A Pixel Shape Index Coupled With Spectral Information for Classification of High Spatial Resolution Remotely Sensed Imagery

Liangpei Zhang, Xin Huang, Bo Huang, and Pingxiang Li

Abstract—**Shape and spectra are both important features of high spatial resolution remotely sensed (HSRRS) imagery, and they are concrete manifestation of textures on such imagery. This paper presents a spatial feature index, pixel shape index (PSI), to describe the shape feature in a local area surrounding a pixel. PSI is a pixel-based feature which measures the gray similarity distance in every direction. As merely the shape feature is inadequate for classifying HSRRS imagery, a transformed spectral feature extracted by independent component analysis is added to the input vectors of our classifier, and this replaces the original multispectral bands. Meanwhile, a fast fusion algorithm that integrates both shape and spectral features using the support vector machine has been developed to interpret the complex input vectors. The results by PSI are compared with some spatial features extracted using wavelet transform, gray level co-occurrence matrix, and the length–width extraction algorithm to test its effectiveness. The experiments demonstrate that PSI is capable of describing shape features effectively and result in more accurate classifications than other methods. While it is found that spectral and shape features can complement each other and their integration can improve classification accuracy, the transformed spectral components are also found to be more suitable for classification.**

Index Terms—**Independent components analysis (ICA), integration of shape and spectra, shape feature, support vector machine (SVM).**

I. INTRODUCTION

HIGH spatial resolution remotely sensed (HSRRS) images with multispectral bands such as QuickBird and IKONOS provide a large amount of information, thus opening up avenues for new remote sensing applications. However, their availability poses challenges to image classification. Due to the complex spatial arrangement and spectral heterogeneity even within the same class, conventional spectral classification methods are grossly inadequate for HSRRS imagery [1]. In order to overcome this inadequacy, spectral features must evidently be complemented by one or the other means. It is by and large agreeable to not only use the spectral information, but also to exploit spatial analysis [2]. Spatial analytical approaches can simply be categorized into spatial features extracted by moving windows or elements and a spatial classifier based on contextual decision criteria with consideration of neighboring pixels inside the classifier [2]. This paper focuses on the spatial features to strengthen the feature space for better classification accuracies.

Several effective spatial features that have been proposed, e.g., wavelet transform (WT) methods, are used to extract texture features [1], [3]. WT allows extraction of different texture features at different scales, thus providing a useful alternative for the spatial features analysis of HSRRS images. The gray level co-occurrence matrix (GLCM) method can introduce spatial information into a spectral classification, in which an image gray value is transformed into the co-occurrence matrix space and various output images are calculated adopting different spatial measures [4], [5]. The Gaussian Markov random field (GMRF) model, which is an accurate and compact version of the Markov random field (MRF) [6], [7], is yet another representative method for extracting spatial information. Notably, Shackelford and Davis [8] proposed an effective algorithm, length–width extraction algorithm (LWEA), to extract the length and width of spectrally similar connected groups of pixels. The notion of LWEA is akin to GLCM, which also measures the spatial features within a spectrally similar neighborhood along a certain direction. In addition to the aforementioned methods, a series of spatial features extracted by morphological profiles (MPs) and differential morphological profiles (DMPs) was proposed [9]–[11]. These morphological features have proven to be effective in extracting multiscale structural information from HSRRS images and, hence, complement spectral information.

Based on the aforementioned work, this paper investigates the following three issues pertaining to spectral and spatial feature extraction and their fusion in the classifier.

1) In addition to textural features, HSRRS imagery also includes ample shape features. The complex raster data structure of remotely sensed images makes the shape description of objects a nontrivial task. Especially, as the spatial resolution increases, sharper spatial features must be interpreted and captured effectively. This paper, therefore, puts forward the pixel shape index (PSI), an improved version of LWEA, to describe the spatial information around the central pixel in its entirety.

2) Quite frequently, multispectral classification utilizes the original spectral bands as the input for classifiers.

Nevertheless, it is very difficult to achieve good classification results this way, as most of the land cover classes in an HSRRS image contain a number of spectrally different pixels or objects [1]. Thus, it would be interesting to investigate if some transformed spectral features can achieve higher accuracy than the original multispectral HSRRS images. For this purpose, the feature extraction method of independent component analysis (ICA) is employed. ICA is an effective approach for blind source separation and has drawn a great deal of attention owing to its tremendous potential in image processing.

3) High-resolution imagery reveals objects in greater detail, depicting their size, shape, structure, and color. Unlike medium-resolution images, a single feature is not sufficient for classifying the high-resolution imagery. Thus, interpretation of such images entails multiple effective features. An appropriate classifier and better decision rules are inevitable for identifying the multiple input features. Unlike the multispectral classification method [2] and the classifiers fusion approach [12]–[14], the multiple spectral and spatial features in this paper are input into 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 information if the input vector contains both spectral and spatial features. This SVM approach of a single classifier with multiple features is chosen because of its fast and effective processing.

The rest of this paper is organized as follows. Section II describes the PSI algorithm. Section III discusses ICA transformation for the original spectral bands. Section IV describes the fusion of spectral and spatial features based on SVM, and Section V details the two experiments with two multispectral HSRRS datasets. The first experiment detects the influence of three parameters on the PSI algorithm, and the second explores fusion of the shape and spectral features. Finally, Section VI concludes the paper.

II. PIXEL SHAPE INDEX

Shackelford and Davis [8] proposed LWEA to examine the context of each pixel and measure the spatial dimensions of groups of spectrally similar connected pixels. LWEA calculated a length and width value for each pixel in the image. These values are found by searching along a predetermined number of equally spaced lines radiating from the central pixel.

PSI adds an extension to LWEA. The shape features are extracted in a pixel-by-pixel manner, wherein each pixel has a feature value. The shape represents the contextual feature along all the directions to some predefined limit around the central pixel. Like LWEA and GLCM, PSI is also based on the spectral similarity of its neighboring pixels.

PSI shape feature extraction consists of three steps.

1) extension of direction lines based on gray level similarity;
2) measurement of the length of each direction line;
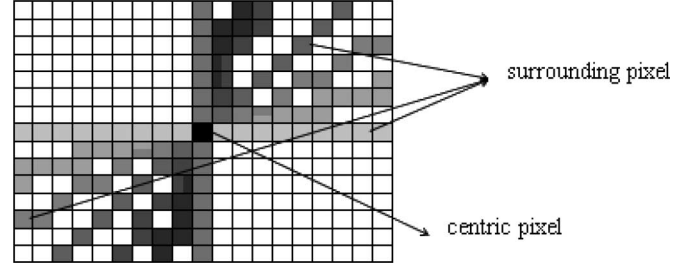3) calculation of PSI.

Fig. 1. Direction lines of PSI algorithm.

A. Extension of Direction Lines

The extended direction lines used to detect the object's overall contour are symmetric around the centric pixel. In one chosen direction, the spectral difference is measured between a pixel and its centric pixel in order to decide if this pixel lies in the homogeneous area around the centric pixel. Some example direction lines are shown in Fig. 1, where different gray levels symbolize different direction lines.

The pixel homogeneity is defined using the following method:

$$\text{PH}_i = \sum_{s=1}^{n} |p_s^{\text{cen}} - p_s^{\text{sur}}| \tag{1}$$

where PH_i represents the spectral homogeneity of the ith direction between the centric pixel and its surrounding pixel, n denotes the number of spectral bands, p_s^{cen} is the spectral value of the centric pixel, as shown in black in Fig. 1, and p_s^{sur} denotes the spectral value of the surrounding pixel, which is on the direction line in different gray levels (Fig. 1).

The ith direction line is extended if the following conditions are met.

1) PH_i is less than a predefined threshold T1.
2) The total number of pixels in this direction line is less than another predefined threshold T2.

The extension of each direction line will cease if either of the above conditions is not met. In this case, the extension of the ith direction line will be terminated, and the algorithm will skip to the $(i + 1)$th direction line.

B. Length of Direction Line

After determining all the direction lines, the length of each and every direction line is calculated. In the ith direction, the length is measured as follows:

$$d_i = \max \left\{ |m^{e1} - m^{e2}|, |n^{e1} - n^{e2}| \right\} \tag{2}$$

where (m^{e1}, n^{e1}) denotes the row and column number of the pixel in one end of the direction line, and (m^{e2}, n^{e2}) denotes the row and column numbers of the pixel in the other end. The maximum city-block distance is adopted here because: 1) it can smooth the spatial feature values in a homogeneous area and let the pixels in the same shape area possess the same or close PSI values and 2) a large amount of computation can be saved with the use of integers instead of decimals in (2).
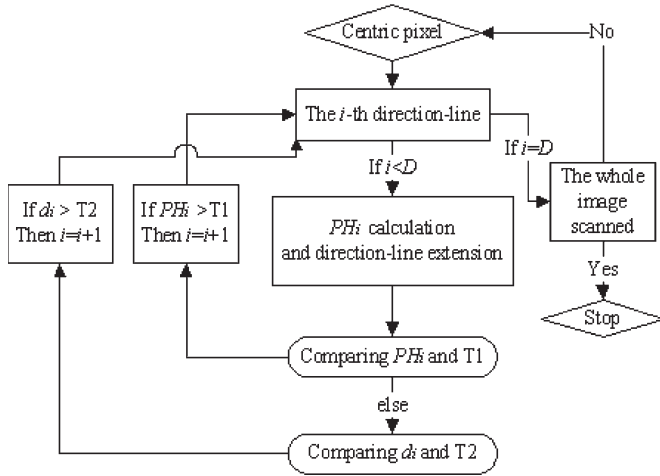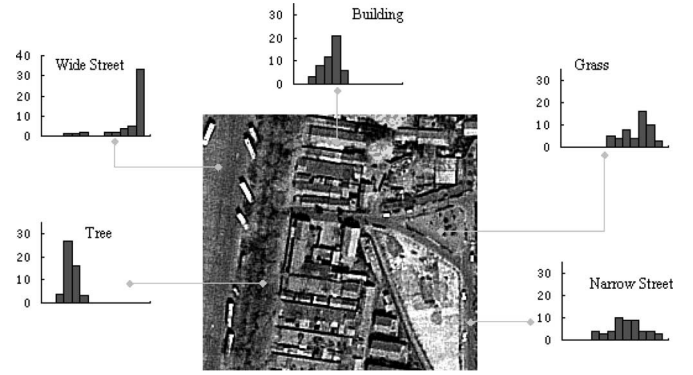
Fig. 2.　Flowchart of the PSI algorithm.

Fig. 3.　PSI feature values for four classes (wide street, narrow street, building, tree, and grass) from a multispectral image of Beijing, China. The PSI parameters used in this figure are $D = 20$, $T_1 = 110$, and $T_2 = 50$. The vertical axis of the histogram represents the occurrence frequency of the PSI value in a local area.

Upon completion of this step, the lengths of all the direction lines are obtained. Thus, the feature sets of the centric pixel can be formulated as $(d_1, d_2, d_3, \ldots, d_D)$, where D denotes the total number of all the directions.

C. Calculation of PSI

The shape index of the centric pixel is the sum total length of all the direction lines, which is defined as

$$\mathrm{PSI}_{m,n} = \sum_{i=1}^{D} d_i \tag{3}$$

where $\mathrm{PSI}_{m,n}$ represents the shape index of the centric pixel (m, n). The flowchart of PSI is shown in Fig. 2.

The spectral similarity is used to construct spectrally connected groups of pixels to describe the shape features. There are three parameters in PSI: the total number of directions D, and the direction line extension thresholds T_1 and T_2.

The parameter D indicates the total number of direction lines of PSI. The value of D can be raised in order to describe the shape features in detail. The values, T_1 and T_2 ought to be determined through experiments as these two parameters are related to the shape and spatial arrangement of objects in the image. T_1 is the threshold for homogeneity, and pertains to the spectral variability in the same shape area. T_2 is the maximum length for one direction line, and relates to the size of one shape area. More details about the parameters of PSI will be discussed later under the experiments section.

It should be noted that PSI differs from LWEA in the following aspects.

1) LWEA extracts length–width features and stores the maximum value as the length and the minimum value as the width after searching all direction lines. However, PSI sums the length of all the direction lines instead of just the maximum length and width for describing the contextual structure around the centric pixel. The features of length and width in LWEA only represent two directions around the centric pixel, which is inadequate to provide sufficient contextual information along all directions. The aim of PSI is to describe the overall neighborhood features and

let all the pixels in the same shape area have the same feature value.

2) In LWEA, the spatial feature values are determined by searching along the predetermined number of equally spaced lines radiating from the centric pixel; however, these values in PSI are calculated along the direction lines throughout the centric pixel. Hence, PSI provides an orientation-independent property, which, in turn, has an advantage to provide more stable spatial features. Moreover, the azimuth sampling in PSI is enhanced from 10 of LWEA ($D = 18$) to 9 ($D = 20$), which improves the descriptive ability of spatial features.

To illustrate the strength of PSI, an example of PSI features for four classes in a multispectral QuickBird image of Beijing, China, is given in Fig. 3. It can be found that different PSI values are achieved within the spectrally similar classes such as building versus road and grass versus tree.

III. INDEPENDENT COMPONENT ANALYSIS

ICA is a multivariate data analysis method for blind source separation of signals. ICA seeks to render the components as statistically independent as possible, and is a useful tool for data mining of remotely sensed data. The use of ICA features for unsupervised signature extraction of IKONOS images has been investigated previously [15]. In this paper, the ICA features of the QuickBird RGB bands are, however, extracted to strengthen multispectral information to reduce the spectral complexity and complement the shape features extracted using the PSI-based algorithm.

The basic model of ICA is

$$\boldsymbol{Y} = \boldsymbol{A}\,\boldsymbol{X} \tag{4}$$

where

$\boldsymbol{Y}$　vector of observed signals;
$\boldsymbol{A}$　mixing matrix to be estimated;
$\boldsymbol{X}$　mutually independent components.

The goal of ICA is to calculate the matrix $\boldsymbol{B}$ such that the sources $\boldsymbol{X} = \boldsymbol{B}\,\boldsymbol{Y}$ can be estimated from the observed vector $\boldsymbol{Y}$ by optimizing the statistical independence criterion.
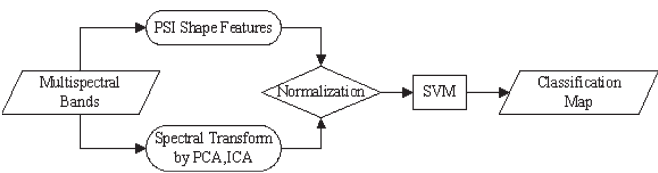
Fig. 4. Fusion of transformed spectral information and PSI shape features using SVM.

Fig. 5. QuickBird image of typical buildings and roads in Beijing.

TABLE I
NUMBER OF TRAINING AND TESTING PIXELS IN FIG. 5

	BUILDING	ROAD	TREE
TRAINING	596	562	413
TESTING	22403	7831	3198

Our paper adopts the fast and robust ICA (Fast-ICA) algorithm proposed by Hyvarinen [16]. It is a fixed-point algorithm based on an optimization of entropy function called negative entropy. Initially, this algorithm was introduced using kurtosis as a function, and was subsequently extended for general contrast functions such as

$$J_G(w) = \left[E\left\{ G(w^{\mathrm{T}}x) \right\} - E\left\{ G(v) \right\} \right]^2 \qquad (5)$$

where w is an m-dimensional vector constrained such that $E\{(w^{\mathrm{T}}x)^2\} = 1$. v is a Gaussian variable of zero mean and unit variance. $G(y)$ is practically any nonquadratic function and is used in its derivation form $g(y)$ in the Fast-ICA algorithm, which can be chosen as follows:

$$G(y) = \frac{1}{4}y^4 \quad g(y) = y^3$$

$$G(y) = \frac{1}{a}\log\left(\cosh(ay)\right) \quad g(y) = \tanh(ay)$$

$$G(y) = -\frac{1}{a}e^{-\frac{ay^2}{2}} \quad g(y) = y\,e^{-\frac{ay^2}{2}}. \qquad (6)$$

TABLE II
CLASSIFICATION ACCURACIES OF DIFFERENT VALUES
FOR D WITH $T_1 = 110$ AND $T_2 = 100$

D	8	12	16	20
Building(%)	87.48	88.06	88.10	88.18
Road(%)	72.83	76.89	77.91	78.61
Tree(%)	97.62	97.53	97.25	97.22
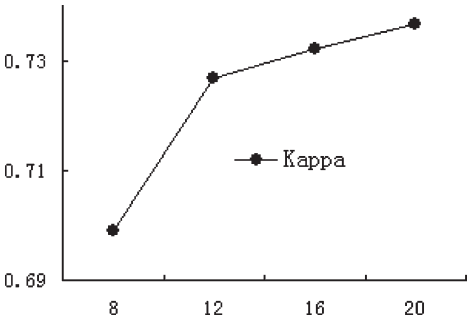
Fig. 6. Relationship between Kappa coefficient and D.
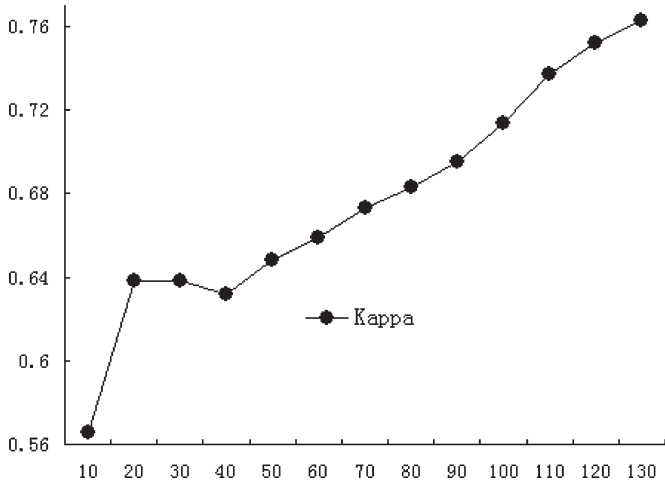
Fig. 7. Relationship between Kappa coefficient and T_1.

Fig. 8. Relationship between Kappa coefficient and T_2.

One by one, the Fast-ICA algorithm estimates each independent component, and the independent component can be obtained subsequent to the optimization process denoted by (5). The RGB bands of the QuickBird are viewed as the input observed vectors, and Fast-ICA is employed to obtain three spectral independent components.

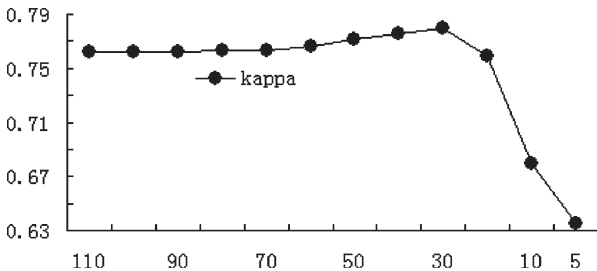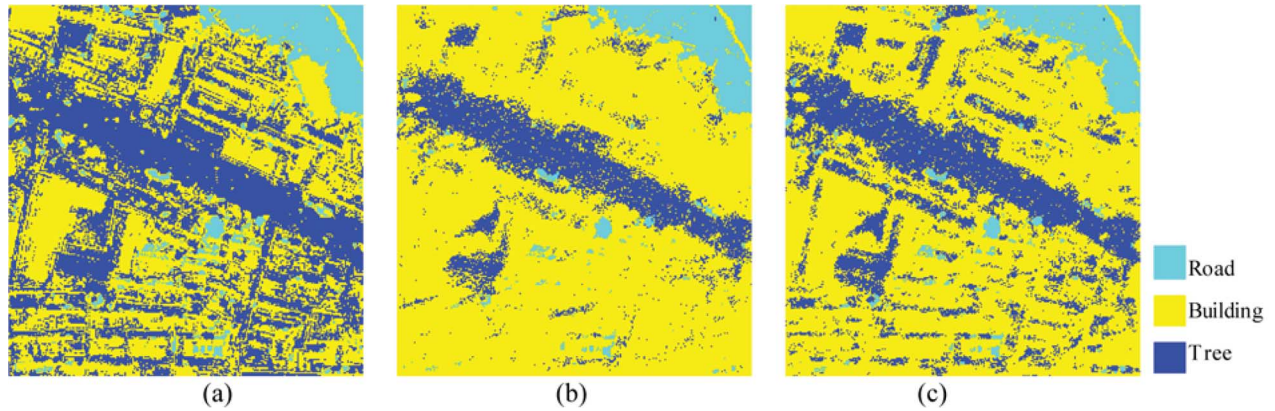Fig. 9. (a) Classification result of spectral features. (b) and (c) The classification results with different parameter values of PSI. The parameters setting of are, respectively, $D = 20$, $T_1 = 100$, and $T_2 = 30$ in (b), and $D = 20$, $T_1 = 100$, and $T_2 = 5$ in (c).
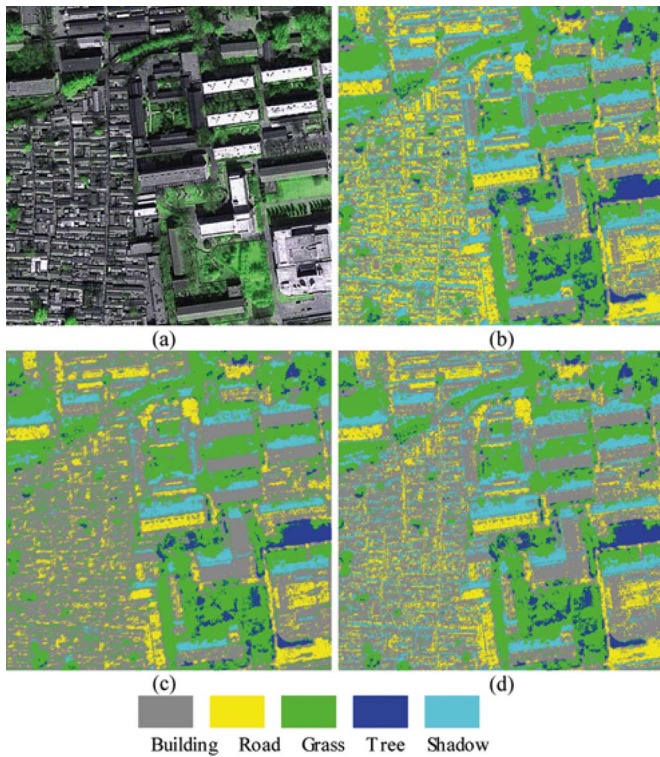

Fig. 10. (a) RGB image, (b) the classification map of original spectral features, (c) the classification result with $T_2 = 5$, and (d) the classification result with $T_2 = 40$. The other two PSI parameters are set to the same $T_1 = 110$ and $D = 20$ in this test.

Fig. 11. Relationship between Kappa coefficient and T_2.

Fig. 12. QuickBird image for the second experimental area in RGB true color.

IV. INTEGRATION OF SHAPE AND SPECTRAL FEATURES USING SVM

This paper examines the feasibility of fusion using PSI integrated with spectral features based on SVM. Primarily, three problems need to be resolved in the fusion: the strategy of integration, normalization prior to classification, and parameters selection of SVM.

Various methods have been proposed to integrate spectral and spatial features. One appropriate method is that of decision-level fusion [2], [12]–[14]. This way, spectral and spatial feature bands can be processed, respectively, in different classifiers, and the final classification map can be obtained by the ap-

proaches based on decision-level fusion. A notable hierarchical fuzzy classification approach has been developed [8], wherein different features were employed for classes with different characteristics. However, as HSRRS data require a significant amount of memory for storage and analysis, computationally light algorithms are of our particular interest [2].

In this paper, the single classifier with multiple features is utilized, which requires less CPU time for classification than the multiple-classifiers method. Owing to the complexity of input features, a nonparametric model needs to be chosen for the classification [17]. Conventional classifiers like the

maximum–likelihood classifier (MLC) are not capable of achieving a satisfactory accuracy. This is because the estimated distribution function usually employs the normal distribution, which may not represent the actual distribution of the data. In recognition of this, fuzzy ARTMAP was tested for the contextual features extracted by the spatial reclassification algorithm [2]. Similarly in [10] and [11], the morphological features after feature selection and the spectrally transformed bands by principal component analysis (PCA) were input to the neural network classifier at the same time. This paper investigates SVM owing to its computational simplicity and superior accuracy when compared to other classifiers.

The second problem is to normalize the spatial and spectral features into [0, 1] in such a way that they can then be input into the classifier. The normalization method of spectral inputs differs from that of spatial ones because of the dissimilar range distribution of spatial feature values. The normalization methods can be formulated as follows:

$$\text{Step1}: \begin{cases} d'_{ij} = \frac{d_{ij} - d_{\min}}{d_{\max} - d_{\min}} \cdot 255 & \text{spectral features} \\ histogram - equalization & \text{spatial features} \end{cases} \tag{7}$$

$$\text{Step2}: d''_{ij} = \frac{d'_{ij} - d'_{\max}}{d'_{\max} - d'_{\min}} \cdot 1 \tag{8}$$

where d_{ij} denotes the original value of the pixel (i, j), and accordingly $d_{\max}, d_{\min}$ represent the maximum and minimum value in that band, d'''_{ij} is the feature value of the pixel (i, j) after normalization.

The third problem relates to the parameters of SVM. SVM classifiers [18], [19] of the form $f(x) = w \cdot \Phi(x) + b$ are learned from the data $\{(x_i, y_i), i = 1, 2, 3, \ldots, N\}$, where x_i is an n-dimensional feature vector, $f(x)$ denotes a hyperplane, which separates samples label $y_i = \pm 1$ on each side, and w and b are the parameters of the hyperplane. The hyperplane calculation can be formulated into a constrained optimization problem as follows:

$$\min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} \xi_i \tag{9}$$

where C is a regularization parameter and ξ_i is slack-variables subject to the constraints

$$y_i (w \cdot \Phi(x_i) + b) \geq 1 - \xi_i \xi_i \geq 0. \tag{10}$$

To simplify the learning of nonlinear SVMs, the objective function is typically expressed in its dual form

$$\min_{0 \leq \alpha_i \leq C} W = \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i Q_{ij} \alpha_j - \sum_{i=1}^{N} \alpha_i + b \sum_{i=1}^{N} y_i \alpha_i. \tag{11}$$

Subject to $0 \leq \alpha_i \leq C$ and $\sum_{i=1}^{N} y_i \alpha_i = 0$, where b is Lagrange multiplier and $Q_{ii} = y_i y_i \Phi(x_i) \Phi(x_i)$. For more

TABLE III
NUMBER OF TRAINING AND TESTING PIXELS

Classes	training pixels	reference pixels
Water	394	18538
Tree	362	15492
Grass	242	1598
Building	468	2215
Bare soil	306	11971
Road	448	2714
Shadow	241	22648

Fig. 13. Available reference data.

details about the SVM, see [18]–[20]. The design of SVM for a specific application needs to handle several issues.

1) *Multiclass classification:* While SVM was originally designed for binary classification, most remote sensing applications involved multiple classes. Two main approaches have been suggested for extending SVM to multiclass classifications [21]. The first is "one against all" (OAA), which consists of a set of binary SVMs. Each classifier is trained to separate one class from the rest and the pixel is allocated to the class for which the largest decision value is determined. The second is "one against one" (OAO), where a series of classifiers are applied to each pair of classes, with the most commonly computed class label reserved for each pixel [21]. In our experiments, the OAO method is employed.

2) *Selection of kernel function:* The commonly used kernel functions are Gaussian radial basis function (RBF) and polynomial function (POLY):

$$\begin{cases} K(x_i, x) = (x_i \cdot x + 1)^p & \text{POLY} \\ K(x_i, x) = \exp\left(-\gamma \|x_i - x\|^2\right) & \text{RBF.} \end{cases} \tag{12}$$

For pattern recognition of HSRRS 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 only has great respondence to local points around the central value [17]. Considering that the feature space of HSRRS images is dispersedly distributed without an obvious clustering center, the POLY kernel is used in this paper.

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 ways exist, which can facilitate
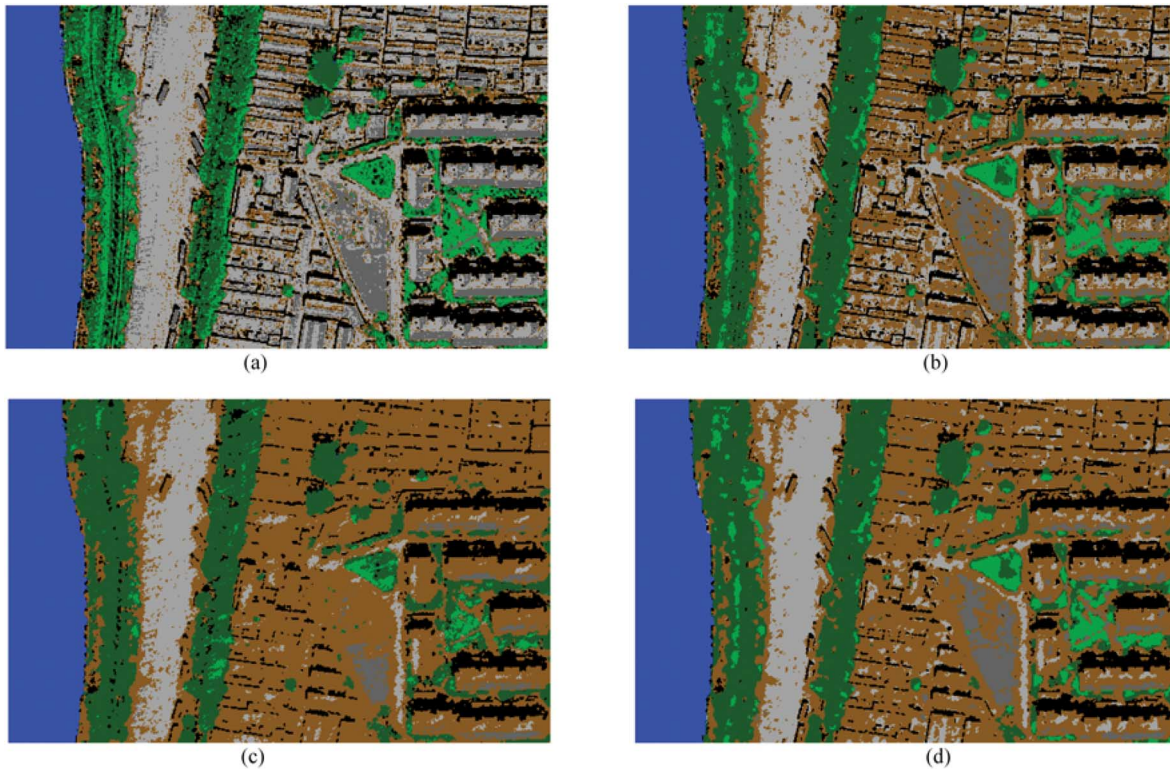
Fig. 14. Classification results of (a) original RGB spectral features, (b) ICA signals, (c) PSI and RGB features, and (d) ICA and PSI features fused by SVM.

TABLE IV
CONFUSION MATRIX OF THE CLASSIFICATION RESULT FOR FIG. 14(a)

	Water	Trees	Grass	Baresoil	Roads	Shadow	Building	Total	%
Water	17147	0	0	0	0	1391	0	18538	92.5
Tree	2	8927	3575	30	65	2098	795	15492	57.6
Grass	0	625	921	15	8	7	22	1598	57.6
Baresoil	0	0	0	1891	281	4	319	2215	85.4
Road	0	11	7	393	10518	36	1006	11971	87.9
Shadow	30	4	0	1	6	2641	32	2714	97.3
Building	0	75	15	8186	10298	628	3446	22648	15.2
Total	17179	9642	4518	10516	21176	6805	5340	75176	
%	99.8	94.0	21.0	13.6	50.1	43.8	70.5		

Overall accuracy = 60.5%, Kappa coefficient = 0.535

the selection of parameters in SVM classifier design such as the well-known leave-one-out (LOO) procedure [21]. In this paper, these parameters are selected based on our experience considering that different parameters setting can be compared in that way. The flowchart of the whole processing is shown in Fig. 4.

V. EXPERIMENTS

Two experiments were performed to test the effectiveness of the proposed methodologies. The first was designed to detect the influence of the PSI parameters. The second was to compare the classification effect of different spatial features. Both the experiments employed QuickBird images of Beijing.

A. Experiment 1—Parameters of PSI

Typically, an urban infrastructure consists of buildings and roads. However, these two types of objects usually present
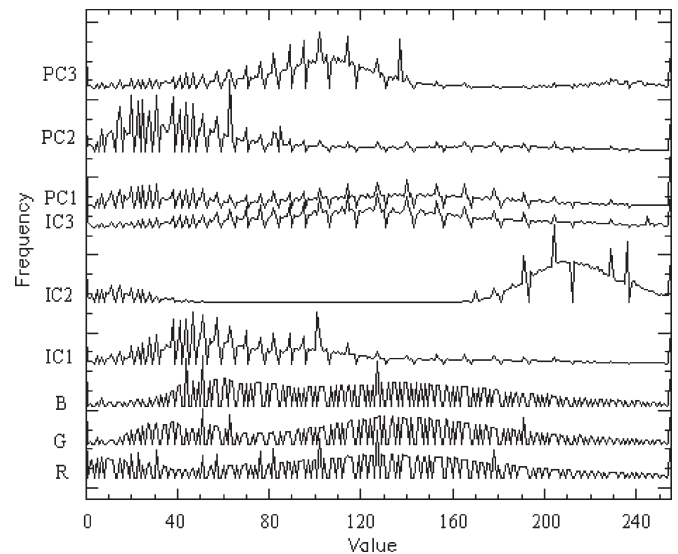

Fig. 15. Histograms of spectral features.

TABLE V
CLASSIFICATION ACCURACIES OF SPECTRAL FEATURES

Features	Water	Tree	Grass	Building	Baresoil	Road	Shadow	OA
RGB	0.925	0.576	0.576	0.153	0.854	0.879	0.973	0.605
PCA	0.925	0.788	0.536	0.616	0.712	0.825	0.969	0.765
ICA	0.925	0.813	0.563	0.605	0.740	0.823	0.971	0.778

analogous spectral features in the HSRRS image. Hence, other features are needed to discriminate them. In Chinese cities, buildings and roads have typical shape characteristics. Fig. 5 illustrates a sample QuickBird image of Beijing city. This experiment tested different values for the three parameters of PSI: T_1, T_2, and D, in order to detect their influence on the algorithm. There are three object types in Fig. 5: buildings, roads, and trees. Although buildings and roads are spectrally similar, they vary in their shapes. The numbers of training and testing pixels of the three types of objects are listed in Table I.

The parameter D indicating the total number of directions of PSI was detected first. As mentioned in Section II, when the value of D increases, the difference of PSI values for pixels in the same shape area decreases. In this experiment, the PSI features are regarded as an additional band of the three spectral RGB bands, the dimension of the input feature vectors is four, and the classifier used herein is the MLC as the performance of feature space is concentrated here. All the input features were preprocessed employing the two steps of normalization. The classification results with different D values are shown in Table II. As D increases, the classification accuracies of buildings and roads are also found to increase. Nonetheless, the accuracy of trees remains nearly at the same level as they have no apparent shape features; the classification results have not been influenced with the input of PSI. This test indicates that D is a parameter representing the capability of PSI for shape feature description, which can be demonstrated in Fig. 6.

Next, the parameter T_1 indicating the maximum spectral distance between the centric pixel and its surrounding pixels was detected. The value of T_1 should be adjusted in accordance with the different image features. If the homogeneous area or the same shape area displayed varying spectral information, the value of T_1 should be set to a larger number. As shown in Fig. 5, the building's roof facing to the sunlight presents a larger spectral value than the other sides with less sunlight, and the vehicles on the roads also influence T_1. This signifies that a lager value of T_1 may result in better classification accuracy. Fig. 7 proves this. Similar to the situation of D, when the value of T_1 ranges from 10 to 130, the classification accuracies of buildings and roads increase clearly, and the accuracy of trees remain almost the same.

Finally, the parameter T_2 indicating the maximum spatial distance between the centric pixel and its surrounding pixels was detected. T_2 value is critical in discriminating objects of different shapes. The relationship between the Kappa coefficient and T_2 is shown in Fig. 8.

When the value of T_2 ranges from 110 down to 30, the Kappa coefficients slowly escalate to the maximum value, and then Kappa decreases sharply to the minimum with the value of T_2
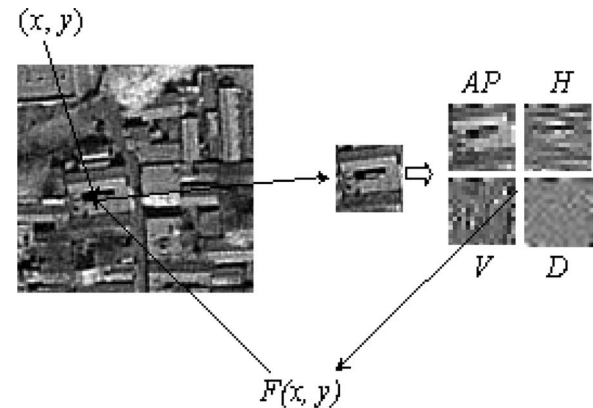

Fig. 16. Orientation-independent WT feature.

being from 30 to 5. It is noted that at 30, T_2 is approximately the average of the width and length of the buildings in Fig. 5. If T_2 is small to some extent, PSI features are not capable of discriminating the objects with different shapes. As Fig. 8 shows, the classification accuracies decrease sharply when the value of T_2 ranges from 30 to 5.

Fig. 9 shows three classification maps of Fig. 5. Fig. 9(b) displays the result using spectral and PSI features with $D = 20$, $T_1 = 100$, and $T_2 = 30$, and Fig. 9(c) the result using spectral and PSI features with $D = 20$, $T_1 = 100$, and $T_2 = 5$.

Fig. 9(a) demonstrates the obvious misclassification of spectral features, whereas the results in Fig. 9(b) and (c) show a considerable improvement over the spectral classification. Comparing Fig. 9(b) and (c), it is apparent that the parameter T_2 is also a scale factor of PSI. A larger value of T_2 [Fig. 9(b)] can identify larger objects but overlooks some smaller objects. A smaller value of T_2 [Fig. 9(c)] can identify smaller objects such as the narrow roads between buildings, whose width approximately equals T_2, about five pixels. Another dataset was added here to discuss the dependence of PSI to the local structural pattern and its multiscale effect. Fig. 10(a) shows a typical residential area in Beijing, consisting of low and dense houses in the old city area, and some communities with high buildings and green covers in the new developing region. The image size is 545×535 pixels, and it displays multiscale shape features of objects. The features employed here also include the RGB and PSI. The classification maps with different values of T_2 are listed in Fig. 10(b)–(d). T_2 was given considerable attention because it was the scale factor of PSI critical to multiscale objects. The relationship between the Kappa coefficient and T_2 is charted in Fig. 11, where similar results are observed: there are apparently misclassifications with only spectral signals as input; when T_2 increases from 3 to 10, the Kappa curve grows to the zenith, and after that the curve fluctuates to another sum-

TABLE VI
CLASSIFICATION ACCURACIES OF SPATIAL FEATURES WITH RGB BANDS

Features	Water	Tree	Grass	Building	Baresoil	Road	Shadow	OA
WT	0.925	0.665	0.586	0.376	0.760	0.857	0.964	0.685
GLCM	0.926	0.815	0.593	0.701	0.706	0.867	0.980	0.814
LWEA	0.925	0.755	0.446	0.705	0.634	0.719	0.930	0.773
PSI	0.926	0.847	0.400	0.881	0.484	0.788	0.969	0.852

TABLE VII
CONFUSION MATRIX OF THE CLASSIFICATION RESULT FOR FIG. 14(d)

	Water	Tree	Grass	Baresoil	Road	Shadow	Building	Total	%
Water	17179	12	0	0	0	1264	83	18538	92.7
Tree	0	13158	1017	0	5	194	1118	15492	84.9
Grass	1	573	969	19	0	0	36	1598	60.6
Baresoil	0	0	0	1666	0	0	549	2215	75.2
Road	0	35	3	39	10140	11	1743	11971	84.7
Shadow	30	8	0	0	1	2646	29	2714	97.5
Building	0	19	15	607	1838	46	20123	22648	88.9
Total	17210	13805	2004	2331	11984	4161	23681	75176	
%	99.8	95.3	48.4	71.5	84.6	63.6	85.0		

Overall accuracy = 87.6%, Kappa coefficient = 0.842

mit with $T_2 = 40$. Subsequently the curve descends gradually to a stable value, which remains above 0.56 even though T_2 ranges greatly from 70 to 200. Note the two T_2 values (10 and 40) corresponding to the two summits in the curve, which are just the sizes of low and high buildings, respectively, in the image. Accordingly, it can be found that PSI with an invariable T_2 cannot describe multiscale characteristics of objects at the same time; however, some multiscale profiles [10], [11] can be constructed to overcome this deficiency. This will be explored in our future work.

In the aforementioned curve, the maximum value is about 0.61, and the minimum is 0.56. Considering the Kappa coefficient of spectral features classification that is 0.48, it can be found that the appropriate parameter range to achieve an apparent accuracy increase is wide comparing to the results of the spectral classification.

B. Experiment2—Fusion of PSI and ICA

The purpose of this experiment is twofold. The first is to compare the classification results using different spatial features to demonstrate the effectiveness of PSI features. Second, it illustrates that the combination of the shape features extracted by PSI and the spectral features extracted by ICA can facilitate HSRRS data classification.

Another QuickBird image of the Beijing city was used. Fig. 12 shows the image (495 × 317 pixels), replete with available ground truth data. This image presents a typical urbanized area in Beijing, which includes a water body, long, and wide roads, low and dense buildings, bare soil, grass, and high trees. The key challenge in classifying this image lies in the confusion among roads and buildings, water and shadow, grass and trees, bare soil and buildings, and bare soil and roads and the uncertainties arising thereof.

In this experiment, the training samples were selected at random. The accuracy estimation was based on the reference pixels
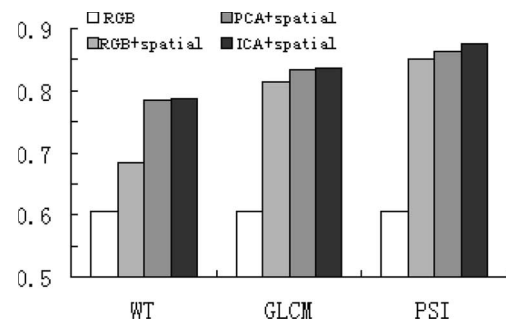

Fig. 17. Increase of accuracy from original RGB spectral bands, to RGB and the spatial information, PCA and the spatial information, and ICA and the spatial information using WT, GLCM, and PSI, respectively.

independent of the training data. The image was classified into seven classes, namely water, tree, grass, building, bare soil, road, and shadow. The numbers of training and testing pixels in each class are listed in Table III, and the available reference data are shown in Fig. 13.

The classification maps using different shape and spectral features are shown in Fig. 14. Fig. 14(a) shows many uncertainties in the classification map that employs only spectral features. These significant misclassifications are between objects of similar spectral signatures such as building–road, and tree–grass. This is also illustrated by the confusion matrix in Table IV. Fig. 14(b) shows the classification map with ICA replacing the original spectral bands. PCA and ICA can be regarded as transformed spectral features, which are utilized to reduce the complexity distribution and heterogeneity of original spectral bands. Three histograms of RGB, PCA, and ICA are compared in Fig. 15, where clear, stable, and less variable spectral distribution are observed from PCA and ICA images. Their classification accuracies are listed in Table V. Upon comparing the results of RGB and ICA, it is noted that the classification accuracies of grasses and trees are improved, hence contributing to an increased accuracy by 17%.

TABLE VIII
CLASSIFICATION ACCURACIES COMPARISON BETWEEN MLC AND SVM

classifier	Water	Tree	Grass	Building	Baresoil	Road	Shadow	OA
MLC	0.925	0.819	0.550	0.784	0.606	0.739	0.834	0.810
SVM	0.926	0.849	0.606	0.889	0.752	0.847	0.975	0.876

Both ICA and PCA can reduce the correlation among the different spectral bands, but the accuracy increase from PCA to ICA is only 1.3%. This is not apparent because only three spectral bands are available, and there is no obvious meaning of statistical independence in this case.

Due to the inadequacy of spectral information for discriminating different objects, spatial features were added to the classifier. For purposes of comparison, the spatial features extracted by GLCM and WT were also employed. The feature images of GLCM were acquired using a 7×7 window displacement of one pixel, and multiple texture measures such as variance, dissimilarity, mean, and contrast [1] were calculated based on the first primary component (PC1) of RGB bands. Four co-occurrence matrices in four directions were averaged, and so just one additional feature band of GLCM can be obtained. The WT feature band was also acquired based on PC1. Here, an 8×8 moving window was utilized, and the entropy measure was employed to extract the features of four subchannels [1]. For the purpose of fair comparison, an orientation-independent WT texture feature was constructed with the parameters illustrated in Fig. 16

$$F(x, y) = \frac{E(AP)}{E(D) + E(H) + E(V)} \quad (13)$$

where E is the entropy measure, $F(x, y)$ represents the WT feature value of pixel (x, y), and $E(AP)$, $E(V)$, $E(H)$, $E(D)$ are the entropy values of the four subimages obtained by the WT at level 1. The spatial features extracted using LWEA were then compared with PSI. The two-band length–width feature images obtained using LWEA were added into spectral bands as the input of the classifier [8].

All the spatial features were normalized using the (7) and (8), and these features were fused with RGB bands based on SVM. Their classification accuracies are listed in Table VI, which shows that the features of GLCM, LWEA, and PSI can supplement the spectral information and describe the structure of objects effectively, whereas WT is unable to.

The highest accuracy was achieved by PSI as its feature band can enhance the homogeneity of neighborhood using the contextual information and describe the shape feature effectively. The classification map of PSI and RGB based on SVM is shown in Fig. 14(c).

The high accuracy of classifying buildings indicates that PSI is a good shape descriptor; however, features based on PSI could not completely exploit the spectral information simultaneously. When PSI and original RGB bands were used together as input, the road–building classification of showed a promising improvement. Nevertheless, the results of tree–grass became worse. Consequently, some transformed spectral bands

TABLE IX
CLASSIFICATION ACCURACIES FOR SPECTRAL FEATURES OF FIG. 10(a)

	RGB	PCA	ICA
OA	0.676	0.730	0.735
K	0.432	0.486	0.493

obtained by PCA or ICA were introduced to attain better spectral classification accuracy. Additionally, this offers strengthened spectral information to reduce the complex distribution and heterogeneity in the original RGB bands and enlarge the contrast between objects and background. The classification map in Fig. 14(d) and the confusion matrix in Table VII reveal an improvement which maintains the accuracy of buildings while increasing those of bare soil and grass.

In this case, the accuracy is enhanced in two stages. In the first stage, the increase is due to the input of spatial features, and in the second stage, the increase is caused by the introduction of transformed spectral information such as PCA and ICA. These enhancements are shown in Fig. 17 with different combinations of spatial features.

For comparisons, MLC was employed as the fusion classifier of PSI and ICA features, and all the inputs were also normalized in advance. The accuracy statistics of Table VIII shows that SVM has a better capability of interpreting complex input features than MLC. This is due to the fact that the distribution of multiple features in the feature space may not be normal and SVM overcomes the defect of MLC where the dominant effects of spectral information may occur.

In order to verify that the integration algorithm based upon both PSI and ICA works stably, the image of Fig. 10(a) was also tested. First, the classification results for spectral features including RGB, PCA, and ICA are shown in Table IX, in which ICA and PCA really acquire the same accuracy. In this test, the accuracy improvement is not very apparent. The Kappa coefficients of PCA and ICA are higher than that of RGB by just 0.05. Subsequently, the spatial features were investigated. In this test, the classification results were obtained based on the respective spatial measures and the RGB bands fused by SVM. Comparable spatial features included the variance, mean, dissimilarity, and contrast measures of GLCM; the two bands LWEA features; the length and width bands that are separate features of LWEA [8]; and also, the orientation-independent WT feature. All the feature bands had to be normalized in advance. The classification statistics are listed in Table X, where the PSI obtains the best result. The accuracy increase from LWEA to PSI verifies that more direction lines can extract the contextual features more effectively. Considering that PSI has only one feature band but LWEA has two, so the proposed PSI is a meaningful extension of the latter.

At last, the integration of PSI and the transformed spectral bands based on SVM were tested. Results are shown

TABLE X
CLASSIFICATION ACCURACIES FOR SPATIAL FEATURES OF FIG. 10(a)

	mean	variance	dissimilarity	contrast	WT	width	length	LWEA	PSI
OA	0.796	0.754	0.799	0.786	0.762	0.749	0.800	0.803	0.825
K	0.577	0.523	0.578	0.561	0.528	0.509	0.580	0.588	0.615

TABLE XI
CLASSIFICATION ACCURACIES FOR THE INTEGRATION
BETWEEN PSI AND PCA OR ICA

	PCA+PSI	ICA+PSI
OA	0.831	0.835
K	0.625	0.632

in Table XI. Similar conclusion is drawn that the Kappa coefficient is slightly improved when ICA is employed.

VI. CONCLUSION

A shape feature index PSI, which presents a novel extension to LWEA, was proposed in this paper. PSI can effectively describe the contextual information through a simple method. This is especially true considering the huge quantity of HSRRS data. The proposed PSI leads to a notable increase of classification accuracy as compared to other spatial measures. The three parameters of PSI were also calibrated.

PCA and ICA have been utilized to extract spectral features in order to replace the original spectral signals in the classifications. In the experiments, ICA performed just slightly better than PCA because only a few spectral bands are available. However, it is exciting to find that the transformed spectral features can improve the classification accuracy with original bands for the multispectral images with the high spatial resolution.

In addition, an efficient algorithm capable of swiftly integrating spatial and spectral features using SVM has been also developed. This is based on the recognition of the fact that only spectral or spatial features are inadequate for classifying HSRRS data. The approach of a single classifier with multiple features was attempted and in this way, the input features should be normalized beforehand, and then SVM is employed to avoid the spectral dominance effect in the feature space. Our experiments showed that the combination of transformed spectral bands such as PCA or ICA and PSI based on SVM performed well. A single classifier cannot fully exploit the spectral and spatial features, yet it is a good approach providing that both computation time and accuracy are considered. While the use of PSI in conjunction with PCA or ICA and SVM improves the classification accuracy considerably, PSI also has some weakness. For example, it provides contextual information with a single scale; however, it could be overcome by the method of multiple profiles [10], [11] with different values of T_2. This will be investigated in our future work.

ACKNOWLEDGMENT

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] C. Zhu and X. Yang, "Study of remote sensing image texture analysis and classification using wavelet," *Int. J. Remote Sens.*, vol. 19, no. 16, pp. 3197–3203, Nov. 1998.

[4] 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.

[5] 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, 1999.

[6] Y. Dong, B. C. Forster, and A. K. Milne, "Segmentation of radar imagery using Gaussian Markov random field model," *Int. J. Remote Sens.*, vol. 20, no. 8, pp. 1617–1639, Oct. 1999.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] J. A. Benediktsson and I. Kanellopoulos, "Classification of multisource and hyperspectral data based on decision fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1367–1377, May 1999.

[13] G. J. Briem, J. A. Benediktsson, and J. R. Sveinsson, "Multiple classifiers applied to multisource remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2291–2299, Oct. 2002.

[14] M. Petrakos, J. A. Benediktsson, and I. Kanellopoulos, "The effect of classifier agreement on the accuracy of the combined classifier in decision level fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2539–2546, Nov. 2001.

[15] P. S. Huang and T. M. Tu, "A target fusion-based approach for classifying high spatial resolution imagery," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2004, pp. 175–180.

[16] 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.

[17] J. Li and R. M. Narayanan, "Integrated spectral and spatial information mining in remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 673–685, Mar. 2004.

[18] G. H. Halldorsson, J. A. Benediktsson, and J. R. Sveinsson, "Source based feature extraction for support vector machines in hyperspectral classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2004, pp. 536–539.

[19] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[20] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.

[21] G. M. Foody and A. Mathur, "A relative evaluation of multiclass image classification by support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1335–1343, Jun. 2004.

Liangpei Zhang received the B.S. degree in physics from Hunan Normal University, ChangSha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

From 1997 to 2000, he was a Professor in the School of the Land Sciences, in Wuhan University. In August 2000, he joined the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, as a Professor and Head of the Remote Sensing section. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence. He is also an Associate Editor of *Geospatial Information Science Journal*.

Prof. Zhang has served as Cochair of the International Society for Optical Engineers (SPIE) Series Conferences on Multispectral Image Processing and Pattern Recognition (MIPPR), the Conference on Asia Remote Sensing in 1999, Editor of the MIPPR01, MIPPR05 Symposium, and Chinese National Committee for the International Geosphere–Biosphere Programme.

Bo Huang is currently an Associate Professor in the Department of Geography and Resource Management, The Chinese University of Hong Kong. He was previously an Associate Professor in the Schulich School of Engineering, University of Calgary. His work focuses on the intersection of geospatial information systems (GIS), artificial intelligence, operations research, and spatial databases. He has authored and coauthored over 30 papers in refereed international journals and developed several GIS, spatial analysis, and optimization software packages. At present, he serves as a Guest Professor of the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China and the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China. His current research interests include spatial optimization, spatial statistics for change analysis, and image processing.

Xin Huang received the B.S. and the M.S. degrees from Wuhan University, Wuhan, China, in 2004 and 2006, respectively. He is currently working toward the Ph.D. degree at Wuhan University.

His current research interests include multi/hyperspectral and high-resolution remotely sensed image processing, spatial feature extraction, neural networks, pattern recognition, and remote sensing applications.

Pingxiang Li received the B.S., M.S., and Ph.D. degrees in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1986, 1994, and 2003, respectively.

Since 2002, he has been a Professor in the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. His research interests include photogrammetry and synthetic aperture radar image processing.