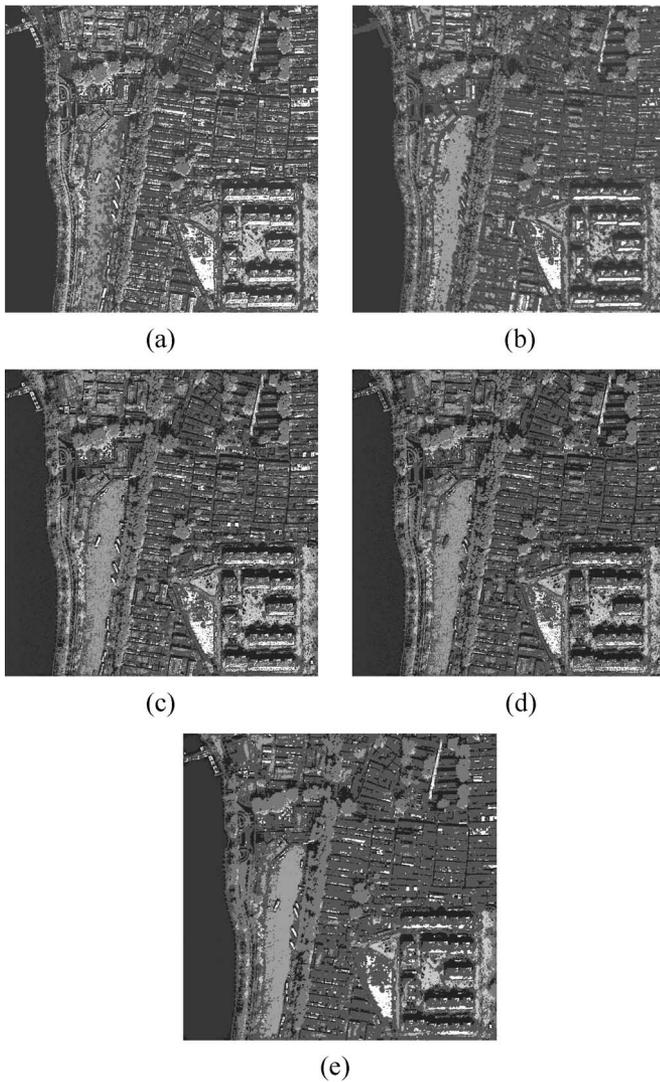


(a)

(b)

(c)

(d)

(e)

Fig. 4. Classification maps for different features in experiment (1). (a)–(e) Classification maps for spectral, GLCM, LWEA, PSI, and SFS features, respectively.

was suppressed by averaging the extracted features over four directions. The S-Index method and SVM classifier were used for this comparison and the dimension of all the different spatial feature sets after feature selection is three. Their classification accuracies with different information classes are provided in Table I and the classification maps for all five feature sets are shown in Fig. 4.
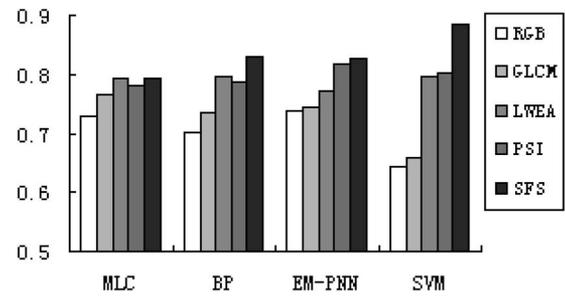
In Fig. 4(a), the obvious misclassifications lie in spectrally similar objects such as tree–grass, building–road–bare-soil, and shadow–water, which illustrates that spectral information is inadequate for the classification of HSRM data. GLCM [Fig. 4(b)] improves the classification of tree–grass slightly and the accuracy of road acquires an obvious increase of 28%; however, it cannot discriminate water–shadow, which is partly attributed to its fixed window processing. It is found that PSI and LWEA achieve similar results, both outperforming GLCM. SFS is an improved version of LWEA and PSI, and it not only combines the spatial measures of them but also introduces some new measures to overcome their defects. Compared with PSI, the improvements for SFS in OA and Kappa coefficient are about 6.2% and 8.6%, respectively.

*3) Comparison of MLC, BP, EM-PNN, and SVM:* For different classifiers, the dimension of all the spatial features after feature selection (S-Index) was set equal to 3. In Fig. 5, the accuracies for four classifiers with five different feature sets are provided, from which it can be observed that the proposed SFS features give better performance independent of the classifiers used.

In order to verify that the proposed algorithms of classification and spatial feature extraction work in a stable manner, another QuickBird multispectral image of Beijing was examined, whose size is 630 × 677 pixels [Fig. 6(a)]. A reference image with different information classes was available [Fig. 6(b)].

In this experiment, the parameters of the direction lines were set as T1 = 90, T2 = 210. The relationship between OA and the number of spatial features for the different feature extraction methods is shown in Fig. 7. The highest overall accuracies achieved by DBFE, S-Index, and ICA are, respectively, 0.832, 0.819, and 0.813. It can be concluded that the S-Index is fast and effective in extracting the SFS features of HSRM data, although, theoretically, the optimal set of features may be the result of a linear or nonlinear transformation of the original features.

Different feature sets were compared again using the SVM classifier and S-index method. Five spatial features were selected from the original feature sets. The classification accuracies with seven information classes for different features are provided in Table II; the classification maps are shown in Fig. 6(c) and (d). Compared with LWEA and PSI, the respective 5.5% and 5.3% improvements of OA via SFS features show the strength of our algorithm, especially in the classification of bare soil, road, and building.
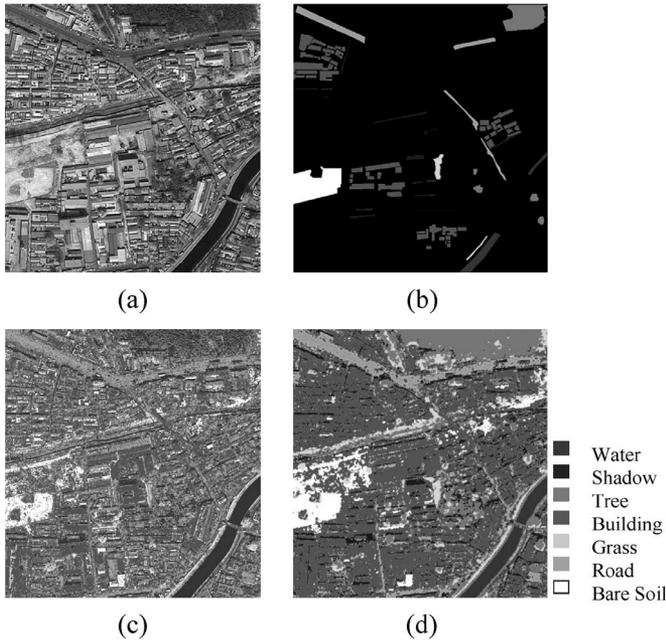
Fig. 6. Classification maps for different features in experiment (2). (a) Image in the experimental area. (b) Reference image. (c)–(d) Classification maps for spectral information and SFS feature set, respectively.
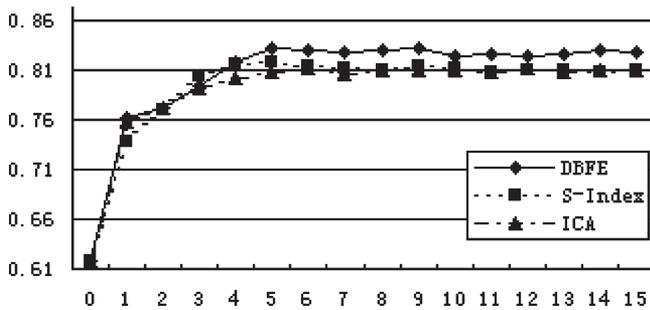


Fig. 7. Accuracy statistics for dimension reduction methods in experiment (2).

TABLE II
CLASSIFICATION ACCURACIES FOR DIFFERENT
SPATIAL FEATURES IN EXPERIMENT (2)

| Class | RGB | GLCM | LWEA | PSI | SFS |
|---|---|---|---|---|---|
| Water | 0.938 | 0.921 | 0.876 | 0.904 | 0.898 |
| Grass | 0.675 | 0.596 | 0.629 | 0.571 | 0.651 |
| Tree | 0.372 | 0.393 | 0.759 | 0.684 | 0.748 |
| Building | 0.598 | 0.624 | 0.738 | 0.851 | 0.833 |
| Road | 0.794 | 0.813 | 0.818 | 0.649 | 0.789 |
| Bare-Soil | 0.716 | 0.681 | 0.756 | 0.728 | 0.861 |
| Shadow | 0 | 0.031 | 0.708 | 0.723 | 0.854 |
| OA | 0.617 | 0.625 | 0.763 | 0.765 | 0.818 |
| Kappa | 0.509 | 0.556 | 0.689 | 0.690 | 0.756 |

## V. CONCLUSION

This letter mainly addresses the three following items. The first one was to introduce some new spatial measures such as $w$-mean, SD, and length–width ratio. The second one was to extract the respective spatial features from each spectral band, and then to use the methods of S-Index, DBFE or ICA to obtain the structural information of multispectral bands. The last one was to compare different features in different classifiers.

The proposed SFS is an extension of previous algorithms. According to the classification maps and accuracies in experiments, the improvements brought about by SFS are apparent. Meanwhile, it should be noticed that not all the spatial measures are indispensable, especially when structural features are extracted from every spectral band. In experiments, DBFE, ICA, and S-Index were successful in reducing the high-dimensional features, but S-Index is more suitable for the SFS features, although a slightly higher accuracy is achieved by DBFE. After that, four kinds of classifiers were used to classify spectral-spatial features in order to test the adaptability of SFS features. An overall conclusion can be drawn that the proposed SFS algorithm obviously improves the previous results and SVM is better at interpreting the hybrid features.

## 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] 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.

[3] D. G. Barber and E. F. LeDrew, "SAR sea ice discrimination using texture statistic: A multivariate approach," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 62, pp. 949–958, 1991.

[4] Y. Zhang, "Optimisation of building detection in satellite images by combining multispectral classification and texture filtering," *ISPRS J. Photogramm. Remote Sens.*, vol. 54, no. 1, pp. 50–60, Feb. 1999.

[5] C. Unsalan and K. L. Boyer, "Classifying land development in high-resolution panchromatic satellite imagery using hybrid structural-multispectral features," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 12, pp. 2840–2850, Dec. 2004.

[6] B. Tso and R. C. Olsen, "A contextual classification scheme based on MRF model with improved parameter estimation and multiscale line process," *Remote Sens. Environ.*, vol. 97, no. 1, pp. 127–136, Jul. 2005.

[7] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.

[8] J. A. Benediktsson, M. Pesaresi, and K. Arnason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1940–1949, Sep. 2003.

[9] A. K. Shackelford and C. H. Davis, "A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1920–1932, Sep. 2003.

[10] L. Zhang, X. Huang, B. Huang, and P. Li, "A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2950–2961, Oct. 2006.

[11] 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.

[12] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 626–634, Mar. 1999.

[13] B. Tian, M. A. Shaikh, M. R. Azimi-Sadjadi, T. H. Vonder, and D. L. Reinke, "A study of cloud classification with neural networks using spectral and textural features," *IEEE Trans. Neural Netw.*, vol. 10, no. 1, pp. 138–151, Jan. 1999.