Classification of Very High Spatial Resolution Imagery Based on the Fusion of Edge and Multispectral Information

Xin Huang, Liangpei Zhang, and Pingxiang Li

Abstract

A new algorithm based on the fusion of edge and multispectral information is proposed for the pixel-wise classification of very high-resolution (VHR) remotely sensed imagery. It integrates the multispectral, spatial and structural information existing in the image. The edge feature is first extracted using an improved multispectral edge detection method, which takes into account the original multispectral bands, the linear NDVI, and the independent spectral components extracted by independent component analysis (ICA). Direction-lines are then defined using the edge and multispectral information. Two effective spatial measures are calculated based on the direction-lines in order to describe the contextual information and strengthen the multispectral feature space. Then, the support vector machine (SVM) is employed to classify the hybrid structural-multispectral feature set. In experiments, the proposed spatial measures were compared with the pixel shape index (PSI) and the gray level co-occurrence matrix (GLCM). The experimental results show that the proposed algorithm performs well in terms of classification accuracies and visual interpretation.

Introduction

At present, commercially available high spatial resolution multispectral images, obtained from QuickBird, Ikonos, etc., can provide a large amount of detailed ground information in a timely manner, thus opening up avenues for new remote sensing applications. However, due to the complex spatial arrangement and spectral heterogeneity even within the same class, traditional spectral methods have proven inadequate for the classification of VHR satellite imagery (Myint *et al.*, 2004). It is generally agreed that combining spectral and spatial information can improve land-use classification from satellite imagery (Dell'Acqua *et al.*, 2004). Therefore, in recent years, many algorithms have been proposed to extract the spatial features, complement the spectral information and improve the pixel-wise classification.

One commonly applied statistical procedure is the gray level co-occurrence 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 (Zhang, 1999; Gong *et al.*, 1992; Barber *et al.*, 1991). A method based on straight lines to assess land development in high-resolution satellite images was introduced and a set of statistical measures were extracted based on the sub-windows showing the regional line distribution (Unsalan *et al.*, 2004). An edge detection method integrating a region-growing approach was used to improve classification for the images of the Indian Remote Sensing Satellite 1C (Sun *et al.*, 2003). Benediktsson *et al.* (2003 and 2005) presented a technique of extended morphological profiles to describe the multi-scale spatial features and to interpret high spatial resolution remote sensing data. Yu (2005) proposed the idea of analyzing lines emanating from a point of interest for segmentation and classification in computer vision. Zhang *et al.* (2006) proposed a pixel shape index (PSI), which extracted structural features based on contextual spectral similarity. The PSI was calculated by a predetermined number of equally spaced lines radiating from the central pixel (called direction-lines).

Based upon the aforementioned work, this paper proposes a methodology that integrates edge and multispectral information to improve the classification accuracy. The flow for the whole process is shown in Figure 1. The salient aspects of our strategy are the following:

1. An improved method of multispectral edge detection is proposed. The original multispectral bands, the linear version of the normalized difference vegetation index (NDVI) (Unsalan *et al.*, 2004), and the spectral independent components extracted by ICA, are combined to produce the edge information. The edge map is a fuzzy image, where each pixel is represented by a value indicating the number of times it has been detected as an edge pixel.
2. The direction-lines are determined based on the fuzzy edge image and contextual spectral similarity in the original multispectral bands. A decision rule is proposed to fuse the edge and multispectral information when direction-lines are being extended. Two effective spatial measures are then proposed to compute the statistical characteristics of the direction-lines for each pixel, called mean and length-width ratio.
3. The proposed structural features are integrated with the multispectral information using the support vector machine (SVM), a relatively new method of machine learning. The notable advantages of SVM include self-adaptability, swift learning pace, and high-dimensional property in feature space. SVM can also reduce the dominance effects of the spectral

The State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, P.R. China (huang_whu@163.com; zlp62@public.wh.hb.cn).

Photogrammetric Engineering & Remote Sensing
Vol. 74, No. 12, December 2008, pp. 1585–1596.

0099-1112/08/7412–1585/$3.00/0

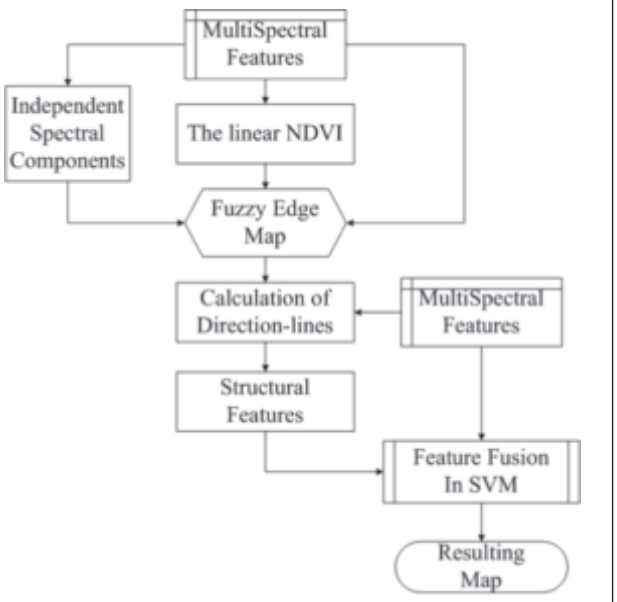

Figure 1. The flow chart for the proposed classification algorithm integrating edge and multispectral information.

information if the input vector contains both spectral and spatial information. This SVM-based classification approach is chosen because of its fast and effective processing.

The rest of this paper is organized as follows. The next section provides the detailed description of the multispectral edge detection method followed by a description of the calculation of direction-lines. The next section discusses the fusion of spectral and structural features based on SVM followed by two experiments to evaluate the performance of the proposed methodology. The last section concludes the paper.

Multispectral Edge Detection

In the procedure discussed in this paper, the algorithm for extracting multispectral edge information is divided into two parts. The first part concerns the information sources for edge detection, including the independent component analysis (ICA), a linear version of NDVI (Unsalan, 2004), and the original multispectral images. The second step describes the edge filtering process and the resulting fuzzy edge image. The establishment of the edge image is illustrated in Figure 2.

Information Sources for Edge Detection

Independent Component Analysis (ICA)

ICA is a multivariate data analysis method for blind source separation of signals. It seeks to render the components as statistically independent as possible, and is a useful tool for data mining of remotely sensed data. The ICA features of the multispectral images are extracted to acquire more information sources for the edge detection. The basic model of ICA is:

$$\boldsymbol{s} = \boldsymbol{W}\,\boldsymbol{q} \tag{1}$$

where $\boldsymbol{s}$ is the vector of observed signals, $\boldsymbol{W}$ is the mixing matrix to be estimated; and $\boldsymbol{q}$ is the mutually independent components.

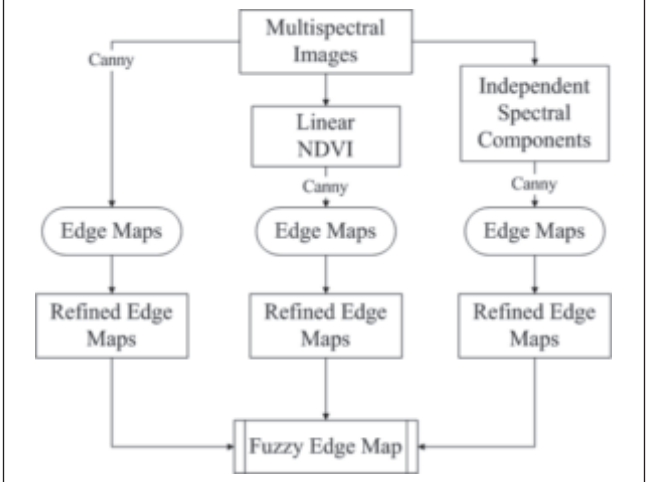

Figure 2. The flow chart for multispectral edge detection based on multiple information sources.

The goal of ICA is to calculate the matrix $\boldsymbol{A}$ such that the sources $\boldsymbol{q} = \boldsymbol{As}$ can be estimated from the observed vector $\boldsymbol{s}$ by optimizing the statistical independence criterion (actually, $\boldsymbol{A}$ is the inverse of the matrix $\boldsymbol{W}$). Starting from an information theoretic viewpoint, the ICA problem is formulated as the minimization of mutual information between the transformed variables, since mutual information is a natural measure of the dependence between random variables. In our study, the Fast-ICA algorithm (Hyvarinen, 1999) is adopted because it presents fast convergence and sequential extraction of the sources. Fast-ICA is a fixed point algorithm based on optimization of the negative entropy (negentropy) function, and the negentropy is used as a measure of mutual information. It can be shown that finding a transformation $\boldsymbol{W}$ that minimizes the mutual information is equivalent to finding directions in which the negentropy is maximized. Readers can refer to Hyvarinen (1999) for more details about Fast-ICA.

The Linear NDVI

In order to strengthen the multispectral information, especially for vegetation areas, the well-known normalized difference vegetation index (NDVI) is used. However, as the degree of vegetation increases, the corresponding NDVI values begin to saturate and cannot represent highly vegetated regions reliably. Therefore, in this study, the linear version NDVI (Unsalan *et al.*, 2004) is used as one of spectral information sources for edge detection. It can be calculated as:

$$\theta = \frac{4}{\pi}\arctan(NDVI) \tag{2}$$

The most significant aspect of the linear NDVI (θ) is that it is highly linear with respect to vegetation density, avoiding the saturation problem found using the NDVI directly.

Edge Detector

In this paper, the following steps describe edge detection for multiple information sources:

Step 1

The linear NDVI image and the independent spectral components are calculated, and they are then combined with the original spectral bands for subsequent processing.

Step 2
Edge detection for the resulting three information sources is performed using the Canny filter; it is considered an optimal edge detector, having a low probability of false or missing edges and a high accuracy of edge positioning (Solaiman *et al.*, 1998);

Step 3
After applying the Canny detector, all the information sources are transformed to binary edge images:

$$\begin{cases} edge_b[i, j] = 1 & \text{If pixel } (i, j) \text{ in band } b \text{ is an edge} \\ edge_b[i, j] = 0 & \text{If pixel } (i, j) \text{ in band } b \text{ is a non-edge} \end{cases}. \quad (3)$$

where $edge_b[i, j]$ denotes the edge value of pixel (i, j) in the band $b(1 \le b \le N)$, and N is the number of the information sources. For multispectral edge detection, a pixel may be considered as an edge pixel in one feature band but as belonging to a homogeneous region in another band. The information of ICA and linear NDVI are included here because performance can be improved by combining edge maps that utilize information from multiple sources;

Step 4
A simple post-processing is then used to remove the noise in all the edge information bands. Given that $count_b[i,j]$ represents the number of edge pixels in the 8-neighborhood of pixel (i,j) in the band b, the isolated edge points can be deleted by:

$$\text{If } edge_b[i, j] = 1 \text{ and } count_b[i, j] = 0$$
$$\text{Let } edge_b[i, j] = 0. \quad (4)$$

Step 5
Each individual feature image is filtered, and therefore N binary edge maps are obtained. The fuzzy edge image is obtained by superimposing all the N edge maps. Each pixel in this image is represented by a value ranging from zero to N, indicating the number of times it has been detected as an edge pixel:

$$edge[i, j] = \sum_{b=1}^{N} edge_b[i, j] \quad (5)$$

where $edge[i, j]$ is the value of pixel (i, j) in the fuzzy edge map.

Structural Feature Extraction

Pixel Shape Index (PSI)

Zhang *et al.* (2006) proposed a pixel shape index (PSI) to examine the context of each pixel and to measure the spatial dimensions of groups of spectrally similar connected pixels. It consists of two steps:

Extension of Direction-lines
The direction-lines used to detect the object's overall contour are symmetric around the central pixel, and are illustrated in Figure 3. In one chosen direction, the spectral difference is measured between a pixel and its central pixel in order to decide if this pixel lies in the homogeneous area around the central pixel.

The pixel homogeneity is defined using the following method:

$$PH_d(k, x) = \sum_{s=1}^{n} |p_s(k) - p_s(x)| \quad (6)$$

where $PH_d(k, x)$ represents the spectral homogeneity of the d^{th} direction between the central pixel k and its surrounding pixel x, n denotes the number of spectral bands, $p_s(k)$ is the spectral value of the central pixel, and $p_s(x)$ is the spectral value of the surrounding pixel, $d = 1, 2, \ldots D$; D is the total number of direction-lines and is invariable in this paper (D = 20). In Equation 6, smaller values of $PH_d(k,x)$ mean increased homogeneity around the central pixel.

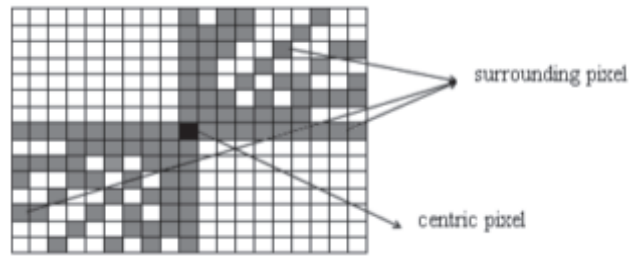

Figure 3. Direction-lines around the central pixel.

The d^{th} direction-line is extended if the following conditions are met at the same time: (a) $PH_d(k,x)$ is less than a pre-defined threshold T1, and (b) the total number of pixels in this direction-line is less than another pre-defined threshold T2. The conditions can be described as:

$$L_d(k) \le T2 \quad \text{and} \quad PH_d(k, x) \le T1 \quad (7)$$

where $L_d(k)$ is the length of the d^{th} direction-line for the central pixel k.

Calculation of PSI
The shape index of the central pixel is the sum of lengths for all the direction lines, which is defined as:

$$PSI(k) = \sum_{d=1}^{D} L_d(k) \quad (8)$$

where $PSI(k)$ represents the shape index of the central pixel k. There are two parameters for PSI: the direction-line extension thresholds T1 and T2. The values T1 and T2 ought to be determined through experiments as these two parameters are related to the shape and spatial arrangement of objects in the image.

Structural Feature Extraction based on PSI and Edge Information

PSI sums the lengths of all the direction-lines as the shape feature value for each pixel. In this study, the PSI algorithm is extended by taking the fuzzy edge information into account, and at the same time two spatial measures, called mean and length-width ratio, are presented. The proposed algorithm for structural feature extraction consists of two steps:

Extension of Direction-lines based on Edge and Spectral Features
The pixel homogeneity in Equation 6 is modified as:

$$PH_d(k, x) = \sqrt{\sum_{s=1}^{n} (p_s(k) - p_s(x))^2}. \quad (9)$$

The city-block distance is replaced by the Euclidean norm, because the former is a simplified distance version. Equation 6 and Equation 9 consider only the multispectral features around the central pixel, and in order to integrate the edge and multispectral information, Equation 9 can be modified to:

$$PH_d'(k, x) = \left(1 + r\frac{edge[x]}{N}\right) PH_d(k, x) \quad (10)$$

where r is a ratio adjusting parameter ($r = 1$ in this paper), and $edge[x]$ denotes the value of pixel x in the fuzzy edge map. It can be seen that when $edge[x] = 0$, Equations 9 and 10 possess the same form. In Equation 10, the value of $\frac{edge[x]}{N}$ represents the estimated probability that pixel x is a potential edge point. When this value increases, the likelihood that the pixel lies in the homogeneous area along the d^{th} direction-line becomes smaller. Therefore, the conditions for extension of direction lines can be modified as:

$$edge[x] \le \lambda N \quad \text{and} \quad PH'_d(k, x) \le T \text{ with } 0 \le \lambda \le 1. \tag{11}$$

The equation means: (a) if $edge[x] > \lambda N$ ($\lambda = 0.7$ in this paper), it may show an obvious discontinuity or a strong spectral difference in the local area and hence the extension of direction-lines should be terminated there, and (b) if $edge[x] \le \lambda N$, Equation 10 is then used to decide whether it should be accepted by its direction-line, where the fuzzy edge values are combined with spectral difference for determining direction lines. The determination of direction lines can be described as follows:

for each pixel k

for direction-line $d = 1$ to D

$L_d(k) = 1$; //the initializing value

while $edge[x] \le \lambda N$ and $PH'_d(k, x) \le T$

$L_d(k) = L_d(k) + 1$;

$x = x + 1$;//go to the next neighboring pixel along the d^{th} direction-line

endwhile

//the extension of the d^{th} direction-line will stop, and the length of this direction-line is stored as $L_d(k)$;

endfor

$k = k + 1$ //go to the next central pixel

endfor

Both Equation 7 and Equation 11 consider the spectral difference and spatial restraint around the central pixel. Comparing their methods of spatial restraint, $edge[x] \le \lambda N$ in Equation 11 and $L_d(k) \le T2$ in Equation 7, it can be found that the edge features are used to limit the extension of direction lines automatically in Equation 11; however, the length for a direction line is limited by a pre-defined threshold T2, and T2 is invariable for each pixel. It is more reasonable for Equation 11 since its spatial extensions of direction lines are based on the spatial features in a local area. Comparing their measurements of spectral difference, $PH_d(k, x)$ and $PH'_d(k, x)$, the fuzzy edge values are included in the latter, while only the spectral features in the local area are considered for the former.

Statistics of Structural Features based on the Direction-lines

Upon completion of the calculation of the lengths of all direction-lines, the direction-lines histogram for central pixel k can be defined as follows, where I represents the whole image.

$$H(k|k \in I)\text{: } \{L_1(k), L_2(k), \cdots L_d(k), \cdots L_D(k)\}. \tag{12}$$

Two statistical measures are proposed to extract the significant features from the histogram:

1. Mean: the mean of the histogram. It aims to describe the overall neighboring features along all the directions and to let all the pixels in the same shape area possess the same feature value:

$$mean(k) = \frac{1}{D}\sum_{d=1}^{D} L_d(k). \tag{13}$$

2. Length-width Ratio: Sorting the lengths of all the direction-lines around one central pixel in sequence, if the e^{th} maximum and the e^{th} minimum are represented as $\underset{\max}{sort^e}[H(k)]$ and $\underset{\min}{sort^e}[H(k)]$, respectively, the ratio can be defined as:

$$ratio(k) = \arctan\frac{\sum_{t=1}^{e} \underset{\min}{sort^t}[H(k)]}{\sum_{t=1}^{e} \underset{\max}{sort^t}[H(k)]}. \tag{14}$$

The length-width ratio can be used to complement the spatial measure of mean. It is meaningful when one considers that a narrow but elongated rectilinear object (for instance a road) could lead to the same mean values as a square feature (for instance an area of grass or a building).

Support Vector Machine for Classification of Spectral-Spatial Features

Conventional classifiers, such as the maximum likelihood classifier (MLC), are not capable of achieving a satisfactory accuracy for VHR data. This is because the estimated distribution function usually employs the normal distribution, which may not represent the actual distribution of the feature space. In recognition of this, the fuzzy ARTMAP (Dell'Acqua *et al.*, 2004), BP neural networks (Benedikttson *et al.*, 2005), probabilistic relaxation (Unsalan and Boyer, 2004), and probability neural networks (Tian *et al.*, 1999) have been tested to integrate the spectral and spatial features. In this paper, SVM is employed to interpret the hybrid structural and multispectral features owing to its computational simplicity and superior accuracy when compared to other classifiers (Foody and Mather, 2004; Melgani and Bruzzone, 2004). Two problems need to be resolved in the fusion process: normalization prior to classification and parameter selection of SVM.

1. The spatial and spectral features are normalized into [0, 255] in such a way that they can then be processed effectively in different classifiers. The normalization method for spectral features differs from that for spatial ones because of their dissimilar distribution range. The normalization methods can be formulated as follows:

$$\begin{cases} d'_k = \dfrac{d_k - d_{\min}}{d_{\max} - d_{\min}} \cdot 255 & \text{spectral features} \\ \textit{histogram–equalization} & \text{structural features} \end{cases} \tag{15}$$

where d_k denotes the original value of the pixel k, and accordingly $d_{\max}$, $d_{\min}$ represent the maximum and minimum values in the image, and d'_k is the feature value after the normalization. It should be noted that the histogram equalization approach is used for structural features because it can increase the contrast of different objects more effectively than the linear stretch method (Zhang *et al.*, 2006).

2. The second problem relates to the parameters of SVM. SVM classifiers of the form $f(a) = w \cdot \Phi(a) + h$ learn from the data $\{(a_i, y_i) \mid a_i \in R^m, y_i \in \{-1, 1\}, i = 1, \ldots, m\}$, where a_i is an m-dimensional feature vector, $f(a)$ denotes a hyperplane that separates samples label y_i on each side, w (weight vector) and h (bias term) are the parameters of the hyperplane (Melgani and Bruzzone, 2004). The hyperplane calculation can be formulated into a constrained optimization problem as follows:

$$\min_{w,h,\xi_i}\left\{\frac{1}{2}||W||^2 + C\sum_{i=1}^{m}\xi_i\right\} \text{ subject to } y_i(w\cdot\Phi(a_i) + h) \geq 1 - \xi_i, \xi_i \geq 0 \quad (16)$$

where C is a regularization parameter, and ξ_i is a slack-variable. Equation 16 is usually solved through its Lagrangian dual problem (Camps-Valls and Bruzzone, 2005). A specific application for SVM needs to handle several issues:

Multiclass Problem

While SVM was originally designed for binary classification, most remote sensing applications involve multiple classes. Some strategies have been proposed for extending SVM to multiclass classifications (Foody and Mather, 2004). In this paper, the one-against-one (OAO) method is employed, where a series of classifiers are applied to each pair of classes, with the most commonly computed class label reserved for each pixel (Foody and Mather, 2004).

Selection of Kernel Function

The commonly used kernel functions are the Gaussian radial basis function (RBF) and the polynomial function (POLY):

$$\begin{cases} poly\text{: } K(x_i, x_j) = (x_i \cdot x_j + 1)^p \\ Rbf\text{: } K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2) \end{cases} \quad (17)$$

For pattern recognition in VHR satellite images, the POLY kernel was found to be better than RBF, because the POLY kernel is a type of function with overall influence, whereas RBF mainly responds to local points around the central value (Zhang *et al.*, 2006). In this study, the POLY kernel is used.

Parameter Optimization

In addition, the kernel based implementation of SVM involves problems pertaining to the selection of multiple parameters, including the kernel parameters (p) and the regularization parameter C. Some standard methods exist that can facilitate the selection of parameters in the SVM classifier design. In our method, these parameters were selected automatically based on the LOO (leave-one-out) algorithm (Chapelle *et al.*, 2002). It is based on the idea of estimating the parameters values so that the estimate of the expected generalization error is minimized. Optimization of the parameters is carried out using a gradient descent search over the space of the parameters.

Experiment

The proposed algorithm was applied and tested on two VHR datasets. In the experiments, the pixel-wise classifications with different spectral and spatial features were evaluated. Some typical information classes in the experiment areas are: roofs (buildings), roads, trails, water, shadow, bare soil, trees, and grass. The proposed structural features extracted by edge and spectral information (ES features) were compared with those extracted by GLCM and PSI. Multispectral information was integrated with spatial features (co-occurrence approaches, shape index, edge features) in every case. At the same time, in addition to the SVM, different classifiers including the maximum likelihood classifier (MLC), backpropagation neural network (BPNN), and probability neural network (PNN) based on EM training (Tian *et al.*, 1999) were employed for different input features in order to arrive at an overall conclusion.

Experiment on HYDICE Data for Washington, D.C.

This experiment analyzed the Hyperspectral Digital Imagery Collection Experiment (HYDICE) airborne hyperspectral data flight line over the Washington, D.C. Mall (Landgrebe, 2003). Two hundred and ten bands were collected in the 0.4 to 2.4 um region of the visible and infrared spectrum. The water absorption bands were then deleted, resulting in 189 channels. The analysis of hyperspectral imagery with high spatial resolution is beyond the scope of this paper; therefore, in this experiment, bands 60, 27, 17, and 72 are used for the red, green, blue, and infrared channels, respectively. The test image comprised 1,280 lines and 307 columns, and is shown in Plate 1a, with the reference data available from Plate 1b. From the reference data, the training and testing samples were selected randomly (see Table 1). As can be seen from Table 1, only 70 samples per class were employed to train the classifier. All other samples were used to test the classification results.

The independent spectral components and the fuzzy edge map produced by the multiple information sources are shown in Plate 1c and 1d, respectively. In the experiments, the value of standard deviation for the Canny filter is set to 0.8, because it was found that smaller values can reduce some additional edges and noise. Plate 1d indicates the number of source bands in which the pixels have been detected as edges. Bright colors indicate strong edges, and dark colors represent weak ones.

The experiment aims to compare the performance of spectral, GLCM, PSI, and ES features for classification of VHR satellite imagery. To allow for fair testing, a similar parameter optimization strategy was used for all the features. The textural measures adopted in GLCM include variance and contrast (Zhang, 1999). These measures were calculated by setting the inter-pixel distance equal to one in all four directions, and finally the directionality was suppressed by averaging the extracted features over four directions. In experiments, the GLCM features were computed using the first principle component with 64 gray levels. As the window size is an important factor for spatial feature analysis, the GLCM measures are calculated using two moving windows, 3×3 and 7×7, respectively. The parameters of PSI were set as: T1 = 120 and T2 = 50; T was set as 120 in the ES algorithm. In the experiment, the multispectral information was integrated with spatial features in SVM and all the input features were normalized beforehand using Equation 15. Their classification accuracies with different information classes are provided in Table 2, and the classification maps are shown in Plate 2. In Table 2, AA represents average accuracy for all the information classes, OA (overall accuracy), Kappa, PA (Producer's Accuracy), and UA (User's Accuracy) are the statistics of the confusion matrix.

In Table 2, the numbers in gray represent the highest accuracies among all the features. This table shows that the proposed ES features achieve the highest PA, except for the classes of roads, trails and trees, and the highest UA, except for roads and roofs. For classification of VHR satellite imagery, the spectral information alone cannot distinguish the spectrally similar objects effectively. This point can be verified in Table 3, which shows the confusion matrix for the classification results of spectral features and ES features. From the table, it can be observed that spectral information cannot discriminate trails-roads-roofs, grass-trees, and water-shadow effectively; however, the ES features overcome the mis-classifications and obviously improved the results. In the experiment, the proposed ES features performed best in average accuracy, overall accuracy and Kappa coefficient. Compared with spectral, GLCM (3×3), GLCM (7×7), and PSI, the improvements for ES are about 11.0 percent, 9.5 percent, 7.8 percent, and 4.1 percent in OA, and 13.1 percent, 10.6 percent, 9.5 percent, and 5.3 percent in Kappa, respectively. Most of the best classification results for different objects were achieved by ES.

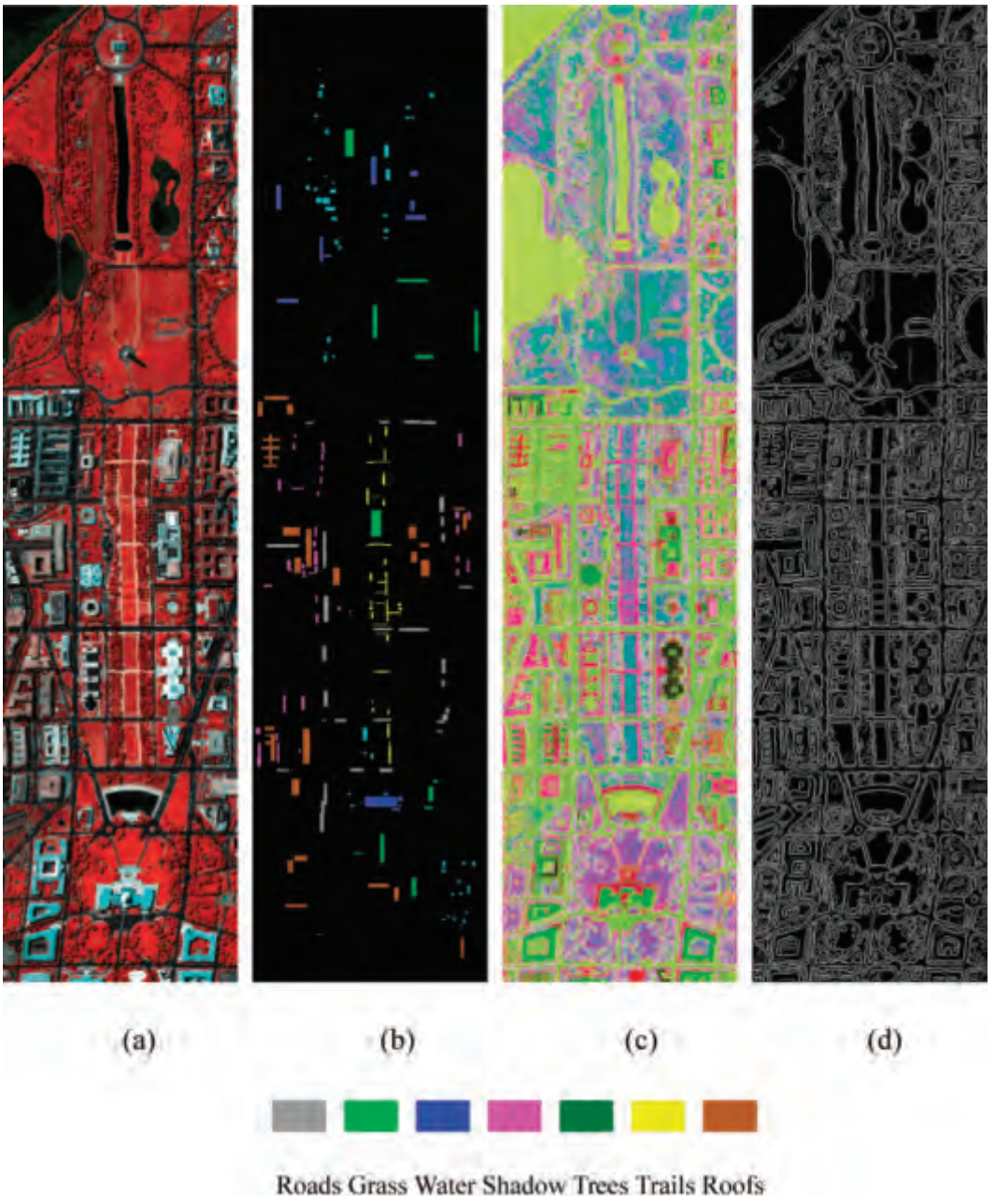

Plate 1. Washington, D.C. data: (a) the three-channel RGB color composite of the area, and (b) the reference data; (c) the false color image with the first three independent spectral components as red, green, and blue bands, and (d) the fuzzy edge map produced by all the available information sources.

By combining the classification map in Plate 2a and the accuracy statistics in Tables 2 and 3, we can observe that many water pixels are wrongly labeled as shadow in the spectral classification map because of the lack of spatial features, while mis-classifications also occur in other spectrally similar objects. For instance, within the local areas highlighted by the rectangular windows in Plate 2a, the obvious errors occur in the classification of roofs, which are wrongly classified as trails, grass and trees. To some extent, the introduction of GLCM features overcomes the mis-classifications compared to the spectral classification; however, their results can be influenced by the size of analysis windows. In the experiment, the smaller window (3×3) obtains higher accuracy for roofs but is poor for roads. The larger window (7×7) is better for roads; however, it cannot discriminate trails and roofs effectively.

Plate 2g through 2j shows a small portion of the classification maps obtained with (g) GLCM (3×3), (h) GLCM (7×7), (i) PSI, and (j) ES features. From the results of GLCM with the 3 $\times$3 analysis window (Plate 2b and 2g), it can be observed that many roads pixels are classified wrongly as shadow. For the 7×7 analysis window (Plate 2c and 2h), although the classification results in the homogeneous areas have been

TABLE 1. THE NUMBERS OF TRAINING AND TEST SAMPLES IN THE WASHINGTON, D.C. MALL EXPERIMENT

Information Class	Number of Training Samples	Number of Test Samples
Roads	70	1952
Grass	70	1893
Water	70	1539
Trails	60	686
Trees	60	1039
Shadow	60	1093
Roofs	70	3355
Total	530	11557

improved, the shape of the grass land, roofs, and water body is not well modeled around the edge areas. From these classification maps, it can also be seen that the mis-classifications occur in the trails-roofs for both 3×3 and 7×7 windows. Coburn and Roberts (2004) showed that there was no single window size that would adequately characterize the range of textural conditions present in satellite images. 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. The PSI and ES algorithms overcome the deficiencies of moving windows by considering the local isotropy and extending the direction lines; therefore, the objects with different sizes and scales can be discriminated more effectively. Comparing Plate 2i and 2j, we can see that the ES map reduces the pepper-salt effect in the grass land, roads, and water, and at the same time it retains the structural and detailed information in the edge regions. The edge and spectral features are considered at the same time for determining the direction lines. That is the reason why ES can improve the classification in both homogeneous and border areas.

It is worth noting that λ is set as 0.7 (Equation 11) in the experiments because it indicates a pixel that is detected as an edge in most information sources. If a pixel is voted as an edge pixel by at least 70 percent of the classifiers, the extension of direction lines should stop at this pixel. For the ES algorithm, λ is an important parameter for spatial restriction of direction lines. If λ increases, some detailed structures may be smoothed; however, when λ becomes smaller, a pepper-salt effect can be found in the homogeneous regions. Therefore, it is a tradeoff between the classification in the homogeneous areas and the detail preservation. λ is set to 0.7 in the experiments because it has been found effective for classification for both edge and homogeneous areas. Figure 4 shows the relationship between λ and the classification accuracies (Kappa coefficients). The statistics show that the reasonable range for λ should be [0.1, 0.7], and it can be seen that the edge maps with larger values of λ cannot control the extension of direction lines effectively, compared to those with smaller values. It is worth noting that the optimal value of λ could not be determined beforehand.

In order to test the performance of the spatial features fully, all the features discussed above were input to different classifiers. In the experiment, the training and test samples were kept the same. In Figure 5, the Kappa coefficients for four classifiers with four different input features are provided, from which it can be observed that the proposed ES features achieve highest accuracies when different classifiers are used. Hence, it can be concluded that the proposed ES features are effective for spatial feature extraction and classification of VHR data independent of the classifiers.

TABLE 2. THE CLASSIFICATION ACCURACIES FOR DIFFERENT FEATURES IN THE WASHINGTON, D.C. MALL EXPERIMENT

	Spectral		GLCM (3×3)		GLCM (7×7)		PSI		ES	
Class	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Roads	91.1	99.4	66.2	88.9	80.3	95.2	73.5	87.0	87.9	97.2
Grass	84.6	95.0	88.2	95.5	93.1	96.9	96.2	97.1	97.0	99.5
Water	72.3	72.5	83.0	85.1	80.2	86.3	90.3	99.0	93.5	99.5
Trails	78.6	56.7	82.6	65.3	85.3	60.8	90.7	69.2	89.1	78.2
Trees	99.6	74.3	92.5	95.9	95.6	90.7	96.5	89.3	99.3	98.2
Shadow	61.7	60.8	80.2	65.0	82.3	66.5	91.6	87.9	93.5	94.5
Roofs	84.4	91.5	85.7	80.6	84.6	79.3	86.9	80.9	94.9	88.8
AA	81.8	78.6	82.6	82.3	85.9	82.2	89.4	87.2	93.6	93.7
OA	82.8%		84.3%		86.0%		89.7%		93.8%	
Kappa	0.793		0.818		0.829		0.871		0.924	

TABLE 3. THE CONFUSION MATRIX FOR PLATE 2A AND 2E. ("S," "G," AND "W" REPRESENT "SHADOW," "GRASS," AND "WATER", RESPECTIVELY)

	Confusion Matrix for Figure **6(a)**								Confusion Matrix for Figure **6(e)**							
	Roads	Grass	W	Trails	Trees	S	Roofs		Roads	G	W	Trails	Trees	S	Roofs	
Roads	1779	0	0	0	0	4	7	99.4	1715	0	40	0	0	10	0	97.2
Grass	0	1601	0	40	0	0	44	95.0	0	1836	0	8	1	0	0	99.5
Water	12	0	1112	0	0	410	0	72.5	0	0	1439	0	0	8	0	99.5
Trails	0	0	0	539	0	0	411	56.7	0	0	0	611	0	0	170	78.2
Trees	2	292	0	0	1035	5	60	74.3	0	5	0	0	1032	14	0	98.2
S	3	0	427	0	4	674	0	60.8	37	0	22	0	0	1022	0	94.5
Roofs	156	0	0	107	0	0	2833	91.5	200	52	38	67	6	39	3185	88.8
	91.1	84.6	72.3	78.6	99.6	61.7	84.4		87.9	97.0	93.5	89.0	99.3	93.5	94.9	
Overall Accuracy = 82.8% Kappa = 0.793									Overall Accuracy = 93.8% Kappa = 0.924							

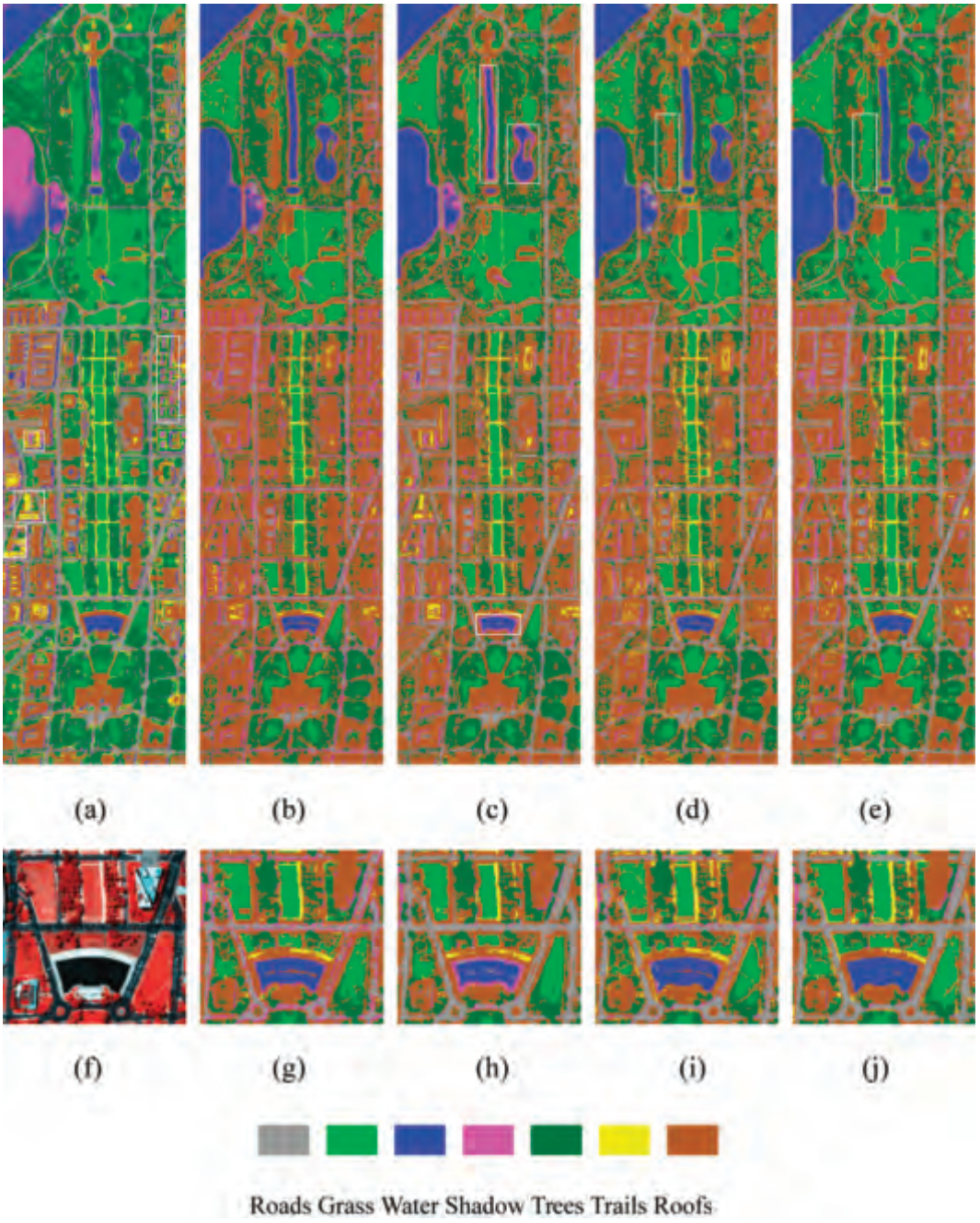

Plate 2. The classification maps for different features: (a) the classification map using only spectral information, (b) the result using GLCM with a 3 × 3 moving window, (c) the result using GLCM with a 7 × 7 window, (d) the result of PSI, and (e) the result acquired by edge-spectral features; the details of classification maps obtained with (g) GLCM (3 × 3), (h) GLCM (7 × 7), (i) PSI, and (j) ES features. The RGB color image is shown in (f).

Experiment on QuickBird Data

In order to verify that the proposed classification algorithm works in a stable manner, a QuickBird multispectral image of Beijing was examined. As is well known, China is a developing country, which has shown rapid economic and social development in recent years. In particular, Beijing is to be the host city of the 2008 Olympic Games, and so additional investment stemming from governments and enterprises are accelerating the city construction, especially for the infrastructure. The VHR images will play an important part in regional planning, project programming, construction monitoring, etc. The test image is shown in Figure 6a, and comprises 516 lines and 512 columns. Its spatial resolution is 2.44 m. A reference image with different information classes is shown in Figure 6b, and the training and test samples were selected randomly from the reference data (Table 4).

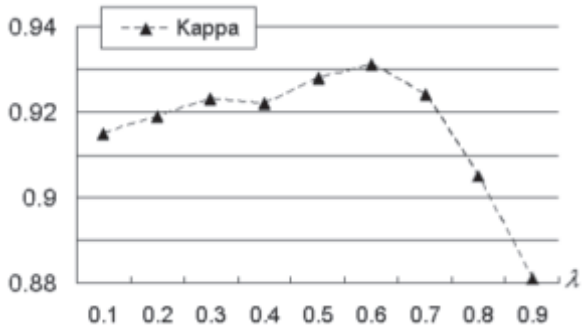

Figure 4. The relationship between λ and the classification accuracies (Kappa coefficients). It can be observed that the reasonable range for λ should be in [0.1, 0.7].

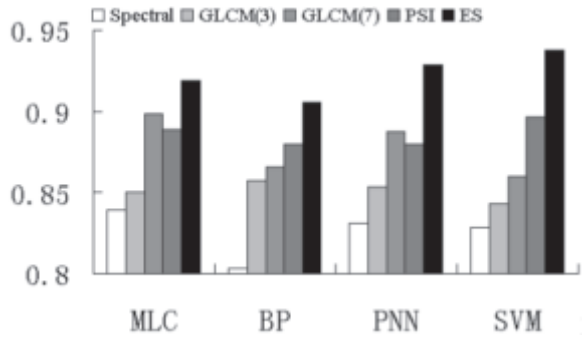

Figure 5. The Kappa coefficients for different classifiers with four features in the Washington, D.C. Mall experiment (the vertical axis represents the Kappa coefficients).

The multispectral bands, independent spectral components, and the linear NDVI were combined to produce the fuzzy edge image, which is shown in Figure 7.

In this experiment, the spatial measures of variance and contrast were again used to extract the GLCM features with a 5×5 analysis window. The parameters of PSI were set as: T1 = 200, T2 = 95, and the value of T in the ES algorithm was 200. The classification accuracies with different classes are available in Table 5 and the classification maps are shown in Figure 8.

From the results in Table 5, it can be seen that nearly all the highest accuracies were achieved by the proposed edge-spectral (ES) features. Compared to the spectral, GLCM and PSI methods, the PA values for building obtained by ES were increased by 26.9 percent, 22.5 percent, and 20.4 percent, respectively; the UA values for roads were increased by 28.7 percent, 28.2 percent, and 25.3 percent, respectively. The results show that ES is a good detector for the spectrally similar classes of building and roads over urban areas, which can be observed clearly from the classification maps shown in Figure 8. The proposed ES features can enhance the classification results for most of the information classes because it integrates the edge and structural features with spectral information effectively. In this QuickBird experiment, compared with

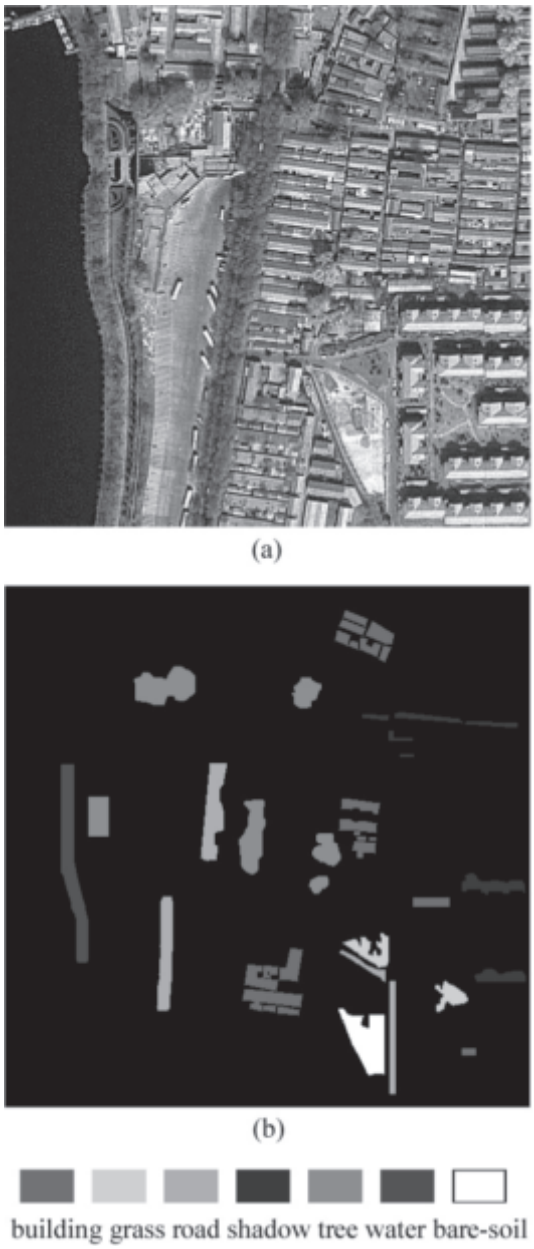

Figure 6. (a) the first principle component for the QuickBird image of Beijing, China, and (b) reference data.

the classification accuracies by the spectral, GLCM, and PSI methods, the improvements in OA are 7.5 percent, 6.7 percent, and 5.6 percent, respectively. The respective improvements in Kappa are 9.1 percent, 8.2 percent, and 6.9 percent. The confusion matrix for the spectral and ES features is provided in Table. 6. Based on the two experiments, it can be concluded that the proposed classification algorithm integrating fuzzy edge maps and multispectral information performed well and in a stable manner.

In this experiment, all the spatial features were again tested in different classifiers. All the training and test samples were the same for the different classifiers. Figure 9 shows the overall accuracies for the four classifiers with

TABLE 4. THE NUMBERS OF TRAINING AND TEST DATA IN THE QUICKBIRD EXPERIMENT

Information Class	Number of Training Samples	Number of Test Samples
Roads	100	1240
Building	95	1835
Water	95	670
Grass	85	244
Bare Soil	100	328
Shadow	100	592
Tree	80	643
Total	655	5552

Figure 7. The fuzzy edge map for the QuickBird test image.

different features. The figure shows that the hybrid spatial-spectral features perform better than the spectral features alone. All the GLCM, PSI, and ES features can strengthen the spectral classification to some extent, and the improvement using ES is the most significant. Compared with the spectral features, the accuracy improvements achieved by ES are 4.7 percent, 8.1 percent, 5.5 percent, and 7.5 percent for the MLC, BP, PNN, and SVM, respectively. Therefore, we can conclude that the proposed classification algorithm gives better performance independent of the classifier used.

Conclusions

In this paper, a novel algorithm integrating the edge and multispectral information is proposed to improve the classification of very high spatial resolution remotely sensed imagery. First, multiple information sources, including multispectral bands, the linear NDVI, and the independent spectral components extracted by ICA, were combined to produce a fuzzy edge map. Secondly, a methodology for the extension of direction-lines was proposed based on the multispectral and the fuzzy edge maps. Two structural features, called mean and length-width ratio, were calculated according to the direction-lines histogram. Lastly, the two structural measures were integrated with the multispectral bands using the support vector machine to obtain the classification result.

In the experiments on the two datasets, the proposed structural features were compared to the multispectral features and several other spatial measures, including GLCM and PSI. The proposed approach performed well in terms of accuracies and the experimental result was promising because it integrated the edge and multispectral information effectively. In experiments, multispectral, GLCM, PSI, and the proposed ES features were input into different classifiers, and the results show that ES features can achieve the highest accuracies with all the classifiers.

From the statistics in Tables 2 and 5, it can be seen that spectral information is still very important in the interpretation of high spatial resolution imagery, especially for vegetation areas. With the current development of sensors, hyperspectral data with high spatial resolution have become available. In our future research, we will work toward combining spectral and spatial classification results more effectively. The possible fusion may lie in the feature level, the decision level and in post-processing.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant No. 40771139. The Washington, D.C. data were obtained from the student CD-ROM which accompanies D.A. Landgrebe's book, *Signal Theory Methods in Multispectral Remote Sensing*. The authors also appreciate the anonymous reviewers for their comments. Their insightful suggestions have significantly improved this paper.

References

Barber, D.G., and E.F. LeDrew, 1991. SAR sea ice discrimination using texture statistic: A multivariate approach, *Photogrammetric Engineering & Remote Sensing*, 57(8):949–958.

TABLE 5. THE CLASSIFICATION ACCURACIES FOR DIFFERENT FEATURES IN THE QUICKBIRD EXPERIMENT

	Spectral		GLCM		PSI		ES	
Class	PA	UA	PA	UA	PA	UA	PA	UA
Roads	0.935	0.697	0.936	0.702	0.927	0.731	0.839	0.984
Building	0.552	0.884	0.596	0.880	0.617	0.883	0.821	0.885
Water	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Grass	0.996	0.935	0.978	0.941	0.980	0.937	0.984	0.972
Bare Soil	0.948	0.555	0.966	0.619	0.960	0.569	1.000	0.572
Shadow	0.902	0.966	0.837	0.958	0.892	0.994	0.929	0.996
Tree	0.967	0.889	0.967	0.877	0.969	0.906	0.989	0.852
AA	0.900	0.846	0.897	0.854	0.906	0.860	0.938	0.895
OA	0.820		0.828		0.839		0.895	
Kappa	0.780		0.789		0.802		0.871	

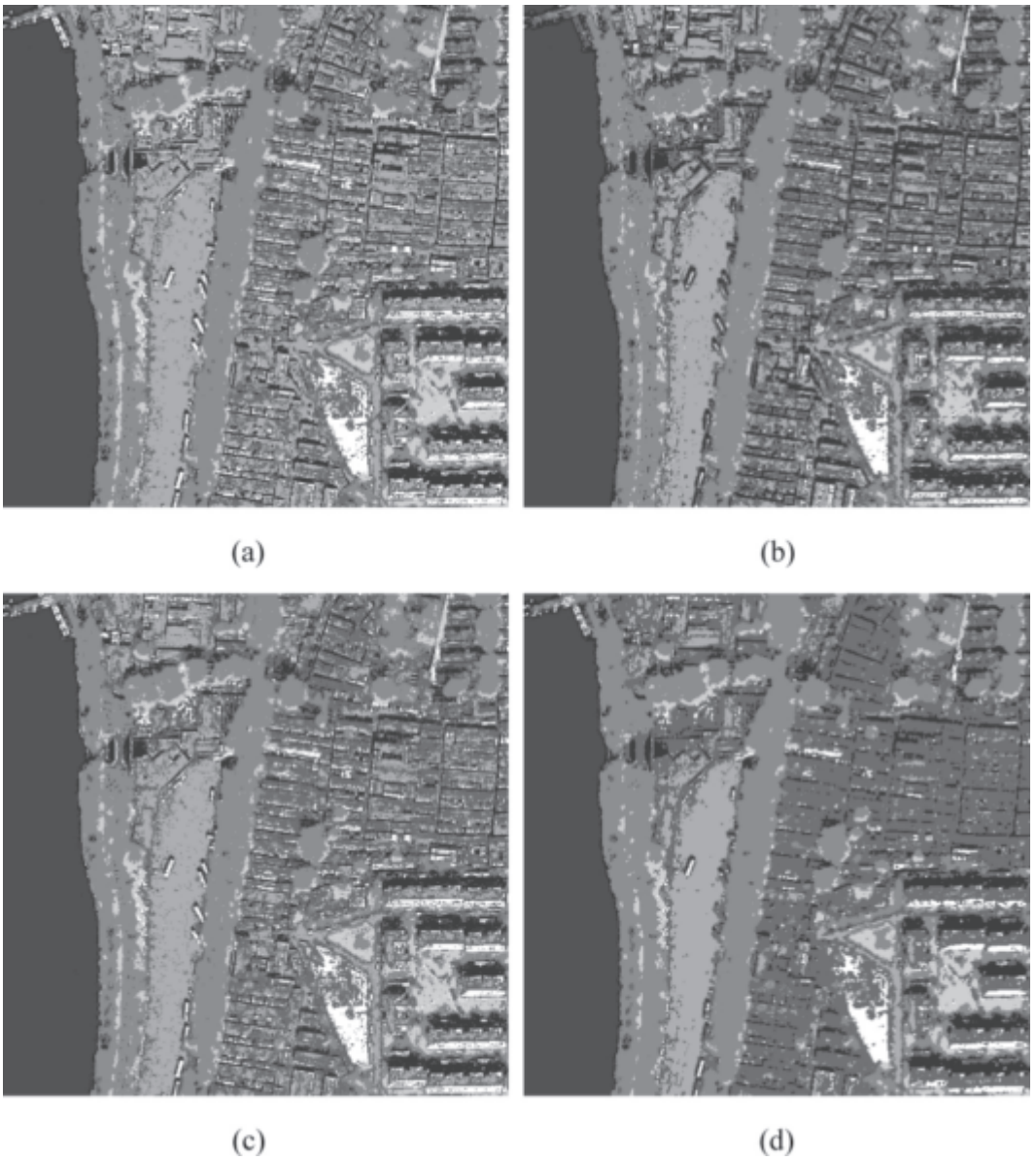

Figure 8. The classification maps for (a) spectral, (b) GLCM, (c) PSI, and (d) ES features. The legend for the maps has been given in Figure 6.

TABLE 6. THE CONFUSION MATRIX FOR FIGURES 8A AND 8D. ("B," "B-S," "W," AND "S" REPRESENT "BUILDINGS," "BARE-SOIL," "WATER," AND "SHADOW", RESPECTIVELY)

	Confusion Matrix for Figure **8a**								Confusion Matrix for Figure **8d**							
	Roads	B	W	Grass	B-S	S	Trees		Roads	B	W	Grass	B-S	S	Trees	
Roads	1159	502	0	0	0	1	0	69.7	1040	17	0	0	0	0	0	98.4
B	79	1014	0	0	15	34	5	88.4	198	1507	0	0	0	0	0	88.4
Water	0	0	670	0	0	0	0	100	0	0	670	0	0	0	0	100
Grass	0	0	0	243	2	0	15	93.5	0	0	0	240	0	0	7	97.2
B-S	1	248	0	0	311	0	0	55.5	1	245	0	0	328	0	0	57.2
Shadow	1	17	0	0	0	534	1	96.6	1	1	0	0	0	550	0	99.6
Trees	0	54	0	1	0	23	622	88.9	0	65	0	4	0	42	636	85.2
	93.5	55.3	100	99.6	94.8	90.2	96.7		83.9	82.1	100	98.4	100	92.9	98.9	
Overall Accuracy = 82.0% Kappa = 0.780									Overall Accuracy = 89.5% Kappa = 0.871							

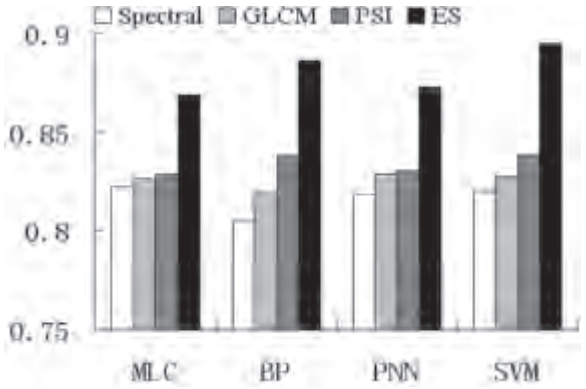

Figure 9. The overall accuracies for different classifiers with four features in the Beijing QuickBird experiment.

Benediktsson, J.A., M. Pesaresi, and K. Arnason, 2003. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations, I*EEE Transactions on Geoscience and Remote Sensing*, 41(9):1940–1949.

Benediktsson, J.A., J.A. Palmason, and J.R. Sveinsson, 2005. Classification of hyperspectral data from urban areas based on extended morphological profiles, *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):480–491.

Camps-Valls, G., and L. Bruzzone, 2005. Kernel-based methods for hyperspectral image classification, *IEEE Transactions on Geoscience and Remote Sensing*, 43(6):1351–1362.

Chapelle, O., V. Vapnik, O. Bousquet, and S. Mukhherjee, 2002. Choosing multiple parameters for support vector machine, *Machine Learning*, 46:131–159.

Coburn, C.A. and A.C.B. Roberts, 2004. A multiscale texture analysis procedure for improved forest stand classification, *International Journal of Remote Sensing*, 25(20):4287–4308.

Dell'Acqua, F., P. Gamba, A. Ferari, J.A. Palmason, J.A. Benediktsson, and K. Arnason, 2004. Exploiting spectral and spatial information in hyperspectral urban data with high-resolution, *IEEE Geoscience and Remote Sensing Letters*, 1(4):322–326.

Foody, G.M., and A. Mather, 2004. A relative evaluation of multiclass image classification by support vector machines, *IEEE Transactions on Geoscience and Remote Sensing*, 42(6):1335–1343.

Gong, P., D.J. Marceau, P.J. Howarth, 1992. A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data, *Remote Sensing of Environment*, 40:137–151.

Hyvarinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis, *IEEE Transactions on Neural Networks*, 10(3):626–634.

Landgrebe, D., 2003. *Signal Theory Methods in Multispectral Remote Sensing*, John Wiley and Sons, Hoboken, New Jersey, 520 p.

Melgani, F., and L. Bruzzone, 2004. Classification of hyperspectral remote sensing images with support vector machines, *IEEE Transactions on Geoscience and Remote Sensing*, 42(8):1778–1790.

Myint, S.W., N.S.N. Lam, and J. Tyler, 2004. Wavelets for urban spatial feature discrimination: Comparisons with fractal, spatial autocorrelation, and spatial co-occurrence approaches, *Photogrammetric Engineering & Remote Sensing*, 70(7):803–812.

Solaiman, B., R.K. Koffi, M.C. Mouchot, and A. Hillion, 1998. An information fusion method for multispectral image classification postprocessing, *IEEE Transactions on Geoscience and Remote Sensing*, 36(2):395–406.

Sun, W.X., V. Heidt, P. Gong, and G. Xu, 2003. Information fusion for rural land-use classification with high-resolution satellite imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 41(4):2840–2850.

Tian, B., M.A. Shaikh, M.R. Azimi-Sadjadi, T.H. Vonder, and D.L. Reinke, 1999. A study of cloud classification with neural networks using spectral and textural features, *IEEE Transactions on Neural Networks*, 10(1):138–151.

Unsalan, C., and K.L. Boyer, 2004. Classifying land development in high-resolution panchromatic satellite images using straight-line statistics, *IEEE Transactions on Geoscience and Remote Sensing*, 42(4):907–919.

Unsalan, C., and K.L. Boyer, 2004. Classifying land development in high-resolution panchromatic satellite imagery using hybrid structural-multispectral features, *IEEE Transactions on Geoscience and Remote Sensing*, 42(12):2840–2850.

Yu, S.X., 2005. Segmentation induced by scale invariance, *Proceedings of the IEEE Computer Society Conference on CVPR 2005*, 20–25 June, San Diego, California, pp. 444–451.

Zhang, L., X. Huang, B. Huang, and P. Li, 2006. A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 44(10):2950–2961.

Zhang, Y., 1999. Optimisation of building detection in satellite images by combining multispectral classification and texture filtering, *ISPRS Journal of Photogrammetry and Remote Sensing*, 54:50–60.

(Received 12 March 2007; accepted 19 June 2007; revised 06 September 2007)