International Journal of Remote Sensing
Vol. 30, No. 12, 20 June 2009, 3205–3221

Taylor & Francis
Taylor & Francis Group

A comparative study of spatial approaches for urban mapping using hyperspectral ROSIS images over Pavia City, northern Italy

XIN HUANG* and LIANGPEI ZHANG
The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, P. R. China

(*Received 11 January 2008; in final form 27 August 2008*)

Urban mapping techniques using high spectral and spatial resolution (HSSR) data were investigated. To this aim, this paper proposes a novel mean shift (MS)-based multiscale method, and different spatial approaches are compared, including differential morphological profiles (DMPs), pixel shape index (PSI), the fractal net evolution approach (FNEA), and the proposed MS method. These spatial features were computed based on a dimensionally reduced representation that was obtained using the non-negative matrix factorization (NMF) transform. The support vector machine (SVM) was then used for classification. These algorithms were evaluated using two HSSR datasets that were obtained by using the Reflective Optics System Imaging Spectrometer (ROSIS) sensor over the urban area of Pavia, northern Italy. The results show that the spatial approaches can effectively complement the spectral features for urban mapping, and the proposed MS-based multiscale algorithm can give comparable or even better results than the FNEA, DMPs and other traditional algorithms.

1. Introduction

In recent years, high spectral and spatial resolution (HSSR) data have become available, and such data provide both detailed structural and spectral information. Most of the recent efforts for land cover/land use mapping in urban areas are related to HSSR imagery. Dell'Acqua *et al.* (2004) compared two spatial analysis techniques, the fuzzy ARTMAP with spatial reclassification (Gamba and Dell'Acqua 2003) and differential morphological profiles (DMPs) for urban mapping using HSSR data. In their experiments, improvements were obtained in terms of classification accuracies compared to pure spectral analysis of the dataset. Gamba *et al.* (2007) proposed an effective mapping procedure exploiting object boundaries, where the boundary and non-boundary pixels were classified separately and then geometrical constraints were enforced to tune the boundary map. The final result was obtained by fusing the two refined maps at the decision level. Akcay and Aksoy (2008) presented a novel method for automatic object detection in HSSR imagery by combining spectral information with structural information exploited by using image segmentation. Benediktsson *et al.* (2005) proposed a derivative of the DMPs for preprocessing of HSSR data. The DMPs were constructed based on the repeated use of openings and closings with a structuring element of increasing size. The morphological profiles were built based on several principal components and

*Corresponding author. Email: huang_whu@163.com

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160802559046

then classified using neural networks. The DMP approach can be extended by applying data fusion to the original data and the morphological information, after feature extraction. Some notable studies on the combination of spectral–morphological information have been provided by Joelsson *et al.* (2005), Palmason *et al.* (2006) and Fauvel *et al.* (2007).

Recent studies show that the exploitation of spatial information is necessary for classification of HSSR imagery, but few such approaches have been proposed (Benediktsson *et al.* 2005). The objective of this research was to propose a new spatial approach for urban mapping and undertake a comparative analysis with some traditional algorithms. In this study, the spatial approaches used can be divided into two categories. One is the spectral–spatial vector stacking method, including the pixel shape index (PSI) (Huang *et al.* 2007*a*) and the grey level co-occurrence matrix (GLCM). The two methods exploit the structural and shape information in the images to complement the spectral feature space. The other approach comprises multiscale techniques, including the fractal net evolution approach (FNEA) object-based analysis, the DMPs and the proposed mean shift (MS)-based approach. These multiscale approaches can effectively exploit the rich spatial information to extract objects with different shapes and sizes, and they have been applied successfully to very high resolution (VHR) imagery (Bruzzone and Carlin 2006, Huang *et al.* 2007*b*). The FNEA, embedded in the commercial software eCognition, is an object-based approach that has been proven to be more efficient than pixel-based methods for VHR image processing in many studies (Wang *et al.* 2004, Gao *et al.* 2006, Yu *et al.* 2006). The DMP algorithm has been used to extract multiscale structural features from HSSR data (Benediktsson *et al.* 2005, Joelsson *et al.* 2005, Palmason *et al.* 2006, Fauvel *et al.* 2007), and thus it is used as a benchmark to evaluate other algorithms. In this study, a novel multiscale MS analysis approach is proposed that computes the multiscale representation of HSSR data by using different spatial bandwidths. Among the aforementioned algorithms, both the PSI and the FNEA have been applied to VHR data classification (Wang *et al.* 2004, Huang *et al.* 2007*a*) but, to the best of our knowledge, they have not yet been applied to HSSR images. Therefore, it is worthwhile evaluating the performance of the PSI, the FNEA and the proposed multiscale MS approach for urban mapping using HSSR data.

Flow charts of the two spatial approaches are shown in figure 1, where (*a*) is the vector stacking approach and (*b*) shows the multiscale procedures. It is necessary to reduce the computation load for HSSR data when spectral and spatial information are considered at the same time. Principal component analysis (PCA) is a commonly used dimension reduction method and has been tested for HSSR data in some studies (Dell'Acqua *et al.* 2004, Benediktsson *et al.* 2005). In this research, the non-negative matrix factorization (NMF) algorithm (Lee and Seung 1999) was used to obtain characteristic images from the hyperspectral data because the NMF is distinguished from other methods by its use of non-negativity constraints. This is in contrast to other methods, such as PCA and vector quantization (VQ), that learn holistic, not parts-based, representations (Lee and Seung 1999). In the experiments, the performance of the NMF transform was evaluated and compared with the PCA and another data transform approach, the decision boundary feature extraction (DBFE), developed by Lee and Landgrebe (1993). After the spectral and spatial feature extraction steps, the support vector machine (SVM; Cortes and Vapnik 1995) was used to classify the spectral–spatial hybrid features as it is intrinsically less

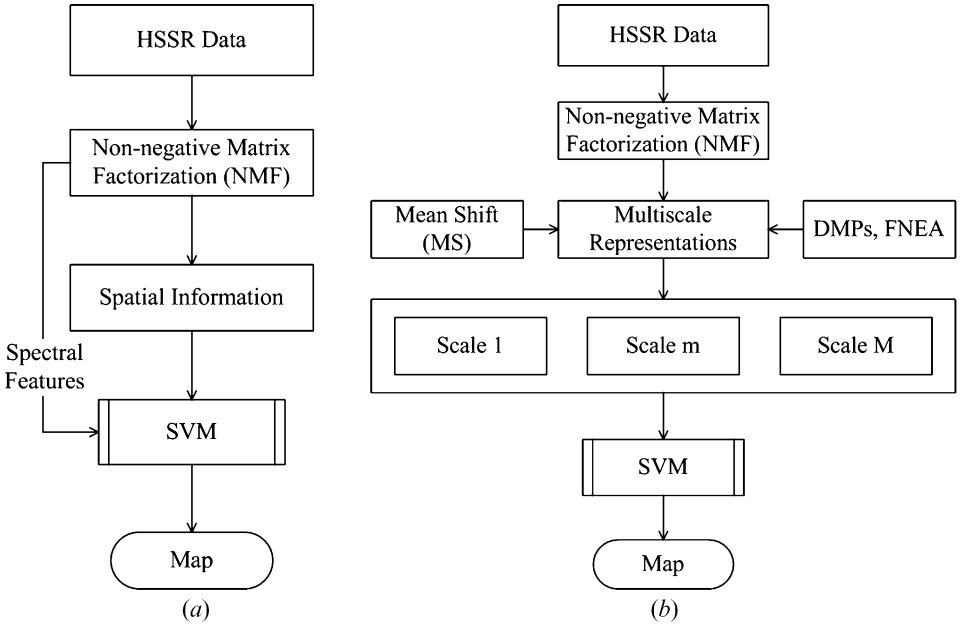


Figure 1. The flow charts of the two spatial approaches for urban mapping using the HSSR data: (a) the vector stacking, and (b) the multiscale approach.

sensitive to the high dimensionality of feature space, and is a non-parametric pattern recognition tool.

The rest of the paper is organized as follows: sections 2 and 3 describe the spectral–spatial vector stacking algorithms and the multiscale approaches, respectively. Section 4 concerns the analysis and comparison of the experimental results and in section 5 we present our conclusions.

2. Spectral–spatial feature stacking algorithms

2.1 *Grey level co-occurrence matrix (GLCM)*

The GLCM has been a standard technique for extracting texture features. It describes the second-order statistical relationships of two pixels in an image by the joint probability density function. In this research, GLCM texture statistics were calculated based on the NMF characteristic images with an interpixel distance of 1 and with different window sizes. The directionality of the GLCM was suppressed by averaging the extracted features over four directions, and the following four commonly used measures were chosen for the co-occurrence matrix:

$$\text{Homogeneity:} \quad \text{HOM} = \sum_i \sum_j \frac{P(i,j)}{1+(i-j)^2} \tag{1}$$

$$\text{Angular second moment:} \quad \text{ASM} = \sum_i \sum_j (P(i,j))^2 \tag{2}$$

$$\text{Entropy:} \quad \text{ENT} = -\sum_i \sum_j P(i,j)\log(P(i,j)) \tag{3}$$

$$\text{Dissimilarity:} \quad \text{DIS} = -\sum_i \sum_j P(i,j)|i-j| \tag{4}$$

where (i, j) are the coordinates in the co-occurrence matrix space and $P(i, j)$ is the co-occurrence matrix value at (i, j). Homogeneity is a measure of lack of variability or the amount of local similarity, and the angular second moment is also a measure of local homogeneity. Entropy and dissimilarity are heterogeneity indices. Entropy is a measure of the degree of disorder in an image, while dissimilarity represents the degree of spread of the grey levels or the average grey level difference between neighbouring pixels. Dissimilarity and homogeneity are inversely correlated.

2.2 *Pixel shape index (PSI)*

The PSI is an efficient spatial feature index for urban mapping with high spatial resolution (Huang *et al.* 2007*a*). PSI values are found by searching along a predetermined number of equally spaced lines radiating from the central pixel, namely direction lines. The extension of direction lines is based on the neighbouring grey level similarity in different directions. Upon completion of the calculation of all the direction lines, the direction-line histogram for each pixel x can be defined as,

$$H(x) = \{d_1(x), \ldots, d_n(x), \ldots, d_N(x)\} \tag{5}$$

where N is the number of direction lines for each pixel ($N=20$ in this paper) and $d_n(x)$ denotes the length of the nth direction line for pixel x. The shape index is the sum of lengths of all the direction lines, which is written as:

$$\text{PSI}(x) = \sum_{n=1}^{N} d_i(x) \tag{6}$$

In the previous literature, the PSI was considered as an additional feature band to complement the inadequacy of spectral information and hence improve the classification performance of VHR images (Zhang *et al.* 2006, Huang *et al.* 2007*a*). In this study, the PSI-NMF feature stacking approach was used for the HSSR data. The hybrid feature set can be written as $\{\text{NMF}_b\}_{b=1}^{B} + \text{PSI}$ where NMF_b represents the NMF-based spectral information at band b.

3. Multiscale classification approach

3.1 *Mean shift (MS)-based multiscale classification algorithm*

The MS procedure is a non-parametric density estimation technique, and its theoretical framework is the Parzen window-based kernel density estimation (Fukunaga and Hostetler 1975). Given n data points x_i ($i=1, \ldots, n$) in d-dimensional space, the kernel density estimator at point x can be written as:

$$\hat{f}_{h,K}(x) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^{n} k\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \tag{7}$$

where $c_{k,d}$ is a normalization constant, h is the bandwidth, and $k(\cdot)$ is the kernel profile that models how strongly the data points are taken into account for the estimation. The key step in the feature space analysis is to find the local maxima of the density $f(x)$, that is the modes of the density that are located among the zeros of the gradient $\nabla f(x)=0$.

The MS procedure is an efficient way to locate these zeros without estimating the density (Comaniciu and Meer 2002). The density gradient estimator can be obtained by differentiating equation (7) and decomposing into two product terms:

$$\hat{\nabla} f_{h,K}(x) = \frac{2}{h^2 c} \hat{f}_{h,G}(x)\, m_{h,G}(x) \tag{8}$$

where $g(x) = -k'(x)$ is the profile of kernel G with $c_{g,d}$ its normalization parameter and c the normalization constant $c = \frac{c_{g,d}}{c_{k,d}}$. In this equation, the first term $\hat{f}_{h,G}(x)$ is the density estimate at x with the kernel G, and the second term is the mean shift:

$$m_{h,G}(x) = \frac{\sum_{i=1}^{n} x_i\, g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)} - x \tag{9}$$

It can be found that the MS is the difference between the weighted mean, using the kernel G for weights, and x, the centre of the kernel. The MS is an efficient spatial feature extraction approach that is capable of delineating arbitrarily shaped clusters because of its nonparametric nature. This characteristic has found MS a potential segmentation and clustering tool for high-resolution remote sensing data. This study presents a multiscale MS-based classification approach for HSSR imagery, which consists of the following steps:

Step 1. Use the NMF transform to reduce the dimension of the original hyperspectral data because it can avoid the empty space phenomenon by which most of the mass in a high-Dimensional space is concentrated in a small region of the space. The characteristic images are written as $\{\mathrm{NMF}_b\}_{b=1}^{B}$.

Step 2. Define a set of bandwidths $\{h_m : 1 \leqslant m \leqslant M\}$ and run the mean shift segmentation algorithm (Comaniciu and Meer 2002). The multiscale representations of the NMF images can then be obtained by the mean shift procedures with different bandwidths. After the multiscale segmentation, the mean grey values of each segment are used to represent the features of each pixel in different scales. Therefore, the multiscale feature set is defined as $\left\{f_b^m : 0 \leq m \leq M\right\}$ where $f_b^0 = \mathrm{NMF}_b$ representing the feature of the pixel level.

Step 3. Compute the multiscale features for all the characteristic images and the multiscale MS-based feature set can be described as $f = \{f_b\}_{b=1}^{B}$ with $f_b = \left\{f_b^m\right\}_{m=1}^{M}$.

Step 4. The multiscale approach always leads to high-dimensional feature space, especially for HSSR data. Therefore, the SVM classifier, which is less sensitive to the high dimensionality of the feature space, is used to interpret the feature set f.

3.2 *Differential morphological profiles (DMPs)*

Opening and closing are two commonly used morphological operators. They are applied to an image with a set of a known shape, called the structuring element (SE). For features with regular shapes that are lighter/darker than the background, the basic approach is to apply openings/closings of increasing sizes. Benediktsson *et al.* (2005) proposed the DMPs that record image structural information. The structural information is collected by applying opening and closing operators with a multiscale approach and by looking at the residues between the multiscale morphologically transformed image and the original one.

Let γ_s and φ_s be the morphological opening and closing operators by reconstruction with structural element SE$=s$. MP_γ and MP_φ are the opening and

closing profiles, respectively, of the image I, and they are defined as vectors

$$\begin{aligned} \mathrm{MP}_{\gamma} &= \left\{ \mathrm{MP}_{\gamma_s} : \mathrm{MP}_{\gamma_s} = \gamma_s(I),\, s \in [0,\, S] \right\} \\ \mathrm{MP}_{\varphi} &= \left\{ \mathrm{MP}_{\varphi_s} : \mathrm{MP}_{\varphi_s} = \varphi_s(I),\, s \in [0,\, S] \right\} \end{aligned} \tag{10}$$

where $\mathrm{MP}_{\gamma 0} = \mathrm{MP}_{\varphi 0} = I$. The DMPs are defined as a vector where the measures of the slope of the opening–closing profiles are stored for every step of an increasing SE series:

$$\mathrm{DMP}_{\lambda} = \{ \mathrm{DMP}_{\lambda_s} : \mathrm{DMP}_{\lambda_s} = |\mathrm{MP}_{\lambda_s} - \mathrm{MP}_{\lambda_{s-1}}|,\, \lambda \in [\gamma,\, \varphi],\, s \in [1,\, S] \} \tag{11}$$

The DMP algorithm has been successfully applied to classification of both VHR (Benediktsson *et al.* 2003) and HSSR (Benediktsson *et al.* 2005) urban remote sensing data. It is worth noting that in the previous literature, DMPs were implemented on several principal component images that always led to high-dimensional feature space; therefore, some dimension reduction techniques were used for the morphological profiles. However, in this study, DMPs are built on the NMF images and a dimension-insensitive classifier, SVM, is used for DMP features without dimension reduction.

3.3 *Fractal net evolution approach (FNEA)-based multiscale object-based approach*

For high-resolution images, the increase in intraclass spectral variation may cause a reduction of statistical separability between classes with traditional pixel-based classification approaches, and object-based classification may be a good alternative (Yu *et al.* 2006). This is because the object-based approaches take advantage of the rich amount of local spatial information present in the irregularly shaped objects in images. The basic idea of this method is to group the spatially adjacent pixels into spectrally homogeneous objects, and in theory this will reduce the local spectral variation and the salt-and-pepper effect (Wang *et al.* 2004, Yu *et al.* 2006). Bruzzone and Carlin (2006) proposed a multilevel context-based system for classification of multispectral high-resolution images, where the FNEA segmentation algorithm embedded in eCognition was used to achieve the multiscale object-based classification. In this paper, this FNEA-based multilevel system is implemented on the HSSR images.

The FNEA is a bottom-up region merging technique starting from a single pixel. In an iterative way, adjacent objects are merged into new larger segments at each subsequent step. The region merging decision is defined as the heterogeneity difference between the new object and the constituent objects,

$$\Delta H = \sum_{b=1}^{B} W_b \left[N_{\mathrm{merge}} \delta_{\mathrm{merge}} - \left(N_{\mathrm{obj1}} \delta_{\mathrm{obj1}} + N_{\mathrm{obj2}} \delta_{\mathrm{obj2}} \right) \right] \tag{12}$$

where W_b controls the weight of band b ($1 \leqslant b \leqslant B$), N_{merge}, N_{obj1} and N_{obj2} represent the number of pixels within the merged object, object 1 and object 2, respectively. δ_{merge}, δ_{obj1} and δ_{obj2} are the corresponding standard deviations. When a possible merge of a pair of image objects is examined, the merge is performed when the criteria index ∇H is smaller than the scale parameter T (i.e. $\nabla H \leqslant T$). The segmentation process stops as soon as this condition is not met by any possible merge. The multiscale object-based representations can then be obtained by the FNEA segmentation with increasing values for T.

4. Experiments and analysis

4.1 *Experimental data and parameter setting*

The HSSR data used in this study were collected in the framework of the HySens project, managed by DLR (the German Aerospace Centre) and sponsored by the European Union. The images were acquired by the Reflective Optics System Imaging Spectrometer (ROSIS) sensor during a flight campaign over Pavia, northern Italy (45°11′ N, 9°9′ E), on 8 July 2002 from 10:30 to 12:00 h. Hyperspectral channels were collected in the 0.43–0.86 μm region of the visible and infrared spectrum with 1.3 m spatial resolution. Two datasets are used for evaluation of different algorithms. One is the Pavia city centre, with dense buildings, open areas and the river Ticino. The other is around the Engineering School at the University of Pavia. Some channels were removed because of noise; hence, the number of spectral bands was 102 for the 'Centre' data and 103 for the 'University'.

The parameters for the spatial approaches in the experiments are detailed as follows:

GLCM: the four measures (equations (1) to (4)) were combined with the spectral features in an SVM classifier. In the experiments, six window sizes were used: 3×3, 5×5, 7×7, 9×9, 11×11 and 13×13.

MS: the bandwidths were set to $h = \{0, 5, 10, 15, 20\}$, with $h = 0$ representing the pixel level, and the bandwidths were selected according to the spatial resolution and the object characteristics in the scene.

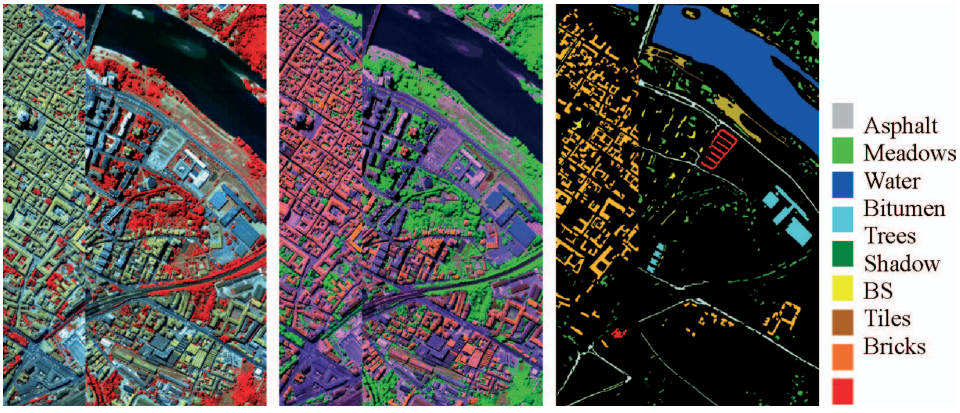
DMPs: the extended MPs use three characteristic bands ($B = 3$) and are based on a disc structuring element, a step of 2 with four openings and four closings (the number of openings/closings is written as NOC). The morphological profiles were then combined with the spectral bands, leading to a 27-dimensional (27-D) feature vector for NOC = 4.

FNEA: the scale parameters were set to $T = \{0, 5, 7, 10, 15\}$ for FNEA segmentation with $T = 0$ representing the pixel level. Note that for a fair comparison between MS and FNEA, their scale parameters should lead to similar average segment sizes (the average number of pixels for each segment).

4.2 *Experiments for the Pavia centre dataset*

The colour composite image of the Pavia centre data is shown in figure 2(*a*), with channel 82 for red, 46 for green and 9 for blue. The Pavia centre image was originally 1096×1096 pixels. A 381-pixel-wide black strip was removed, resulting in a 'two-part' 1096×715 image (figure 2). The 3-D NMF characteristic image is shown in figure 2(*b*), and the reference map is provided in figure 2(*c*). The numbers of training and test samples are listed in table 1. The training and test samples and the reference data were provided by the University of Pavia.

Figure 3 shows the hyperspectral signatures of different classes for the Pavia centre data. It can be seen that although the hyperspectral bands are available, some classes are still spectrally similar because they are made of similar materials. The spectrally confused or similar class pairs include water–shadow, asphalt–bitumen, tree–meadows, and brick–tiles–bare soil (BS). Therefore, the spatial information is exploited to complement the spectral feature space. Improvements in mapping accuracies can be expected when shape, texture and spatial coherence are considered or integrated with the spectral features.



(*a*) (*b*) (*c*)

Figure 2. (*a*) The RGB image for the Pavia centre dataset, (*b*) 3-D NMF bands, and (*c*) the reference map provided by the University of Pavia.

The class-specific accuracies for different spectral features are listed in table 2. It can be seen that the NMF characteristic images gave 95.6% overall accuracy (OA), higher than the original 102-D hyperspectral data (94.8%) and the 29-D DBFE features (94.5%), which showed that the NMF transform was efficient to represent the hyperspectral data although only 3-D features were extracted. The accuracies of different spatial approaches are provided in table 3, where the accuracy of the GLCM was obtained using a 13×13 window size as it resulted in the highest accuracy among all the windows tested. The 21-D and 27-D DMPs were produced by NOC=3 and NOC=4, respectively. The first comment on table 3 is that the proposed multiscale MS approach achieved the same and the highest OA with the 27-D DMPs (98.4%). This is a promising result considering that only 15-D features are available for the MS approach. From table 3 it can also be seen that: (1) all the spatial approaches (GLCM, PSI, DMPs, FNEA and MS) outperformed the spectral approaches (hyperspectral, DBFE, NMF), verifying the necessity of the exploitation of spatial information for HSSR data classification; (2) the multiscale algorithms gave higher accuracies than the feature stacking approaches. This phenomenon shows that the multiscale approach is a better spectral–spatial classification method and it can extract the rich spectral and spatial information from HSSR data more

Table 1. The training and test samples for the Pavia centre data.

Class	No. of training sets	No. of test sets
Water	824	65 971
Trees	820	7598
Asphalt	816	9248
Bricks	808	2685
Bitumen	808	7287
Tiles	1260	42 826
Shadow	476	2863
Meadows	824	3090
Bare soil	820	6584
Total	7456	148 152

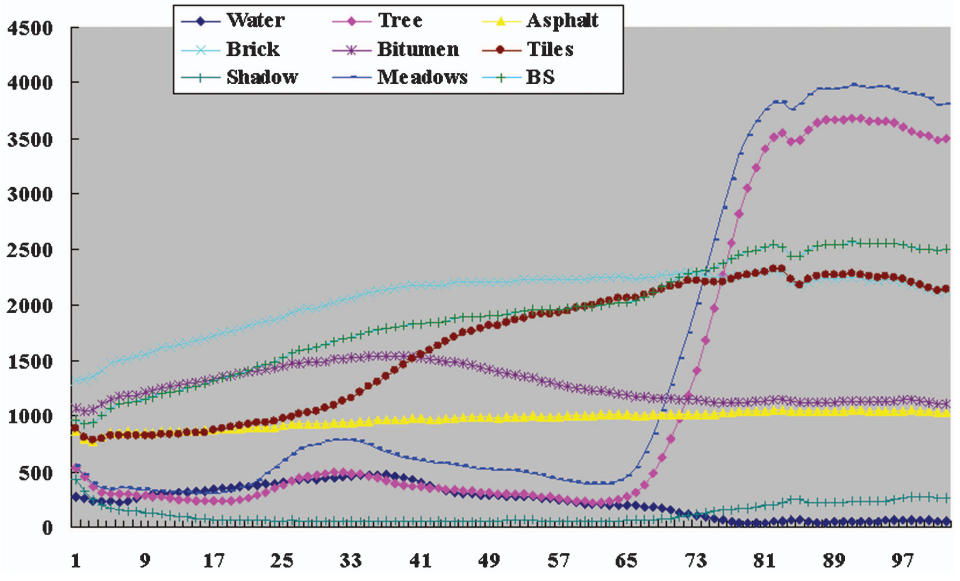


Figure 3. The 102-channel spectral signatures of different classes for the Pavia centre data.

effectively; and (3) all the multiscale algorithms gave satisfactory OAs (97.4% for FNEA, 97.9% and 98.4% for DMPs with NOC=3 and 4, respectively, and 98.4% for MS); in addition, they gave comparable and better class-specific accuracies for most of the classes compared with the purely spectral method.

Some classification maps with their subsamples are presented in figure 4. This figure shows that the spatial approaches can provide more accurate results for both quantitative statistics and visual inspection. The DMPs and MS algorithms avoided the salt-and-pepper effect in the pixel-level classification and they showed improvements in homogeneous regions such as roofs, shadows and roads. Figure 4(*d*) shows that some shadow pixels were wrongly identified as water for 27-D DMPs, and the same errors also occurred for 21-D DMPs. However, we did not observe the misclassification in the MS-based mapping. This phenomenon reveals that the multiscale MS algorithm can describe the shape and border more

Table 2. The class-specific accuracies (%) of different spectral feature sets for the Pavia centre data.

	Hyperspectral	DBFE	NMF
Feature	102-D	29-D	3-D
Water	92.1	91.5	98.6
Trees	92.9	92.0	79.0
Asphalt	93.7	94.4	94.9
Bricks	90.0	86.7	77.7
Bitumen	96.5	96.4	83.9
Tiles	99.3	99.3	98.7
Shadow	92.5	92.3	99.8
Meadows	98.3	97.7	90.9
Bare soil	96.0	95.6	86.5
OA (%)	94.8	94.5	95.6
Kappa	0.928	0.926	0.938

Table 3. Class-specific accuracies (%) of different spatial feature sets for the Pavia centre data.

	Spectral–spatial feature stacking		Multiscale method			
Class	GLCM (7-D)	PSI (4-D)	DMPs (21-D)	DMPs (27-D)	FNEA (15-D)	MS (15-D)
Water	99.4	99.1	99.0	99.1	98.5	99.1
Trees	89.6	81.8	87.9	89.4	86.5	88.6
Asphalt	91.4	94.8	97.5	99.3	97.9	99.5
Bricks	81.6	68.4	99.1	99.7	86.2	87.5
Bitumen	83.4	87.2	93.2	98.0	94.9	99.1
Tiles	99.3	99.0	99.3	99.4	99.7	99.3
Shadow	99.9	100.0	97.5	95.4	100.0	99.9
Meadows	89.2	91.3	91.0	91.0	95.0	98.6
Bare soil	89.2	88.8	99.2	99.8	90.2	97.7
OA	96.6	96.2	97.9	98.4	97.4	98.4

regularly and accurately, and it verifies the advantage of the MS approach for clustering with arbitrary shapes and textures.

To evaluate the proposed algorithms, the results of some existing approaches are reported in table 4, including (1) the spectral–morphological integration method (Palmason *et al.* 2006), which combines the spectral principal components and

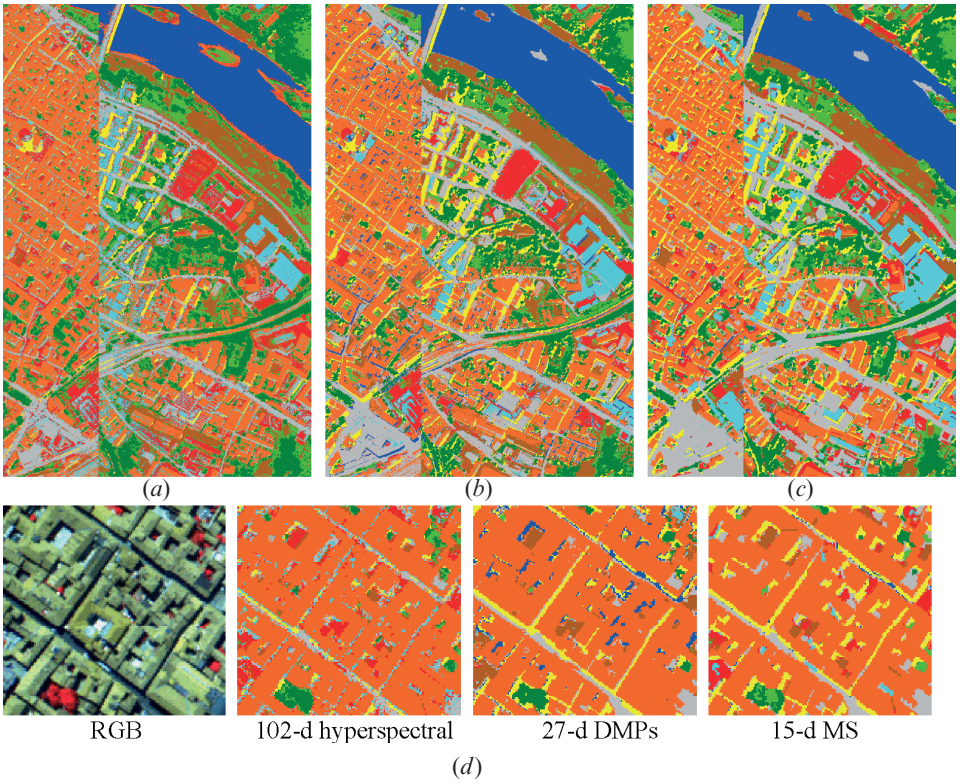


Figure 4. The classification results for (*a*) the 102-D data (94.8%), (*b*) DMPs (98.4%), and (*c*) the multiscale MS approach (98.4%). (*d*) Subsamples shown for a detailed comparison.

Table 4. Class-specific accuracies (%) of some existing algorithms for the Pavia Centre dataset.

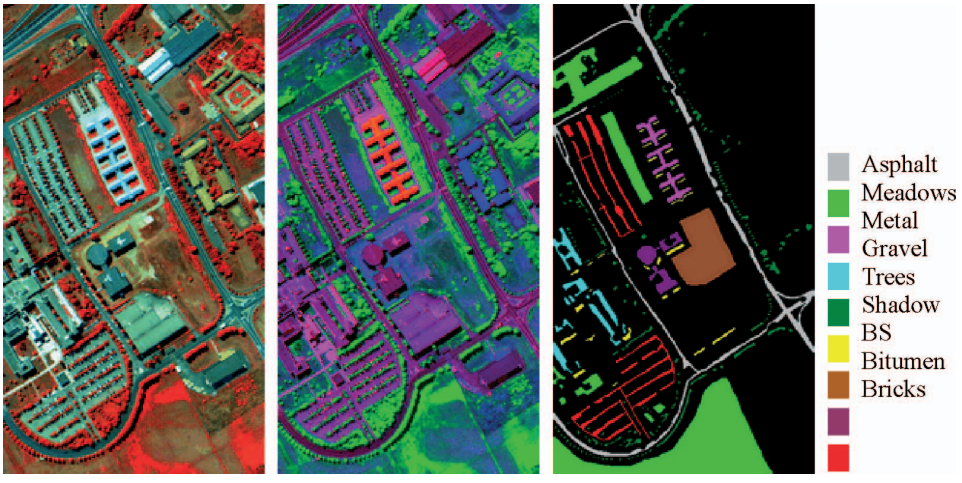
	Joelsson *et al.* 2005	Palmason *et al.* 2006	Fauvel *et al.* 2007
Class	Random forest	Spectral–morphological	Spectral–spatial classification
Water	99.3	99.8	99.2
Trees	89.6	82.4	90.1
Asphalt	94.1	95.6	97.8
Bricks	72.4	99.3	97.5
Bitumen	92.7	95.2	96.4
Tiles	97.8	96.4	99.3
Shadow	99.5	97.4	99.2
Meadows	97.5	94.4	91.7
Bare soil	87.5	92.0	96.1
OA	96.9	97.0	98.2

DMPs (NOC=4) using a neural network classifier (namely PCA-DMPs-NN), (2) the random forest classification (Joelsson *et al.* 2005), which uses a collection of tree-like classifier systems to analyse the hyperspectral data (namely RF-Hyperspectral) and (3) the spectral and spatial classification using SVMs and morphological profiles (namely spectral/spatial classifier) by Fauvel *et al.* (2007). The above approaches provided satisfied OAs (97.0%, 96.9% and 98.2%, respectively) that were slightly lower than those of the NMF-DMPs-SVM and the multiscale MS algorithms. By comparing tables 3 and 4, it can be seen that (1) the NMF transform is suitable for representing the HSSR data, and the SVM classifier is suitable for the hyperdimensional morphological profiles; and (2) the proposed multiscale MS algorithm is an efficient spectral–spatial classification approach in terms of the higher accuracy compared with the existing algorithms.

4.3 *Experiments for the University dataset*

The colour composite image and the 3-D NMF features for the University data are shown in figures 5(*a*) and 5(*b*). The reference map is provided in figure 5(*c*). The numbers of training and test samples are listed in table 5.

The hyperspectral signatures of different classes for the University data are shown in figure 6. It can be seen that the trees, shadow and metal sheets are easy to distinguish because of their reflectance characteristics, while the hyperspectral curves for bitumen–asphalt, gravel–brick and bare soil–meadows are very similar, and therefore it is difficult to discriminate them using spectral information alone. The accuracies for the spectral and spatial feature sets are listed in tables 6 and 7, respectively, where 'Ave' denotes the averaged accuracy of all the information classes. By analysing table 6, the 31-D DBFE features (99% criterion) gave the highest accuracy (77.9%), and the NMF transform provided comparable OA with the 103-D hyperspectral data (70.0% for NMF and 71.1% for hyperspectral data). Considering that only 3-D NMF bands were used and significantly outperformed the 3-D PCA bands (64.2%), it can be concluded that the NMF transform is an efficient dimensionality reduction algorithm for HSSR data. In table 7, the GLCM approach was implemented by combining the NMF bands and the four texture measures in equations (1) to (4). Three window sizes (9×9, 11×11, 13×13) were tested, and the



(*a*) (*b*) (*c*)

Figure 5. (*a*) The RGB image for the University dataset, (*b*) 3-D NMF bands, and (*c*) the reference map provided by the University of Pavia.

overall accuracies were 76.9%, 76.6% and 76.4%, respectively, and thus the class-specific accuracies of the 9×9 window are reported in the table. The first comment on table 7 is that the 27-D DMPs gave the highest OA (89.3%), and the multiscale MS obtained a comparable accuracy with the 21-D DMPs (82.7% and 83.9%, respectively). Table 7 also shows that the multiscale MS approach achieved satisfied averaged accuracy (90.4%), showing its ability for spectral–spatial information extraction from objects with different sizes and scales. It should be noted that it provided more accurate results than the spectral approach for all the classes except bare soil. Especially for the classes of bitumen, bricks and gravel, the MS approach gave obviously higher accuracies than other methods. The classification maps for 102-D hyperspectral data, 27-D DMPs and the 15-D multiscale MS algorithms are shown in figure 7.

From the above experiments, it is clear that the spatial information can complement the spectral feature space effectively and discriminate the classes with similar spectral responses. The averaged accuracies of different approaches for the three spectrally similar class pairs are listed in table 8. It can be seen that the

Table 5. The training and test samples for the University data.

Class	No. of training sets	No. of test sets
Trees	524	3064
Asphalt	548	6631
Bitumen	375	1330
Gravel	392	2099
Metal sheets	265	1345
Shadow	231	947
Bricks	514	3682
Meadows	540	18 649
Bare soil	532	5029
Total	3921	42 776

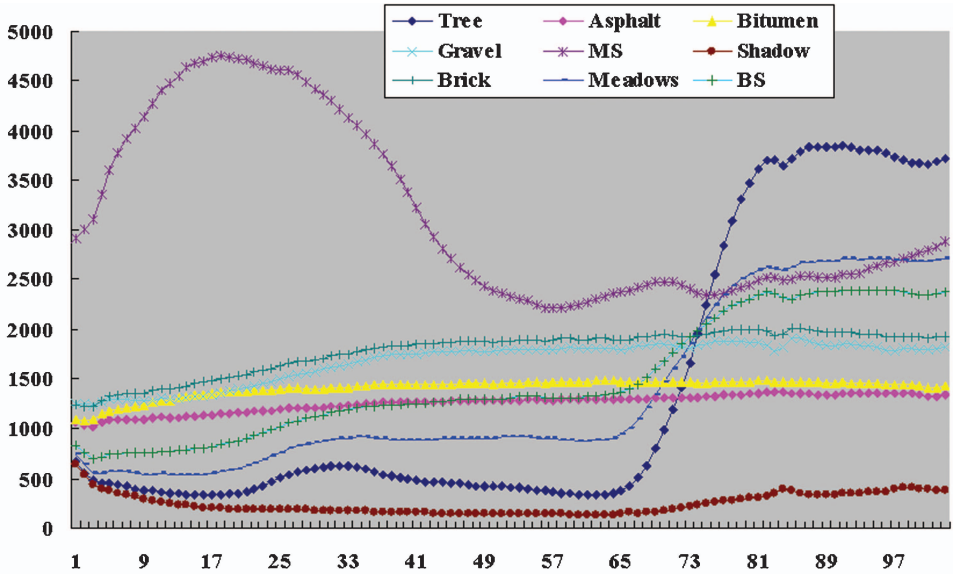


Figure 6. The 103-D spectral signatures of different classes for the University dataset.

multiscale MS technique showed good discrimination ability for bitumen–asphalt (94.5%) and gravel–brick (99.1%).

The experimental results were also compared with the two existing approaches: the spectral–morphological integration method of Palmason *et al.* (2006) and the spectral–spatial classification using SVMs and DMPs by Fauvel *et al.* (2007). Their overall and averaged accuracies are listed in table 9. Comparing tables 7 and 9, it can be seen that the multiscale MS approach gave higher averaged accuracy (90.4%) although the spectral–morphological algorithm and the spectral–spatial classification gave higher OA values. This observation shows that the MS algorithm is an efficient spectral–spatial feature extraction approach for classes with different sizes and scales.

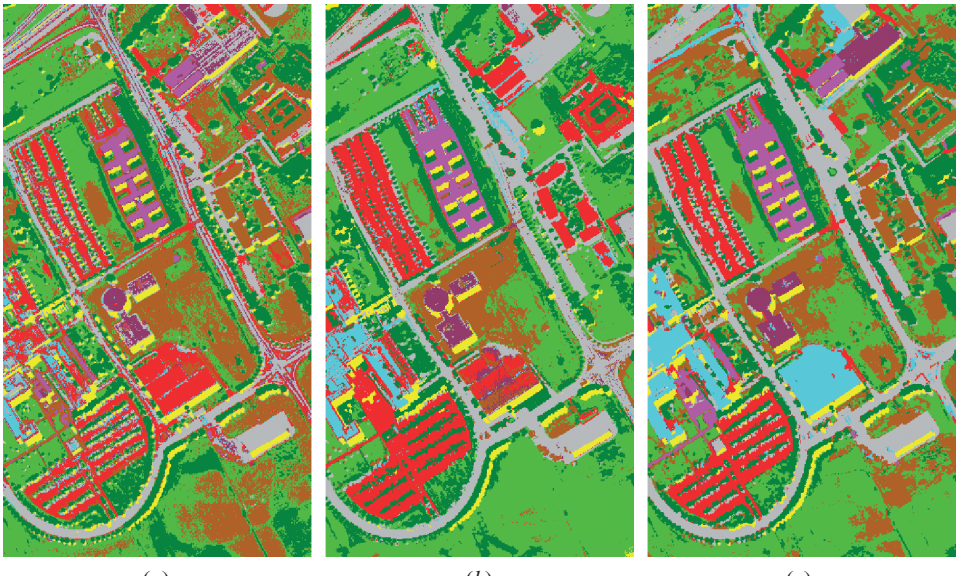
The two experiments show that the multiscale MS is similar to the DMP algorithm in its essence, which has been proved by the comparable performance on the test sets. Finally, the computational time for the two approaches was compared. The two

Table 6. Class-specific accuracies (%) of different feature sets for the University data.

	Hyperspectral	DBFE	PCA	NMF
Feature	103-D	31-D	3-D	3-D
Asphalt	70.9	72.9	68.3	69.8
Bare soil	67.9	91.2	41.0	32.0
Bitumen	83.5	82.0	68.4	80.2
Bricks	93.1	82.2	92.9	84.9
Gravel	43.5	65.5	24.1	34.2
Metal	98.6	100.0	98.9	99.4
Meadows	61.0	71.4	57.7	72.4
Shadow	99.6	90.8	99.9	99.8
Trees	98.3	97.6	98.6	98.9
OA	71.1	77.9	64.2	70.0
Ave	79.6	83.7	72.2	74.6

Table 7. Class-specific accuracies (%) of different spatial feature sets for the University data.

	Spectral–spatial feature stacking		Multiscale method			
Class	GLCM (7-D)	PSI (4-D)	DMPs (21-D)	DMPs (27-D)	FNEA (15-D)	MS (15-D)
Asphalt	80.1	80.4	79.9	93.6	75.8	90.3
Bare soil	49.1	54.0	57.0	64.9	35.4	50.2
Bitumen	67.5	71.3	93.8	97.1	94.4	98.6
Bricks	96.2	84.9	98.0	99.5	98.8	99.7
Gravel	43.4	58.1	72.5	77.0	65.5	98.4
Metal	99.9	99.5	99.7	99.7	99.1	99.8
Meadows	77.6	79.2	85.9	90.5	80.1	77.8
Shadow	92.3	97.8	98.2	98.2	99.8	99.6
Trees	99.7	99.0	99.5	98.5	98.7	98.8
OA	76.9	78.1	83.9	89.3	77.9	82.7
Ave	78.4	80.5	87.2	91.0	83.1	90.4



(*a*) (*b*) (*c*)

Figure 7. Classification maps for (*a*) the 103-D data (71.1%), (*b*) 27-D DMPs (89.3%), and (*c*) the multiscale MS approach (82.7%).

Table 8. Averaged accuracies (%) for the spectrally similar class pairs.

Feature	Bitumen–asphalt	Gravel–brick	Bare soil–meadows
103-D hyperspectral data	77.2	68.3	64.5
GLCM	73.8	69.8	63.4
PSI	75.9	71.5	66.6
21-D DMPs	86.9	85.3	71.5
27-D DMPs	95.3	88.3	77.7
Multiscale FNEA	87.6	82.2	57.8
Multiscale mean shift	94.5	99.1	64.0

Table 9. Class-specific accuracies (%) of some existing algorithms for the Pavia University dataset.

	Palmason *et al.* 2006	Fauvel *et al.* 2007
Class	Spectral–morphological	Spectral–spatial classification
Asphalt	88.3	90.9
Bare soil	48.2	85.3
Bitumen	90.5	95.2
Bricks	86.9	95.8
Gravel	52.2	57.9
Metal	98.7	99.5
Meadows	92.9	85.9
Shadow	92.6	95.1
Trees	98.5	99.2
Overall accuracy	85.6	88.0
Average accuracy	83.2	89.4

approaches were implemented on a Windows XP system with a 1.5 GHz processor and 512 MB RAM. The DMP and MS algorithms were implemented based on Matlab software, while the LIBSVM approach (www.csie.ntu.edu.tw/~cjlin/libsvm) was used for the SVM classification. For the Pavia centre dataset (1096×715), it took 163 s to generate the 27-D DMPs, while the SVM-based classification took 349 s. However, the 15-D MS feature extraction took 179 s and the time for SVM classification was 186 s. With respect to the University dataset (610×340), the 27-D DMPs calculation and its SVM classification took 32 and 165 s, respectively, while the CPU time for the 15-D MS extraction and classification was 52 and 99 s, respectively. From the above statistics, we find that the MS calculation took more CPU time than the DMPs, while the latter needs more computational time for classification because of its higher dimensionality.

5. Conclusions

This paper studied the detailed urban mapping techniques using spatial approaches based on HSSR imagery. The main contributions of this research are the following: (1) a multiscale mean shift approach was proposed for analysis of HSSR data, and the multiscale representations were obtained using different bandwidths; (2) the performance of the NMF transform for dimensionality reduction of hyperspectral data was evaluated; and (3) different spatial mapping approaches, including spectral–spatial vector stacking (GLCM and PSI) and multiscale approaches (FNEA, DMPs and MS), were compared and analysed for detailed urban mapping using HSSR imagery.

All of the algorithms were implemented and evaluated on the two ROSIS hyperspectral datasets over Pavia City, in northern Italy. By analysing the experimental results, we reached the following conclusions:

(1) The NMF transform is feasible for hyperspectral data reduction especially when the hyperdimensional data are transformed to a small-dimensional feature space. This conclusion is based on the observations that 3-D NMF features gave comparable or even higher accuracies than the 102-D hyperspectral image and the 29-D DBFE in the Pavia Centre dataset, and

they also gave comparable results to the 103-D hyperspectral data in the University dataset.

(2) In all of the experiments, the spatial approaches provided more accurate mapping than the spectral methods. It can be concluded that spatial information is necessary to complement the spectral features for high-resolution imagery although hyperspectral data are available. This is because some different objects are made of similar materials with similar spectral responses and the rich spatial information should be exploited to discriminate them.

(3) The multiscale approaches (FNEA, DMPs and MS) outperformed the feature stacking approaches (GLCM and PSI), demonstrating that the multiscale techniques can extract the rich spectral–spatial information from HSSR imagery more efficiently. For the multiscale approaches used in this paper, the DMP algorithm proved to be successful for urban feature extraction. In the experiments, four openings/closings (NOC=4) gave higher accuracies than three openings/closings (NOC=3). We also found that the SVM classifier is appropriate for the high-dimensional morphological profiles because of its dimensionality-insensitive character, and the NMF transform is a good alternative to produce basis images for morphological profiles.

(4) The proposed multiscale MS approach achieved satisfactory results in both experiments. In the Pavia centre dataset, it gave the same overall accuracy as the 27-D DMPs, while in the University dataset, it gave a slightly higher averaged accuracy than the 27-D DMPs. In addition, the multiscale MS clearly outperformed the multilevel FNEA algorithm with similar parameters. Therefore, it can be concluded that the mean shift is a potential analysis approach for spectral–spatial feature extraction.

Acknowledgements

Dr Huang Xin would like to thank Prof. Paolo Gamba and Prof. Fabio Dell'Acqua, University of Pavia, for kindly providing the ROSIS datasets and the ground reference data. This work was supported by the 863 High Technology Program of China under Grant 2007AA12Z148 and by the National Science Foundation of China under Grant 40771139.

References

AKCAY, H.G. and AKSOY, S., 2008, Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE Transactions on Geoscience and Remote Sensing*, **46**, pp. 2097–2111.

BENEDIKTSSON, J.A., PESARESI, M. and ARNASON, K., 2003, Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 1940–1949.

BENEDIKTSSON, J.A., PALMASON, J.A. and SVEINSSON, J.R., 2005, Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, **43**, pp. 480–491.

BRUZZONE, L. and CARLIN, L., 2006, A multilevel context-based system for classification of very high spatial resolution images. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, pp. 2587–2600.

COMANICIU, D. and MEER, P., 2002, Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, pp. 603–619.

CORTES, C. and VAPNIK, V., 1995, Support vector networks. *Machine Learning*, **20**, pp. 273–297.

DELL'ACQUA, F., GAMBA, P., FERARI, A., PALMASON, J.A., BENEDIKTSSON, J.A. and ARNASON, K., 2004, Exploiting spectral and spatial information in hyperspectral urban data with high resolution. *IEEE Geoscience and Remote Sensing Letters*, **1**, pp. 322–326.

FAUVEL, M., CHANUSSOT, J., BENEDIKTSSON, J.A. and SVEINSSON, J.R., 2007, Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. In *Proceedings of the Geoscience and Remote Sensing Symposium 2007 (IGARSS '07), 2007 IEEE International*, Barcelona, Spain, pp. 4834–4837 (Piscataway, NJ: Institute of Electrical and Electronics Engineers).

FUKUNAGA, K. and HOSTETLER, L.D., 1975, The estimation of gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory*, **21**, pp. 32–40.

GAMBA, P. and DELL'ACQUA, F., 2003, Improved multiband urban classification using a neuron-fuzzy classifier. *International Journal of Remote Sensing*, **24**, pp. 827–834.

GAMBA, P., DELL'ACQUA, F., LISINI, G. and TRIANNI, G., 2007, Improved VHR urban area mapping exploiting object boundaries. *IEEE Transactions on Geoscience and Remote Sensing*, **45**, pp. 2767–2682.

GAO, Y., MAS, J.F., MAATHUIS, B.H.P., ZHANG, X. and VAN DIJK, P.M., 2006, Comparison of pixel-based and object-oriented image classification approaches: a case study in a coal fire area, Wuda, Inner Mongolia, China. *International Journal of Remote Sensing*, **27**, pp. 4039–4055.

HUANG, X., ZHANG, L. and LI, P., 2007*a*, Classification and extraction of spatial features in urban areas using high resolution multispectral imagery. *IEEE Geoscience and Remote Sensing Letters*, **4**, pp. 260–264.

HUANG, X., ZHANG, L. and LI, P., 2007*b*, An adaptive multiscale information fusion approach for feature extraction and classification of IKONOS multispectral imagery over urban areas. *IEEE Geoscience and Remote Sensing Letters*, **4**, pp. 654–658.

JOELSSON, S.R., BENEDIKTSSON, J.A. and SVEINSSON, J.R., 2005, Random forest classifiers for hyperspectral data. In *Proceedings of the Geoscience and Remote Sensing Symposium 2005 (IGARSS '05), 2005 IEEE International*, Seoul, Korea, pp. 160–163 (Piscataway, NJ: Institute of Electrical and Electronics Engineers).

LEE, C. and LANDGREBE, D.A., 1993, Feature extraction based on decision boundaries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **15**, pp. 388–400.

LEE, D.D. and SEUNG, H.S., 1999, Learning the parts of objects by non-negative matrix factorization. *Nature*, **401**, pp. 788–791.

PALMASON, J., BENEDIKTSSON, J.A., SVEINSSON, J. and CHANUSSOT, J., 2006, Fusion of morphological and spectral information for classification of hyperspectral urban remote sensing data. In *Proceedings of the Geoscience and Remote Sensing Symposium 2006 (IGARSS '06), 2006 IEEE International*, Denver, CO, pp. 2506–2509 (Piscataway, NJ: Institute of Electrical and Electronics Engineers).

WANG, L., SOUSA, W.P. and GONG, P., 2004, Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. *International Journal of Remote Sensing*, **25**, pp. 5655–5668.

YU, Q., GONG, P., CLINTON, N., BIGING, G. and SCHIROKAUER, D., 2006, Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering and Remote Sensing*, **72**, pp. 799–811.

ZHANG, L., HUANG, X., HUANG, B. and LI, P., 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**, pp. 2950–2961.