Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Classification of hyperspectral urban data using adaptive simultaneous orthogonal matching pursuit

Jinyi Zou
Wei Li
Xin Huang
Qian Du

SPIE.

Classification of hyperspectral urban data using adaptive simultaneous orthogonal matching pursuit

Jinyi Zou,[a] Wei Li,[a,*] Xin Huang,[b] and Qian Du[c]

[a]Beijing University of Chemical Technology, College of Information Science and Technology, Beijing 100029, China

[b]Wuhan University, State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan 430079, China

[c]Mississippi State University, Department of Electrical and Computer Engineering, Mississippi 39762, United States

Abstract. Simultaneous orthogonal matching pursuit (SOMP) has been recently developed for hyperspectral image classification. It utilizes a joint sparsity model with the assumption that each pixel can be represented by a linear combination of labeled samples. We present an approach to improve the performance of SOMP based on *a priori* segmentation map. According to the map, we first build a local region where within-segment pixels are preserved while between-segment pixels are excluded. Hyperspectral pixels in the preserved region around the test pixel are then simultaneously represented by a linear combination of training samples, whose weights are recovered by solving a sparsity-constrained optimization problem. Finally, the label of the test pixel is determined to be the class that yields the minimal total residuals between the test samples and the approximations. Experimental results demonstrate that the proposed adaptive SOMP (ASOMP) is superior to some existing classifiers, such as the original SOMP and the recently proposed weighted-SOMP (WSOMP). For example, the ASOMP performed with an accuracy of 95.53% for the ROSIS University of Pavia data with 120 training samples per class, while SOMP obtained an accuracy of 87.61%, an improvement of approximately 8%. © *2014 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JRS.8.085099]

Keywords: hyperspectral image; sparse representation; segmentation; simultaneous orthogonal matching pursuit.

Paper 14163SS received Mar. 22, 2014; revised manuscript received Jun. 6, 2014; accepted for publication Jun. 11, 2014; published online Jul. 18, 2014.

1 Introduction

With the rapid development of sensing technology, hyperspectral imagery (HSI) obtained by remote sensing systems has been investigated in many practical applications.[1,2] HSI is a specific image which records more spectral information than the common RGB pictures. Usually, HSI covers hundreds of spectral channels.[3] With a large number of spectral bands, HSI provides wealthy information to distinguish different physical materials and objects.[4] With spatial resolution being improved, HSI can provide more accurate classification for heterogeneous urban image scenes.

In pattern recognition literatures, there are different types of classification methods. One of the most classic classifiers is based on the Bayesian theory,[5] which relies on the characteristic of statistical probability. The k-nearest-neighbor (K-NN) classifier[6] employs the Euclidean distance between training and testing samples and assigns the class label according to the most frequent class label in the k-nearest range. In recent years, nonlinear neural networks (NN)[7] have also been applied for classification tasks. One of the most important and famous classifiers is support vector machine (SVM),[8] which provides superior and stable performance in HSI classification by learning an optimal decision hyperplane to best separate the training samples using a kernel function to map

*Address all correspondence to: Wei Li, E-mail: liwei089@ieee.org

the nonlinear samples into a high-dimensional feature space. The key part of the SVM is the kernel function.[9] Some popular kernels include linear, polynomial, and radial basis function (RBF), in which the RBF kernel function is the most widely used one for the SVM in HSI.[10]

In recent years,[11,12] combination of the spectral and spatial information into classification has drawn increasing attention for HSI classification. One strategy involves spatial features or spatial smoothness, such as morphological filtering, morphological leveling,[13] Monte Carlo optimization, and Markov random fields,[14] which has offered better accuracy compared to using pixel-wise features only. Another strategy is to use a composite kernel (CK)[15] where both spectral and spatial information are added into kernel functions to generate a new kernel. In Refs. 3 and 16, combination of a segmentation map with SVM classification results through majority voting was proposed. In this method, several segmentation or clustering approaches[17] were employed, such as k-means, expectation maximization,[18] and iterative self-organizing data analysis (ISODATA).[19,20] In Refs. 3 and 16, clustering-based segmentation of hyperspectral images was explored, and a taxonomy and survey of clustering techniques can be found in Ref. 21. Multinomial logistic regression[22] is another widely used classification method which employs the logistic function to provide the posterior probability. A fast algorithm for sparse multinomial logistic regression has been developed in Ref. 23. In Ref. 24, a generalized CK framework was presented based on multinomial logistic regression, combining the spectral and the spatial information in HSI.

In Ref. 25, simultaneous orthogonal matching pursuit (SOMP), which relied on the observation that a hyperspectral pixel can be represented by a joint sparsity model (JSM) of a linear combination of training samples, has been developed for hyperspectral image classification. In SOMP, the sparsity of the input data with respect to a given overcomplete training dictionary is used, and hyperspectral pixels in a small neighborhood around the test pixel are simultaneously represented by a linear combination of training samples. The represented test pixels can be recovered by solving an optimization problem constrained by the sparsity level.[26,27] Finally, the label of the test pixel is determined to be the class that yields the minimal total residuals between the test samples and the approximations. SOMP has aimed at the similarity between neighboring pixels by adopting a JSM. With a JSM, SOMP was originally designed to exploit the spatial correlation across neighboring pixels. The algorithm is based on the assumption that neighboring pixels often consist of similar materials and share a common sparsity pattern, which makes the simultaneous sparse recovery possible. However, we observe that SOMP is based on a fixed region (e.g., 3×3, 5×5, etc.) of each test pixel that may also include between-class pixels especially for complex urban image scenes. In these areas, materials even within a small region may greatly change, which indicates the disadvantage of SOMP with a fixed window. In Ref. 28, a kernel-version of SOMP was discussed. In Ref. 29, a nonlocal weighted joint sparse representation classification method was proposed. In the weighted-SOMP method (denoted as WSOMP), for each pixel located in a local region, a similarity matrix between the central pixels and its neighboring pixels was calculated, and the matrix was viewed as nonlocal weights for further sparse representation.

In this context, we propose an adaptive approach to improve the SOMP algorithm based on *a priori* segmentation map. From the segmentation map, we exclude the pixels of different materials from the central pixel in the region. Then, we represent the pixels by a sparse coefficient vector and calculate the residuals between the original and the reconstructed pixels on a class-specific dictionary. This approach effectively avoids the aforementioned shortcomings of SOMP. In order to demonstrate its superiority, the proposed adaptive-SOMP (ASOMP) will be compared with the state-of-the-art approaches, such as the original SOMP and the recently proposed WSOMP, using real ROSIS hyperspectral data.

The remainder of this paper is organized as follows. In Sec. 2, the original SOMP, the WSOMP, and the proposed ASOMP are introduced. In Sec. 3, the experimental results are presented. Finally, concluding remarks are made in Sec. 4.

2 Proposed Method

2.1 *Simultaneous Orthogonal Matching Pursuit*

The SOMP (Ref. 25) classifier considers the fact that neighboring pixels tend to belong to the same class with high probability. Assume a given hyperspectral data set with training samples

$\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^{N}$ in R^d with M classes, where d is the dimension of the image and N is the total number of training samples. $\mathbf{X}^m$ is a matrix with a size of $d \times N_m$, where N_m is the number of available training samples for class m, and $\sum_{m=1}^{M} N_m = N$.

The SOMP classifier[25] is based on a JSM[26,30] to extract the contextual information in HSI. It is assumed that pixels for the same material in a region share a common sparsity pattern. Thus, the similar neighboring pixels can be sparsely represented by a linear combination of a few shared atoms. If test samples $\mathbf{y}_i$ and $\mathbf{y}_j$ are the adjacent pixels with similar spectral signatures, $\mathbf{y}_i$ can be presented as

$$\mathbf{y}_i = \mathbf{X}\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_{i,\lambda_1}\mathbf{x}_{\lambda_1} + \boldsymbol{\alpha}_{i,\lambda_2}\mathbf{x}_{\lambda_1} + \cdots + \boldsymbol{\alpha}_{i,\lambda_K}\mathbf{x}_{\lambda_K}, \tag{1}$$

where $\mathbf{X}$ is the dictionary of structured training samples and $\Lambda_K = \{\lambda_1, \lambda_2, \ldots, \lambda_K\}$ is the support of the sparse vector $\boldsymbol{\alpha}_i$. Thus, $\mathbf{y}_j$ should have a similar expression as

$$\mathbf{y}_j = \mathbf{X}\boldsymbol{\alpha}_j = \boldsymbol{\alpha}_{j,\lambda_1}\mathbf{x}_{\lambda_1} + \boldsymbol{\alpha}_{j\lambda_2}\mathbf{x}_{\lambda_1} + \cdots + \boldsymbol{\alpha}_{j,\lambda_K}\mathbf{x}_{\lambda_K}. \tag{2}$$

The JSM can be applied on a small region of pixels. Figure 1(a) illustrates a region with nine connected pixels, where $\mathbf{y}_1$ is the central pixel, and $\mathbf{y}_2$, $\mathbf{y}_3$, $\mathbf{y}_4$, $\mathbf{y}_5$, $\mathbf{y}_6$, $\mathbf{y}_7$, $\mathbf{y}_8$, and $\mathbf{y}_9$ are the eight neighboring pixels. We represent $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4, \mathbf{y}_5, \mathbf{y}_6, \mathbf{y}_7, \mathbf{y}_8, \mathbf{y}_9]$ as a set of spatial neighborhood pixels. According to the SOMP algorithm,[25] $\mathbf{Y}$ can be expressed as

$$\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2 \ldots \mathbf{y}_9] = [\mathbf{X}\boldsymbol{\alpha}_1, \mathbf{X}\boldsymbol{\alpha}_2 \ldots \mathbf{X}\boldsymbol{\alpha}_9] = \mathbf{X}[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_9] = \mathbf{XS}, \tag{3}$$

where $\mathbf{S}$ is a sparse matrix with only K nonzero rows and it consists of the sparse vectors $\{\boldsymbol{\alpha}_i\}_{i=1,\ldots,9}$ corresponding to each pixels. It is assumed that each sparse vector $\{\boldsymbol{\alpha}_i\}_{i=1,\ldots,9}$ shares the same support Λ_K because the neighboring pixels are most likely similar.

In SOMP,[25] the sparse matrix $\mathbf{S}$ can be recovered by solving the following problem:

$$\min \|\mathbf{S}\|_{\text{row},0} \quad \text{subject to: } \mathbf{XS} = \mathbf{Y}, \tag{4}$$

where $\|\mathbf{S}\|_{\text{row},0}$ denotes the nonzero rows of $\mathbf{S}$. This optimization problem can reformulated by adding a slack tolerance,[25] described as

$$\tilde{\mathbf{S}} = \arg \min \|\mathbf{S}\|_{\text{row},0} \quad \text{subject to: } \|\mathbf{XS} - \mathbf{Y}\|_F \le \sigma \tag{5}$$

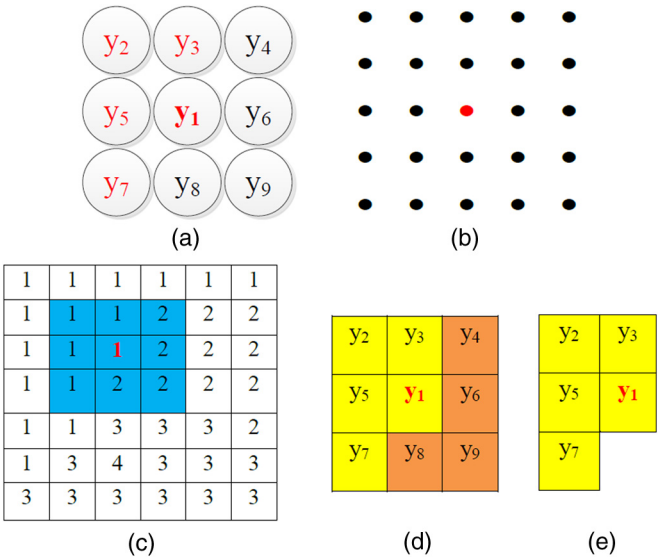

Fig. 1 (a) Nine connected regions, (b) 25 connected regions, (c) segmentation map with a patch of region size 3×3, (d) a patch of region size 3×3, and (e) chosen data for the central pixel from the patch of region.

or denoted as the equivalent optimization problem

$$\tilde{\mathbf{S}} = \arg\ \min \|\mathbf{XS} - \mathbf{Y}\|_F \quad \text{subject to: } \|\mathbf{S}\|_{\text{row},0} \leq K_0, \tag{6}$$

where K_0 denotes as the sparsity level and $\|\cdot\|_F$ denotes the Frobenius norm. This simultaneous sparse recovery problem is NP-hard and can be solved by OMP (Refs. 26 and 31) greedy algorithm.

Once the sparse matrix $\tilde{\mathbf{S}}$ is obtained, we can calculate the total error residuals between the original test samples and the recovered approximations obtained from class-specific dictionaries

$$r^m(\mathbf{Y}) = \|\mathbf{Y} - \mathbf{X}^m\tilde{\mathbf{S}}^m\|_F, \tag{7}$$

where $\tilde{\mathbf{S}}^m$ denotes the N_m rows corresponding to class m and the minimal total residual of the training dictionary will win this label of the central pixel $\mathbf{y}_1$,

$$\text{Class}(\mathbf{y}_1) = \arg \min_{m=1\ldots M} r^m(\mathbf{Y}). \tag{8}$$

2.2 *Weighted-SOMP*

For each central pixel, WSOMP adopted a similarity matrix as an adaptive weight for its neighboring pixels. Suppose in the nine connected neighboring pixels $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2 \ldots \mathbf{y}_9]$, $\mathbf{y}_1$ is the central pixel and others are neighboring pixels. For $\mathbf{y}_1$, the spectral similarity with y_i is defined as $\|y_i - y_1\|$, $i = 1, 2, 3, \ldots, 9$, and the weights coefficient is represented as

$$w(i, 1) = f(\|y_i - y_1\|), \tag{9}$$

where the function $f(\cdot)$ can be a Gaussian kernel. For two very similar samples, the weight in Eq. (9) is close to one; for two very dissimilar samples, the weight approaches zero. Once the weights of the surrounding pixels are obtained, a thresholding process is then applied to discard all the pixels whose weights are smaller than a preset threshold value w_1; for all the pixels whose weights are large than a threshold w_2, then their weights will be set to be a constant 1. The thresholding process can be defined as

$$W(i, 1) = \begin{cases} 0 & 0 < w(i, 1) \leq w_1 \\ w(i, 1) & w_1 < w(i, 1) \leq w_2 \\ 1 & w_2 < w(i, 1) \leq 1 \end{cases} \qquad i = 1, 2, 3, \ldots, 9. \tag{10}$$

Then Eqs. (5) and (6) are further described as

$$\tilde{\mathbf{S}} = \arg\ \min \|\mathbf{S}\|_{\text{row},0} \quad \text{subject to: } \|\mathbf{XS} - \mathbf{YW}\|_F \leq \sigma, \tag{11}$$

$$\tilde{\mathbf{S}} = \arg\ \min \|\mathbf{XS} - \mathbf{YW}\|_F \quad \text{subject to: } \|\mathbf{S}\|_{\text{row},0} \leq K_0, \tag{12}$$

where $\mathbf{W} = \text{diag}[W(1, 1), W(2, 1), \ldots, W(9, 1)]$ represents a diagonal matrix, and each diagonal element denotes the weight of the corresponding neighboring pixel to the sparse representation of the central test pixel. Obviously, a larger weight indicates more contribution in the representation, which is mainly from similar pixels according to Eq. (10).

Once the sparse matrix $\tilde{\mathbf{S}}$ is obtained, the total error residuals are calculated between the testing samples and the recovered approximations obtained from class-specific dictionaries

$$r^m(\mathbf{Y}) = \|\mathbf{YW} - \mathbf{X}^m\tilde{\mathbf{S}}^m\|_F, \tag{13}$$

and the minimal total residual of the training dictionary will be assigned to the label of the central testing pixel $\mathbf{y}_1$, which is represented as

$$\text{Class}(\mathbf{y}_1) = \arg \min_{m=1\ldots M} r^m(\mathbf{Y}). \tag{14}$$

2.3 *Adaptive-SOMP*

In SOMP, a greater window size is preferred for large homogeneous regions in HSI while a smaller window size is more appropriate for heterogeneous regions with complex distribution. Unfortunately, there is always a fixed window size for the current SOMP, which is an obvious weakness. The reason is that for a too-small window size, the spatial information cannot be fully utilized; for a too-big window size, it is inevitable that pixels in the block are confused among different materials. In this paper, we mainly investigate how to adaptively choose neighboring pixels to improve the performance of SOMP. Specifically, we improve the SOMP algorithm based on *a priori* segmentation map, from which within-segment pixels are preserved while excluding between-segment pixels. In other words, we first select a large window size, no matter if it is homogeneous or heterogeneous; and then, an *a priori* segmentation map facilitates the adjustment so that only samples with the similar spectral characteristics to the central pixel in this region will be maintained.

Here, we take an example of nine connected pixel regions as shown in Fig. 1(d). Figure 1(c) illustrates a part of the segmentation map of the data, which the same digital number (e.g., 1, 2, 3, 4) in the map indicates the similar material. In Fig. 1(d), $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$, $\mathbf{y}_5$, and $\mathbf{y}_7$ represent the pixels belonging to the same material and the rest are other materials. In the ASOMP, we exclude the pixels belonging to different materials from the region, and $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$, $\mathbf{y}_5$, and $\mathbf{y}_7$ are preserved as the new neighboring pixels $\tilde{\mathbf{Y}}$, which can be further expressed as

$$\tilde{\mathbf{Y}} = [\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_5, \mathbf{y}_7] = [\mathbf{X}\boldsymbol{\alpha}_1, \mathbf{X}\boldsymbol{\alpha}_2, \mathbf{X}\boldsymbol{\alpha}_3, \mathbf{X}\boldsymbol{\alpha}_5, \mathbf{X}\boldsymbol{\alpha}_7] = \mathbf{XS}. \tag{15}$$

This joint sparse recovery problem is NP-hard and can be approximately solved by greedy pursuit algorithm similar to the original SOMP

$$\min \|\mathbf{S}\|_{\text{row},0} \quad \text{subject to: } \mathbf{XS} = \tilde{\mathbf{Y}}, \tag{16}$$

where $\mathbf{S}$ is a sparse matrix with only few nonzero rows, and for each central pixel, the row of $\mathbf{S}$ is transformable according to the pixels in the chosen neighboring region. Similar to SOMP, this problem can be compactly presented as

$$\hat{\mathbf{S}} = \arg \min \|\mathbf{S}\|_{\text{row},0} \quad \text{subject to: } \|\mathbf{XS} - \tilde{\mathbf{Y}}\|_F \leq \sigma \tag{17}$$

or

$$\hat{\mathbf{S}} = \arg \min \|\mathbf{XS} - \tilde{\mathbf{Y}}\|_F \quad \text{subject to: } \|\mathbf{S}\|_{\text{row},0} \leq K_0. \tag{18}$$

After the sparse recovery problem is solved, the error residuals between the original test samples and the reconstructed pixels from class-specific dictionaries are calculated as

$$r^m(\tilde{\mathbf{Y}}) = \|\tilde{\mathbf{Y}} - \mathbf{X}^m\tilde{\mathbf{S}}^m\|_F, \qquad m = 1, 2, \ldots, M. \tag{19}$$

The label of the central pixel is determined to be the minimal residual of the class

$$\text{Class}(\mathbf{y}_1) = \arg \min_{m=1\ldots M} r^m(\tilde{\mathbf{Y}}). \tag{20}$$

.

The ASOMP demonstrates an adaptive feature extraction and pixel clusters process. Also for each pixel classification, the set of test samples is self-adapted by its neighboring region according to the presegmentation map. The segmentation map actually performs as the guided dictionary, from which dissimilar pixels are removed from the corresponding HSI region. In the proposed ASOMP, the segmentation map is obtained by the popular mean-shift clustering.[32]

3 Experiments and Analysis

The image scene used for experiments were recorded by the reflective optics system imaging spectrometer (ROSIS) over University of Pavia. The data consists of 610×340 pixels and the spatial resolution is 1.3 m. The number of spectral bands is 115 with the spectrum coverage range from 0.43 to 0.86 μm. After removing the 12 most noisy channels, 103 bands are preserved in our experiments. There are nine classes in the image scene, which are asphalt, meadows, gravel, trees, metals sheets, bare soil, bitumen, bricks, and shadows. We randomly select 120 training samples for each class from the ground truth map. Details on the training samples and the test data set are listed in Table 1. The second HSI data was recorded by the same sensor over the Center of Pavia during a flight campaign over Pavia, northern Italy.[12] The number of spectral bands is 102 for Pavia Center. The original data are a 1096×1096-pixel image. However, some of the samples contain no information and have to be discarded before analysis. So, the experimental image consists of 1096×492 pixels and each pixel has 102 spectral bands

Table 1 Training and test data for the University of Pavia data as well as the classification accuracies of three different classifiers (120 training samples per class).

Class		Samples		Accuracy (%)		
No	Name	Test	Train	SOMP	WSOMP	ASOMP
1	Asphalt	6631	120	69.33	85.69	90.08
2	Meadows	18,649	120	92.98	97.59	98.70
3	Gravel	2099	120	97.52	99.24	99.60
4	Trees	3064	120	69.91	87.83	89.69
5	Metal sheets	1345	120	99.93	96.95	71.23
6	Bare soil	5029	120	94.77	99.32	95.17
7	Bitumen	1330	120	97.74	99.70	99.62
8	Bricks	3682	120	91.34	95.71	99.08
9	Shadow	947	120	60.82	78.88	97.78
Total		42,779	1080	87.61	94.80	95.53

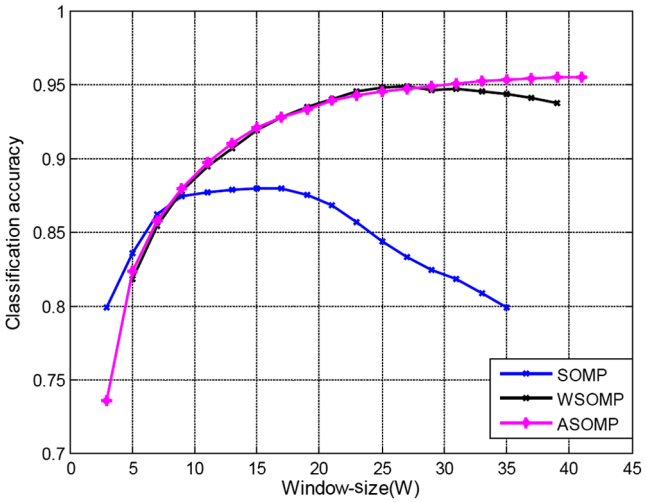

Fig. 2 Classification accuracy versus different window sizes for three classifiers (120 training samples per class) using the University of Pavia data.

after 13 noisy bands are discarded. There are also nine classes called water, trees, meadow, brick, soil, asphalt, bitumen, tile, and shadow, and more details are listed in Table 3.

We mainly compare the performance of the proposed method (i.e., ASOMP) with the original SOMP and the recently proposed WSOMP. As for parameter tuning, after our empirical study, the error tolerance σ is fixed to be 0.1 for all the SOMP, WSOMP, and ASOMP; in addition, the maximum iterations have been set to be 100 that is associated with the level of sparsity K_0.[25] For the WSOMP, the lower bound (i.e., threshold) in Eq. (10) is set to be 0.14 and the upper bound is set to be 0.88, which are the same as in the original paper.[29]

First of all, we investigate the classification performance of three classifiers with different window sizes using the University of Pavia data as illustrated in Fig. 2. For SOMP, it achieves the highest accuracy (i.e., 87.61%) when the window size is 15×15; nevertheless, it decreases when the window size becomes larger, which is due to the large window region containing many different materials. We observe that the WSOMP improves the classification performance due to the discarded pixels which are viewed as less correlated to the central pixels. The WSOMP achieves the best performance when the window size is 21×21. The best performance of the proposed ASOMP apparently outperforms the original SOMP, and it offers the accuracy of 95.53%, an

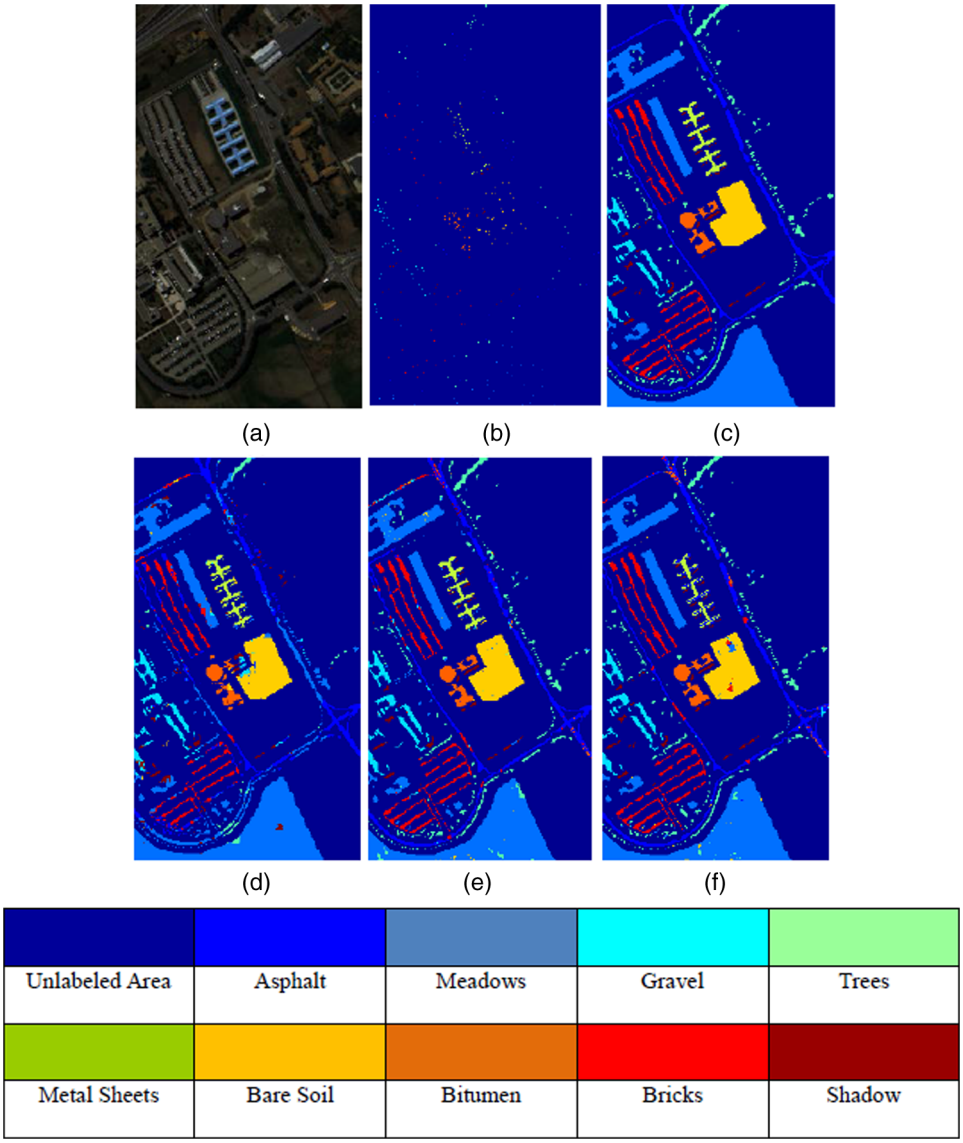

Fig. 3 For the University of Pavia data (120 training samples per class): (a) the false-color image, (b) train set, (c) test set, classification maps obtained by (d) SOMP, (e) WSOMP, and (f) ASOMP.

improvement of approximately 8% to the SOMP, and approximately 1% to the WSOMP. The improvement owes much to the benefit of the adaptively selected neighborhood. As mentioned before, a local region where within-segment pixels are preserved while excluding between-segment pixels has been selected. The selected region has some advantages: on one hand, it utilizes the spatial information in HSI; on the other hand, it guarantees that remaining data only contains one type of materials and keeps the purity even within a large window size. This is also the reason that the ASOMP holds the steady accuracy when the window size increases, which is very different from the other two classifiers. Figure 3 further illustrates the thematic maps resulting from these three classifiers. From Fig. 3, it is apparent that the classification map of the ASOMP is less noisy than those of SOMP and WSOMP, especially for bricks and shadow class, which is actually consistent to the results listed in Table 1. It is worth noting that for WSOMP, two thresholds have to be chosen empirically. When thresholding is not applied (that is to say, when the lower bound is 0 and the upper bound is 1), the method cannot

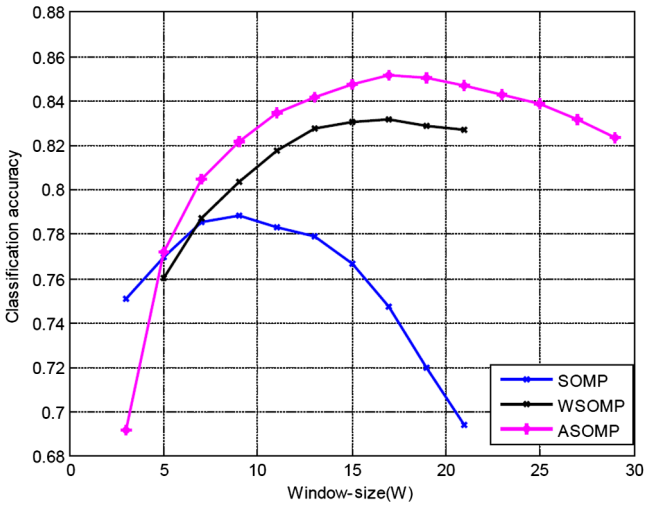

Fig. 4 Classification accuracy versus different window sizes for three classifiers (60 training samples per class) using the University of Pavia data.

Table 2 Training and test data for the University of Pavia data as well as the classification accuracies of three different classifiers (60 training samples per class).

Class		Samples		Accuracy (%)		
No	Name	Test	Train	SOMP	WSOMP	ASOMP
1	Asphalt	6631	60	50.04	57.67	60.94
2	Meadows	18,649	60	85.97	87.14	89.84
3	Gravel	2099	60	94.00	97.47	98.62
4	Trees	3064	60	73.53	85.70	89.13
5	Metal sheets	1345	60	99.41	97.10	69.67
6	Bare soil	5029	60	86.34	90.57	94.45
7	Bitumen	1330	60	96.17	96.77	98.05
8	Bricks	3682	60	76.45	86.94	89.30
9	Shadow	947	60	39.07	50.37	78.35
Total		42,779	540	78.83	83.16	85.60

demonstrate its superiority. Furthermore, it becomes suboptimal if two fixed thresholds are used for the whole scene. Comparatively, our method avoids this issue by directly removing the dissimilar pixels according to the segmentation map.

Next, we experiment on the same data while changing the number of training samples per class to be 60 as illustrated in Fig. 4. The proposed ASOMP still has the highest accuracy when the window size is 17×17, the optimal window size for the SOMP is 9×9, and the one for the WSOMP is 17×17. The improvement between the ASOMP and the WSOMP is consistently larger than 2% for different window sizes. Table 2 lists the classification accuracy for each class under optimal parameters for these three classifiers, and Fig. 5 illustrates the classification maps. We further evaluate the sensitiveness of the proposed method to varying numbers of training samples per class as shown in Fig. 6. When the number of training samples per class is low (e.g., 40), the ASOMP and WSOMP performs worse than the SOMP because too few training samples cannot construct a complete training subdictionary. If the subdictionary size is too small, the training samples may be insufficient to faithfully represent the subspace associated with each class, leading to a lower classification accuracy. Nevertheless, the proposed ASOMP is superior to the SOMP and the WSOMP when the number of training samples per class is larger.

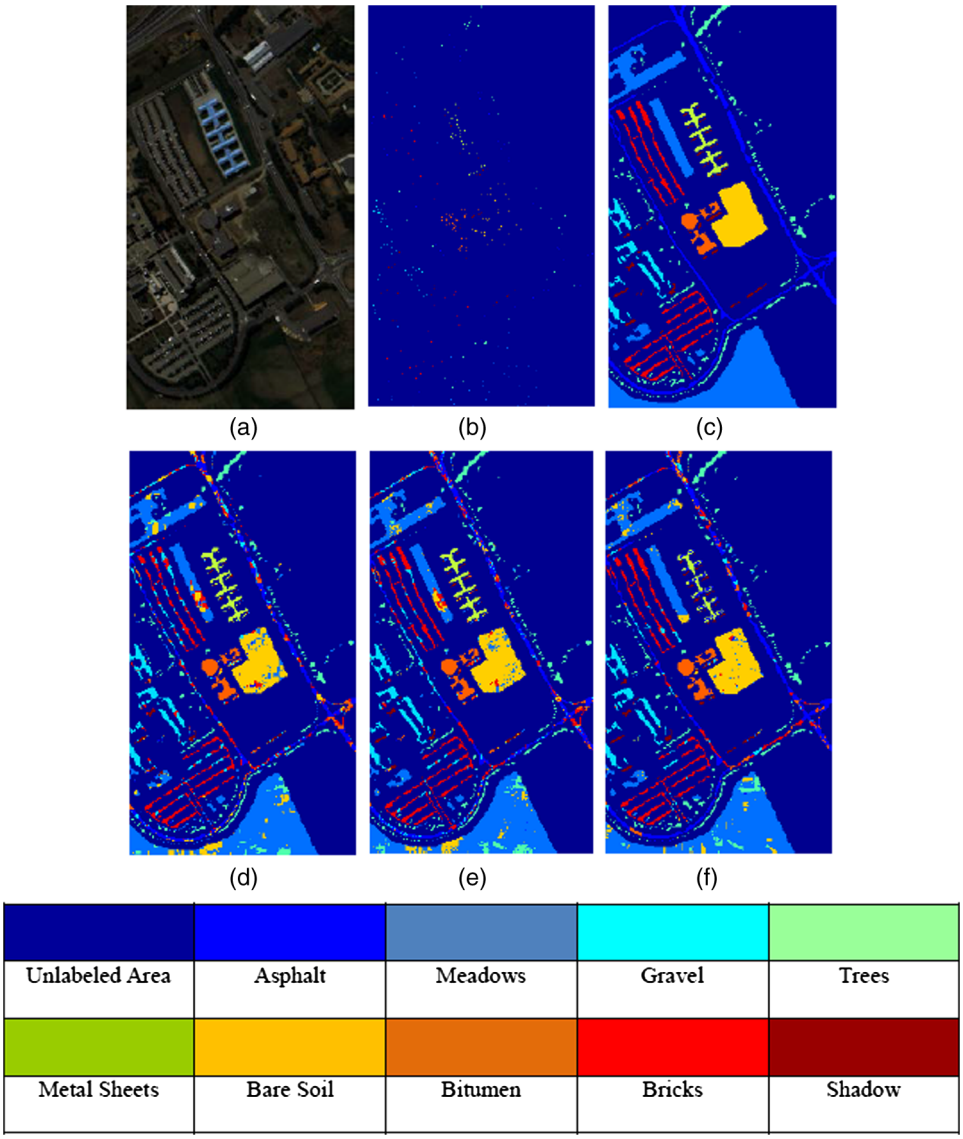

Fig. 5 For the University of Pavia data (120 training samples per class): (a) the false-color image, (b) train set, (c) test set, classification maps obtained by (d) SOMP, (e) WSOMP, and (f) ASOMP.

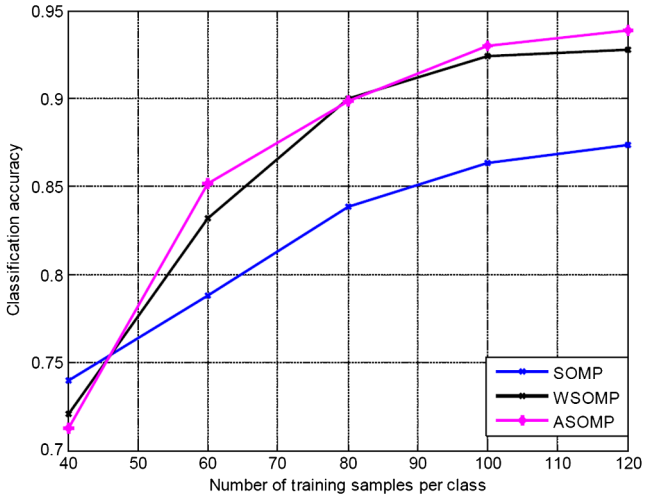

Fig. 6 Classification accuracy versus different numbers of training samples per class for three classifiers (for SOMP, $W = 9$; for WSOMP and ASOMP, $W = 17$) using the University of Pavia data.

Then, we experiment on the Pavia Center data as listed in Table 3, and the classification maps are illustrated in Fig. 7. With optimal parameters, the proposed ASOMP still outperforms the SOMP and the WSOMP. From the results in Fig. 7, it is apparent that classification maps of the ASOMP are less noisy than those of SOMP, especially for bitumen and trees class. In Table 3, there are six classes having higher class-specific accuracy of the proposed ASOMP. It is worth mentioning that the overall accuracy of the proposed ASOMP is not far more superior to the SOMP in this urban area than in the previous data. One of the reasons is that most of the regions in the University of Pavia scene are heterogeneous; however, the Center of Pavia data have more homogeneous regions, especially, the class water occupies nearly 66% of the total data set.

Finally, we provide the computational complexity of the SOMP and the proposed ASOMP. The computational complexity is upper-bonded by $O(TBNK_0)$ for both SOMP and ASOMP, and T is the number of the simultaneously represented pixels. Although for the same window size, the number of the simultaneously represented pixels in the proposed ASOMP is less than the

Table 3 Training and test data for the Center of Pavia data as well as the classification accuracies of three different classifiers (120 training samples per class).

Class		Samples		Accuracy (%)		
No	Name	Test	Train	SOMP	WSOMP	ASOMP
1	Water	65,278	120	99.96	99.95	99.85
2	Trees	6508	120	86.03	84.97	91.90
3	Meadow	2905	120	95.80	91.88	97.83
4	Brick	2140	120	93.88	91.64	96.26
5	Soil	6549	120	93.24	92.17	93.68
6	Asphalt	7585	120	79.67	84.13	77.35
7	Bitumen	7287	120	89.23	91.08	94.41
8	Tile	3122	120	99.26	98.85	99.90
9	Shadow	2165	120	91.36	91.18	87.71
Total		103,539	1080	95.98	96.12	96.54

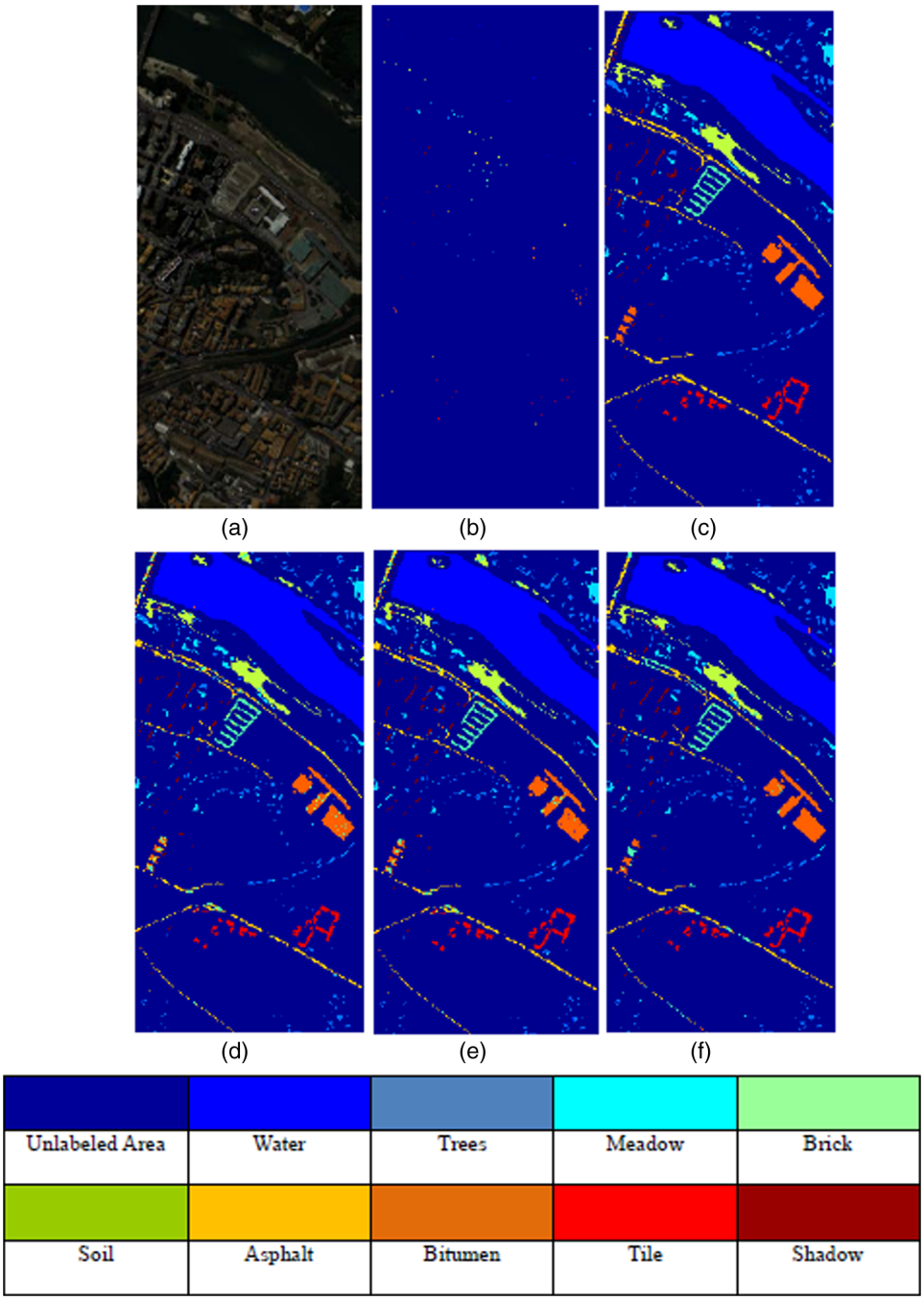

Fig. 7 For the Center of Pavia data (120 training samples per class): (a) the false-color image, (b) train set, (c) test set, classification maps obtained by (d) SOMP, (e) WSOMP, and (f) ASOMP.

SOMP. Then, we demonstrate the computing time of the two algorithms. All the programs are implemented in MATLAB on an Intel(R) Core(TM) 2 Duo CPU machine with 4 GB of RAM. Tables 4–6 list the computing time (in seconds) of the original SOMP and the proposed ASOMP with different window sizes. From these results, it can be seen that the computational speed of the proposed method is relatively faster than the SOMP, even with a large window size (e.g., $W = 15$). We note that with the same window size, the proposed ASOMP only requires the neighboring pixels similar to the central pixel for sparse representation; however, the SOMP sparsely represents the central pixel using all the pixels in the local region. Thus, for each iteration, the computational cost of the sparse matrix $\mathbf{S}$ in the ASOMP is much less than the SOMP due to the smaller size of the sparse matrix $\mathbf{S}$.

Table 4 Computing time (in seconds) of the proposed ASOMP and the SOMP under the same conditions (120 training samples) for the University of Pavia data with varying window sizes.

	$W = 3$ (s)	$W = 5$ (s)	$W = 7$ (s)	$W = 9$ (s)	$W = 11$ (s)	$W = 13$ (s)	$W = 15$ (s)
SOMP	477	1144	2443	4801	17,647	25,546	35,091
ASOMP	430	917	1742	3109	6350	11,487	16,494

Table 5 Computing time (in seconds) of the proposed ASOMP and the SOMP under the same conditions (60 training samples) for the University of Pavia data with varying window sizes.

	$W = 3$ (s)	$W = 5$ (s)	$W = 7$ (s)	$W = 9$ (s)	$W = 11$ (s)	$W = 13$ (s)	$W = 15$ (s)
SOMP	453	1202	2265	3651	5144	7601	13,327
ASOMP	258	494	1393	2239	3241	4521	5908

Table 6 Computing time (in seconds) of the proposed ASOMP and the SOMP under the same conditions (120 training samples) for the Pavia Center Data with varying window sizes.

	$W = 3$ (s)	$W = 5$ (s)	$W = 7$ (s)	$W = 9$ (s)	$W = 11$ (s)	$W = 13$ (s)
SOMP	10,922	9173	17,423	31,095	63,492	114,878
ASOMP	5991	7492	14,064	24,592	40,313	66,791

4 Conclusion

In this paper, a new ASOMP method was proposed for hyperspectral image classification. The proposed method combined *a priori* segmentation map with the sparse-representation-based classification to ensure that a testing sample be represented by its neighboring similar pixels only. The segmentation map actually performs as the guided dictionary, from which dissimilar pixels are removed from the corresponding local region. The experiments using hyperspectral ROSIS urban data demonstrated that our proposed ASOMP method yielded superior classification results compared to the original SOMP and the recently proposed WSOMP.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant No. NSFC-61302164.

References

1. W. Li et al., "Nearest regularized subspace for hyperspectral classification," *IEEE Trans. Geosci. Remote Sens.* **52**(1), 477–489 (2014).
2. D. A. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.* **19**(1), 17–28 (2002).
3. Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering technology," *IEEE Trans. Geosci. Remote Sens.* **47**(8), 2973–2987 (2009).
4. J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, Springer-Verlag, New York (1999).
5. H. J. Zhang et al., "Textual and visual content-based anti-phishing: a Bayesian approach," *IEEE Trans. Neural Networks* **22**(10), 1532–1546 (2011).

6. M. Hasanlou and F. Samadzadegan, "Comparative study of intrinsic dimensionality estimation and dimension reduction techniques on hyperspectral images using K-NN classifier," *IEEE Geosci. Remote Sens. Lett.* **9**(6), 1046–1050 (2012).
7. D. Muchoney and J. Williamson, "A Gaussian adaptive resonance theory neural network classification algorithm applied to supervised land cover mapping using multitemporal vegetation index data," *IEEE Trans. Geosci. Remote Sens.* **39**(9), 1969–1977 (2001).
8. F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.* **42**(8), 1778–1790 (2004).
9. N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines*, Cambridge University Press, Cambridge, UK (2000).
10. F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.* **42**(8), 1778–1790 (2004).
11. A. Plaza et al., "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.* **113**(Suppl. 1), S110–S122 (2009).
12. X. Huang and L. Zhang, "An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.* **51**(1), 257–272 (2013).
13. M. Fauvel et al., "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.* **46**(11), 3804–3814 (2008).
14. C. H. Chen "Hyperspectral image data unsupervised classification using Gauss–Markov random fields and PCA principle," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing Symposium*, Vol. 3, pp. 1431–1433, IEEE, Toronto, Canada (2002).
15. G. Camps-Valls et al., "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.* **3**(1), 93–97 (2006).
16. H. Yang, Q. Du, and B. Ma, "Decision fusion on supervised and unsupervised classifiers for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.* **7**(4), 875–879 (2010).
17. Y. Tarabalka et al., "Segmentation and classification of hyperspectral data using watershed," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing Symposium*, Vol. 3, pp. 652–655, IEEE, Boston, Massachusetts (2008).
18. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. B* **39**(1), 1–38 (1977).
19. G. Ball and D. Hall, "ISODATA, a novel method of data analysis and classification," Tech. Rep. AD-699616, Stanford University, Stanford, California (1965).
20. N. B. Venkateswarlu and P. S. V. S. K. Raju, "Fast ISODATA clustering algorithms," *Pattern Recognit.* **25**(3), 335–342 (1992).
21. A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Comput. Surv.* **31**(3), 266–323 (1999).
22. D. Bohning, "Multinomial logistic regression algorithm," *Ann. Inst. Stat. Math.* **44**(1), 197–200 (1992).
23. B. Krishnapuram et al., "Sparse multinomial logistic regression: fast algorithms and generalization bounds," *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(6), 957–968 (2005).
24. J. Li et al., "Generalized composite kernel framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.* **51**(9), 4816–4829 (2013).
25. Y. Chen and N. M. Nasrabadi, "hyperspectral image classification using dictionary-based sparse representation," *IEEE Trans. Geosci. Remote Sens.* **49**(10), 3973–3985 (2011).
26. J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: greedy pursuit," *Signal Process.* **86**(3), 572–588 (2006).
27. E. van den Berg and M. P. Friedlander, "Theoretical and empirical results for recovery from multiple measurements," *IEEE Trans. Inf. Theory* **56**(5), 2516–2527 (2010).
28. Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification via kernel sparse representation," *IEEE Trans. Geosci. Remote Sens.* **51**(1), 217–231 (2013).
29. H. Zhang et al., "A nonlocal weighted joint sparse representation classification method for hyperspectral imagery," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **PP**(99), 1–10 (2013).
30. S. F. Cotter et al., "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.* **53**(7), 2477–2488 (2005).

31. J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory* **53**(12), 4655–4666 (2007).
32. X. Huang and L. Zhang, "An adaptive mean-shift analysis approach for object extraction and classification from urban hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **46**(12), 4173–4185 (2008).

Jinyi Zou received his BS degree from Beijing University of Chemical Technology, Beijing, China, in 2013. He is currently pursuing the MS degree at the same university. His supervisor is Dr. Wei Li.

Wei Li received his PhD degree in electrical and computer engineering from Mississippi State University, Starkville, in 2012. Subsequently, he spent one year as a postdoctoral researcher at the University of California, Davis. He is currently with the College of Information Science and Technology at Beijing University of Chemical Technology, Beijing, China. His research interests include statistical pattern recognition, hyperspectral image analysis, and data compression.

Xin Huang received his PhD degree in remote sensing from Wuhan University, Wuhan, China, in 2009. Currently, he is a professor in the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing at the same university. His research interests include hyperspectral data analysis, pattern recognition, and remote sensing applications. In 2010, he received the Boeing Award for the best paper in image analysis and interpretation from the American Society for Photogrammetry and Remote Sensing.

Qian Du received her PhD degree in electrical engineering from University of Maryland, Baltimore County, in 2000. Currently, she is Bobby Shackouls professor in the Department of Electrical and Computer Engineering at Mississippi State University. Her research interests include hyperspectral remote sensing image analysis, pattern classification, data compression, and neural networks. She is currently the chair for the Remote Sensing and Mapping Technical Committee of International Association for Pattern Recognition.