

Received October 15, 2018, accepted October 29, 2018, date of publication November 12, 2018, date of current version December 18, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2879963

Unsupervised Band Selection of Hyperspectral Images via Multi-Dictionary Sparse Representation

FEI LI^{1,2}, PINGPING ZHANG¹, AND HUCHUAN LU¹, (Senior Member, IEEE)

¹School of Information and Communication Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China

²Engineering Training Center, Shenyang Aerospace University, Shenyang 110136, China

Corresponding author: Fei Li (alinafeili@mail.dlut.edu.cn)

This work was supported in part by the Scientific Research Foundation of the Liaoning Education Department under Grant L201605 and Grant L201627 and in part by the Aeronautical Science Foundation of China under Grant 2015ZB54007.

ABSTRACT Band selection is a direct and effective method to reduce the spectral dimension, which is one of popular topics in hyperspectral remote sensing. Recently, a number of methods were proposed to deal with the band selection problem. Motivated by the previous sparse representation methods, we present a novel framework for band selection based on multi-dictionary sparse representation (MDSR). By obtaining the sparse solutions for each band vector and the corresponding dictionary, the contribution of each band to the raw image is derived. In terms of contribution, the appropriate band subset is selected. Although the number of dictionaries is increasing, the efficiency of the algorithm is much higher than the previous due to the reduction of the dictionary self-learning process. Five state-of-the-art band selection methods are compared with the MDSR on three widely used hyperspectral datasets (Salinas-A, Pavia-U, and Indian Pines). Experimental results show that the MDSR achieves marginally better performance in hyperspectral image classification and better performance in average correlation coefficient and computational time.

INDEX TERMS Hyperspectral image, band selection, sparse representation.

I. INTRODUCTION

Hyperspectral data has both spectrum and spatial information of observed material, and the number of spectral bands usually reaches hundreds or even thousands [1]. This powerful advantage enhances and expands the potential application of hyperspectral imaging (HSI), including environment monitoring, precision agriculture, geological mapping, mineral exploration and so on. However, there is some redundant information in hyperspectral image cubes due to the spectral reflectance of most materials' changes only gradually over certain spectral regions and many contiguous bands are highly correlated [2]. In addition, some new challenging problems exist in the imagery processing owing to high spectral dimensionality, such as problem of storage and transmission, increment of computation, Hughes phenomenon (curse of dimensionality) [3]–[9]. Furthermore, for a given problem, not all spectral bands are useful and necessary in a hyperspectral image. Alternatively, a given spectral band may be a useful feature in a problem, but not for others [10]. Therefore, dimensionality reduction is a necessary method for efficient HSI applications to dispose of the above issues.

In general, there are two dimensionality reduction methods: feature extraction and band selection. The former usually generates a low-dimensional spectral data by using a transformation matrix based on a certain criteria. For instance, the transformation criterion can be principal component analysis (PCA) [11], linear discriminant analysis (LDA) [12], non-parametric weighted feature extraction (NWFE) [13], and so on. However, aforementioned methods are classical unsupervised or supervised feature extraction methods, and they inevitably change the physical meaning of the original bands. The latter is to identify the best subset of bands from all the original bands based on an adaptive selection criterion [10]. In this paper we focus on the band selection methods because they can preserve the spectral physical characteristics which are extremely significant in some applications.

According to whether the priori information is used, the band selection algorithm can also be divided into two categories: supervised and unsupervised. Supervised band selection methods require some prior knowledge about pixels' labels or objects' information [14]. For example, [15] derived minimum misclassification canonical analysis (MMCA);

Jeffries-Matusita distance and its extensions [16], contract divergence [17], class space separability [18], maximum likelihood classification accuracy [19], were adopted as the selection criteria; Du and Younan [20] proposed particle swarm optimization (PSO) based supervised band selection algorithm; Yang *et al.* [21] adopt sequential forward selection (SFS) search strategy method to select bands based on class spectral signatures only without performing classification; Cao *et al.* [22] integrated the local spatial information of the hyperspectral image into band selection method. Although the selected bands obtained by the supervised techniques are more effective than those obtained by unsupervised techniques in some specific applications, the necessary prior knowledge may be not easy to obtained in many cases. In this study, we focus on unsupervised band selection techniques because they are more practical than supervised ones regardless of any prior knowledge on pixels or objects to be analyzed. Concerning the unsupervised band selection technologies, many researchers have made a lot of efforts, and put forward some useful algorithms. Reference [14] compared a series of early unsupervised band selection methods, including information entropy, first and second spectral derivative, ratio, correlation and principal component analysis ranking based algorithms. Some literatures applied endmember extraction algorithms to band selection [23], [24]. In [23], the concept of NDINDR was applied to band selection and obtained promising results. Reference [24] proposed the application of similarity based endmember extraction algorithms: linear prediction (LP) and orthogonal subspace projection (OSP) for band selection, and improve their performance. Furthermore, [25] integrated graphics processing unites (GPU) parallel implementations with [23] and [24] to alleviate the computational burden. Sui *et al.* [26], [27] proposed an unsupervised band selection method which integrates the overall accuracy and redundancy into the band selection process, and it designed a balance parameter to trade off the overall accuracy and redundancy through an optimization model.

Recently, the idea of sparse representation (SR) was introduced into band selection [28]–[32]. There are mainly two groups, some algorithms represent the original image by multiplication of the dictionary and sparse coefficient matrix which is derived by K-SVD or orthogonal matching pursuit (OMP) [28], [31]; And other SR based methods pre-clustered basis bands with similar spectral properties [30], or pre-select bands using linear prediction (LP) algorithm and further refine them [29]. However, there are some issues existed in these BS methods. One is that the preprocessing will consume time, since it needs to pre-select or cluster the original bands. Another is that all dictionaries should be overcomplete dictionaries. But in the actual image, due to the length of the band vector is far greater than the number of bands, it is necessary to ensure that this is a part of the loss of image space information. Inspired from some previous work [28]–[34], in this paper we propose a novel unsupervised band selection algorithm

for hyperspectral image processing that utilizes the sparsity of the input sample vector. Our proposed algorithm is based on an effective sparse representation model in which the spatial image of a band is approximately represented by linear combinations of a few spatial image group (called atoms) from the entire hyperspectral image. The sparse vector, representing the atoms and associated weights for the input sample, can be obtained by solving an optimization problem constrained by the sparsity level and reconstruction accuracy. The specific bands can be determined by the value of associated weights which are not zeros under multiple given overcomplete dictionaries. We sort bands in descending order by the value of non-zero weights, and the first several bands are selected as the final reduced representation. In order to evaluate the performance of the proposed method, extensive comparisons against several unsupervised band selection methods are presented. These compared methods are well chosen to cover recent tendencies in this field. Experimental results on three widely used hyperspectral image classification datasets show that our proposed algorithm achieves superior performance and significantly outperforms other state-of-the-art unsupervised band selection methods.

In summary, our work has three main contributions:

(1) A novel multi-dictionary sparse representation method is proposed for hyperspectral band selection. The proposed method formulates the band selection as the sparse solutions for each band vector and the corresponding dictionary.

(2) We propose an effective method that can solve the sparse coefficients of multiple dictionaries and improve the computational efficiency.

(3) Extensive experiments on three benchmarks have validated the feasibility of our proposed methods, and show that our approach performs favorably against other state-of-the-art methods.

The rest of this paper is organized as follows. In Section II, we introduce the proposed unsupervised band selection method. Then we present the experimental results on some well known hyperspectral images in Section III. Finally, we make conclusions in Section IV.

II. SPARSE REPRESENTATION BASED BAND SELECTION

In recent years, sparse representation comes to be one of the hotspots in digital signal processing. In brief, sparse representation is the decomposition process of the original signal, and in this decomposition process the input signal will be represented as a linear approximation of the dictionary [35]. In the field of image processing, sparse representation has been successfully applied in image denoising [36], image restoration [37], [38], image recognition [39], [40] and so on. In the same way, the sparse representation method has also achieved great success in hyperspectral image processing, such as classification and target detection. In this section, we first introduce sparse representation based image processing algorithm that represents the sample through a sparse linear combination of samples from a dictionary. We then

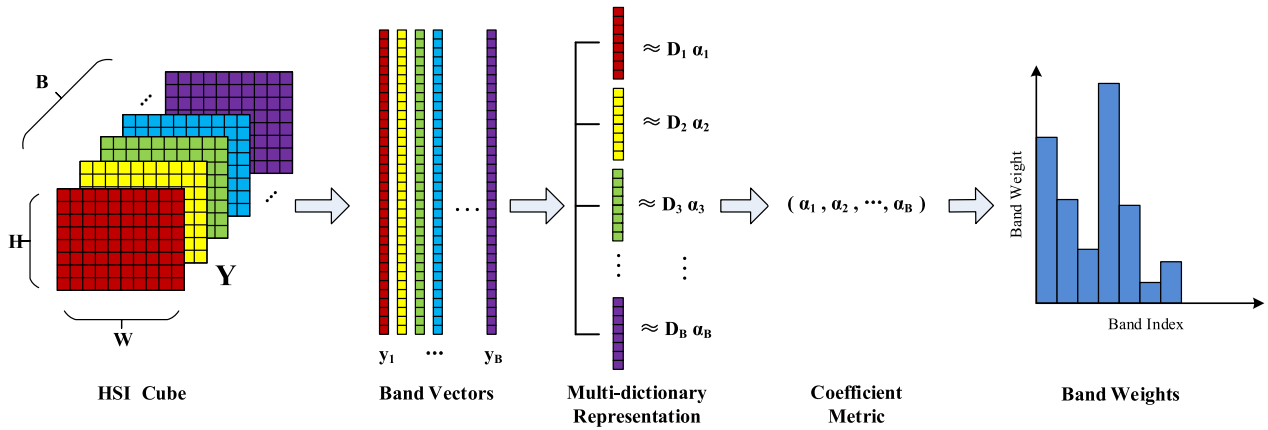


FIGURE 1. The schematic pipeline of the proposed band selection method. First, each band of the three-dimension hyperspectral data Y is stretched into a high-dimension vector. Then, the dimension of the vector is reduced by random sampling, so that the dimension of the vector is far less than the number of bands. Next, multi-dictionary learning is applied to the band vectors, and each band is approximately represented by the linear combination of other bands. Finally, the weight of each band is calculated by the sparse coefficient. The first several bands with large weights are selected as the final representation.

present the proposed multi-dictionary sparse representation method to solve the band selection problem.

A. SPARSE REPRESENTATION

The sparsity of signals has become a particularly powerful prior in many signal processing tasks, especially in the area of computer vision and pattern recognition [41]. Recently, sparse representation methods have been extended to the field of hyperspectral image processing, such as hyperspectral image classification [42]–[45] and hyperspectral object detection [46]. In methods of hyperspectral image classification, pixels are the performed objects which can be represented by a dictionary or basis with specific sparse constraints. A few coefficients derived from the representation can carry important information of the pixels, and can be treated as the feature of pixels. At last, a pixel can be signed a label by searching the minimum reconstruction error through the coefficients and dictionary elements. Formally, for the hyperspectral images classification problem, the sparse representation model can be casted as the following equation,

$$\begin{aligned}
 x &\approx A^1 \alpha^1 + A^2 \alpha^2 + \dots + A^M \alpha^M \\
 &= \underbrace{[A^1 \quad \dots \quad A^M]}_A \underbrace{\begin{bmatrix} \alpha^1 \\ \dots \\ \alpha^M \end{bmatrix}}_\alpha = A\alpha, \tag{1}
 \end{aligned}$$

where x is a pixel in a hyperspectral image, and A is the dictionary including M sub-dictionaries. α is the coefficient vector of this linear combination. With the selected sub-dictionaries as inputs, naive classifiers, like random forest [47], SVM [48], Bayesian network [49], can be used for further classification.

B. SPARSE REPRESENTATION MODEL FOR BAND SELECTION

The main purpose of band selection is to find an optimal or suboptimal band subset instead of the original

hyperspectral image, which can be used in sequential applications. In another word, the subset of bands is the set that can approximately represent the original bands under some measures or the subset of bands is the collection of bands which mainly contribute to the whole hyperspectral image. Hence, we should find out the contribution of each band to the whole image, and then choose the band according to its contribution. Sparse representation is an effective method to rank the contribution. When a band image is approximated by a linear combination of a dictionary, which consists of other band images, the sparse coefficients or weights will represent the contribution of each dictionary atom to the target band image. If the weight is large, the band will make great contribution to the target band, while if the weight is small, the band will make little contribution to the target band. We calculate sparse representation for each band by corresponding dictionaries and get a series of weights. The contribution of each band to the whole image can be obtained by statistical weights. As a result, the bands with larger weights are the selected bands.

When we describe the band selection problem of hyperspectral images, the spatial image of each band becomes the performed object. Figure 1 illustrates the schematic pipeline of the proposed band selection method. Let $Y \in R^{W \times H \times B}$ be the original hyperspectral image. The spatial image of every band is stretched into a real-value vector instead of a two-dimensional matrix, thus we get

$$Y = [y_1, y_2, \dots, y_B], \tag{2}$$

where $y_i \in R^L$ ($i = 1, 2, \dots, B$) is the image vector of the i -th band, B is the number of spectral bands, $L = W \times H$ is the number of pixels in the image, and W and H are the width and height of the band image, respectively. The purpose of band selection is to select a best subset from the original band set. The resulting number of bands in that subset is less than B . If there is a best subset which can be found from original band set, the subset can approximately represent the original band

set. We can use the following equation to present the relation,

$$\mathbf{y}_i \approx \mathbf{D}_i \alpha_i, \quad \mathbf{D}_i \subset \mathbf{Y}, \quad (3)$$

where \mathbf{D}_i is the band subset which is exclusive of \mathbf{y}_i and α_i is the linear combination coefficient for the i -th image vector on the subset \mathbf{D}_i . It should be noted that there is a trivial solution for the above equation. In the whole band set \mathbf{Y} , there is a column vector, e.g., \mathbf{y}_i , therefore, one can use the whole band set \mathbf{Y} as the basis, leading to the coefficient α_i is one, and others are zeros. As a result, the coefficient vector α will be an identity vector which lead to same weights for all band vectors. In order to avoid the case, we assign each band vector to a band subset \mathbf{D}_i (also called sub-dictionary). We will use the new sub-dictionaries \mathbf{D}_i hereafter. In addition, if we describe all bands by the defined band subset, the above equation can be represented as

$$\mathbf{Y} \approx \mathbf{D} \alpha, \quad (4)$$

where \mathbf{D} is the set $\{\mathbf{D}_i\}(i = 1, 2, \dots, B)$ and α is the set of coefficient weight α_i .

According to the previous expressions, we can convert every represented band vector as

$$\mathbf{y}_i = \mathbf{D}_i \alpha_i + \beta_i, \quad (5)$$

where β_i is the approximation error vector. By evaluating the value of α_i , we can determine the contribution of each band in \mathbf{D}_i to the target image vector \mathbf{y}_i . With sparse constraints on α , which means that if \mathbf{y}_i is independent with the j -th column of \mathbf{D} , the value of j -th element in α is zero, we can select the most important bands based on the approximation error. If we calculate the sparse representation of all bands in \mathbf{Y} , a sparse coefficient matrix will be obtained. Each band vector's weight is the sum of the corresponding row in the coefficient matrix. The best α_i value can be found by the following optimization problem. We will elaborate it in detail.

C. MULTI-DICTIONARY LEARNING

To find the optimal linear combination coefficients, we propose a novel multi-dictionary learning method. Formally, the sparse coefficient α_i can be derived by solving the constrained optimization problem,

$$\begin{aligned} \hat{\alpha} &= \arg \min \|\alpha\|_0 \\ \text{s.t. } \mathbf{y}_i &= \mathbf{D}_i \alpha_i + \beta_i \end{aligned} \quad (6)$$

where $\|\cdot\|_0$ is the L_0 norm, which is the total number of non-zero elements in a vector. Because the approximation error is usually restricted to a controllable interval in empirical data, the constraint in previous equation can be relaxed to the inequality form,

$$\begin{aligned} \hat{\alpha} &= \arg \min \|\alpha\|_0 \\ \text{s.t. } \|\mathbf{D}_i \alpha_i - \mathbf{y}_i\|_2 &\leq \sigma \end{aligned} \quad (7)$$

where σ is the error tolerance. Thus, the above optimization problem can be further interpreted as minimizing the approximation error at a certain sparse

level,

$$\begin{aligned} \hat{\alpha}_i &= \arg \min \|\mathbf{D}_i \alpha_i - \mathbf{y}_i\|_2 \\ \text{s.t. } \|\alpha\|_0 &\leq K_0 \end{aligned} \quad (8)$$

where $\|\cdot\|_2$ is the L_2 norm and K_0 is an upper bound of given sparsity degrees. Though this minimization problem is still regarded as a NP-hard problem by computer scientists, and almost impossible to solve, α_i can be approximately derived by greedy pursuit algorithms, such as orthogonal matching pursuit (OMP) [50], [51] or subspace pursuit (SP) [52]. In this paper, we use the OMP algorithm to solve the above optimization problem. The OMP algorithm is a refinement of the matching pursuit (MP) [53], and its basic idea is described as follows. A sparse approximation is constructed by selecting an atom matching the signal \mathbf{y}_i from a dictionary matrix \mathbf{D}_i . The residual error of signals can be derived, then the atom matching the residual error best is selected continuously. The iteration is carried on until the residual error is negligible or the pre-defined iteration number is achieved. In contrast to the MP algorithm, the OMP algorithm request that all atoms selected at each step of the decomposition should be orthogonal. The convergence speed of the OMP algorithm is faster than the original matching pursuit when the same accuracy is required.

For our band selection method, after each band image is represented by the corresponding dictionaries, we get a coefficient matrix $\mathbf{X} \in R^{B \times B}$. In \mathbf{X} , most of entries would be equal to zero, since the sparsity constraint K_0 is added to the equation (8). The greater the absolute value of coefficients is, the more important the corresponding basis of dictionary is in forming the original hyperspectral image. For each row of the coefficient matrix \mathbf{X} , we count the number of entries which are not zero. Then we get a histogram of the corresponding indices which are the band number. If h denotes the histogram, it is computed by

$$h = \sum_{i=1}^B g(\alpha_i) / B \quad (9)$$

where $g(x) = 1$ if $x \neq 0$, $g(x) = 0$ if $x = 0$. The bands with the first \tilde{B} greater values of histogram series should be selected as the target bands. The following **Algorithm 1** gives the overall computing workflow for our band selection method.

III. EXPERIMENTS

In this section, we perform extensive experiments to verify the effectiveness of the proposed band selection algorithm. First, we introduce the evaluation metrics and several other methods that achieve state-of-art performance in hyperspectral image classification. Then, we present three widely used hyperspectral image datasets for our comparison. Finally, experimental results and detailed discussions are provided to demonstrate the effectiveness and advantages of the proposed band selection approach.

Algorithm 1 Multi-Dictionary Sparse Representation Method

Require:

Hyperspectral image $\mathbf{Y} \in R^{W \times H \times B}$, number of pixels N ($N \ll B$), sparsity level K_0 , target band dimension \tilde{B} .

Ensure:

Reduced-band hyperspectral image $\tilde{\mathbf{Y}} \in R^{W \times H \times \tilde{B}}$;

- 1: Convert \mathbf{Y} into $\tilde{\mathbf{Y}} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_B]$, $\mathbf{y}_i \in R^{WH}$;
 - 2: Randomly select N pixels in $\tilde{\mathbf{Y}}$ to form a new sample set $\hat{\mathbf{Y}} = [\hat{\mathbf{y}}_1, \hat{\mathbf{y}}_2, \dots, \hat{\mathbf{y}}_B]$, $\hat{\mathbf{y}}_i \in R^N$;
 - 3: Construct a group of overcompleted dictionaries, *i.e.*, $\mathbf{D}_i = [\hat{\mathbf{y}}_1, \hat{\mathbf{y}}_2, \dots, \hat{\mathbf{y}}_{i-1}, 0, \hat{\mathbf{y}}_{i+1}, \dots, \hat{\mathbf{y}}_B]$;
 - 4: Apply OMP algorithm to Equ.(8) and find the coefficient matrix $\tilde{\mathbf{X}} \in R^{B \times N}$;
 - 5: Count nonzero entries in every row of $\tilde{\mathbf{X}}$ by Equ.(9);
 - 6: Sort the number of nonzero entries in descending order.
 - 7: Select the bands with the first \tilde{B} indices.
 - 8: Build $\tilde{\mathbf{Y}}$ with selected bands.
-

A. EVALUATION METRICS AND COMPARED METHODS

In this paper, the proposed band selection method is mainly used for hyperspectral image classification. Therefore, the overall classification accuracy (OCA) is the main evaluation metric used in the experimental analysis. In addition, to compare the performance of each method in eliminating the data redundancy, the correlation coefficient of selected bands and the time of computation are also calculated and compared.

In order to fairly evaluate the performance of the proposed method and other state-of-the-art techniques, unsupervised comparison settings without using labeled information have been made [24], [25]. More specifically, we compare several classical band selection methods, *i.e.*, linear prediction (LP) based band selection [24], [25], orthogonal subspace projection (OSP) based band selection [24], cluster based band selection [54], and sparse representation based band selection (SpaBS) [28]. In addition, the comparison also includes the traditional dimensionality reduction methods, such as PCA based feature extraction.

1) LP-BASED BAND SELECTION [24]

This kind of band selection methods are based on the criterion of band dissimilarity. Two bands are first selected randomly or based on the dissimilarity among the original band set. Using the selected bands, the methods estimate the next candidate band by performing linear prediction and least squares solution. The band that yields the maximum error is considered as the most dissimilar band to the selected band set, and can be selected for the next iteration.

2) OSP-BASED BAND SELECTION [24]

This kind of methods first compute the orthogonal subspace of selected bands. Then it compute all other bands' projection in this subspace. The band that yields the maximum

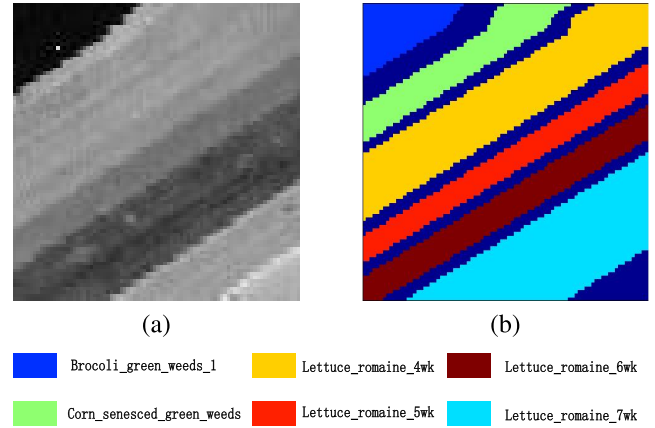


FIGURE 2. Sample image in the Salinas-A Scene dataset. (a) Grayscale image. (b) Corresponding pseudo color ground truth.

orthogonal component is the most dissimilar band to the selected bands, and can be selected as a new target band.

3) CLUSTER-BASED BAND SELECTION [54]

In this kind of methods, every band image is considered a performed object. First, all bands are clustered using a hierarchical clustering algorithm. Then, the methods select the most representative band in every cluster based on divergence measures. These representative bands can be considered as the selected bands.

4) SpaBS [28]

In this method, a sparse representation of the hyperspectral image data is pursued through an existing algorithm, K-SVD, that decomposes the image data into the multiplication of an overcomplete dictionary (or signature matrix) and the coefficient matrix. By calculating the histogram of the coefficient matrix, the top K bands is selected.

B. HYPERSPECTRAL DATASETS

Three real-world hyperspectral image datasets [55] are used for experimental verifications. All of images in these three datasets are earth observation images taken from air-bornes or satellites over some public available hyperspectral scenes.

1) SALINAS-A SCENE DATASET

The Salinas-A Scene dataset is a subset of the Salinas Scene dataset, gathered by AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) sensor in 1998, over Salinas Valley in California. It is characterized by a high spatial resolution (3.7-meter pixels). The covered area comprises 512 lines by 217 samples. It includes vegetables, bare soils, and vineyard fields. In Salinas-A, there are 86×83 pixels with 224 bands including 20 bands of water absorption. Following previous works [56], we discarded the 20 water absorption bands. The groundtruth of this dataset contains 6 classes. Figure 2 shows

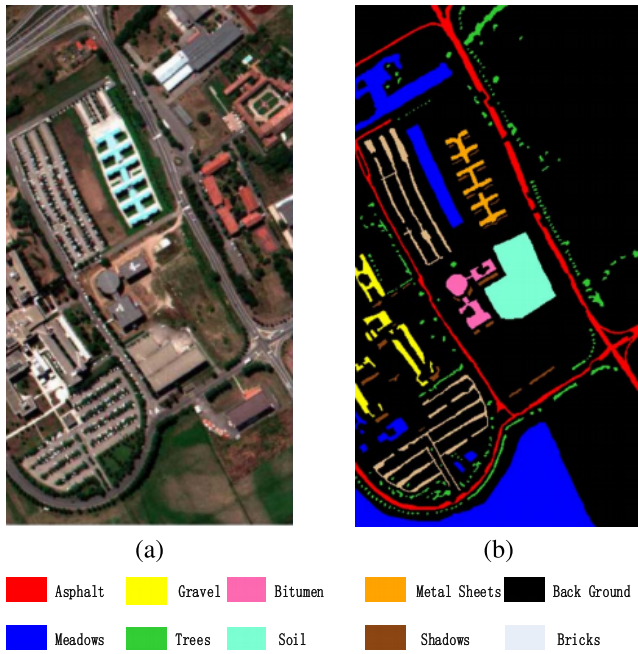


FIGURE 3. Sample image in the Pavia University Scene dataset. (a) True color image. (b) Corresponding pseudo color ground truth.

the Salinas-A grayscale image and its corresponding pseudo color ground truth.

2) PAVIA UNIVERSITY SCENE DATASET

The Pavia University Scene dataset was acquired by the ROSIS (Reflective Optics System Imaging Spectrometer) sensor over the Engineering School, Pavia University, northern Italy. The number of spectral bands is 103, and the number of pixels is 610×340 . The geometric resolution of this scene is 1.3 meters. Pavia University Scene dataset employed in this paper is the one with groundtruth land cover map, which includes 9 classes. Figure 3 shows the image of this dataset and its corresponding pseudo color ground truth.

3) INDIAN PINES SCENE DATASET

This dataset was also gathered by AVIRIS sensor in 1992, over the Indian Pines test site in north-western Indiana. It consists of 145×145 pixels and 224 spectral reflectance bands in the wavelength range of $0.4 - 2.5 \mu\text{m}$. The Indian Pines Scene dataset contains two-thirds agriculture, and one-third forest or other natural perennial vegetation. There are two major dual lane highways, a rail line, as well as some low density housing, other built structures, and smaller roads. This scene is a subset of a larger scene. The ground truth available is designated into 16 classes and is not all mutually exclusive. In the experiments, 12 classes are selected. Following previous works [42], [57], we also reduce the number of bands to 200 by removing bands covering the region of water absorption. Figure 4 shows this scene's grayscale image and its ground truth.

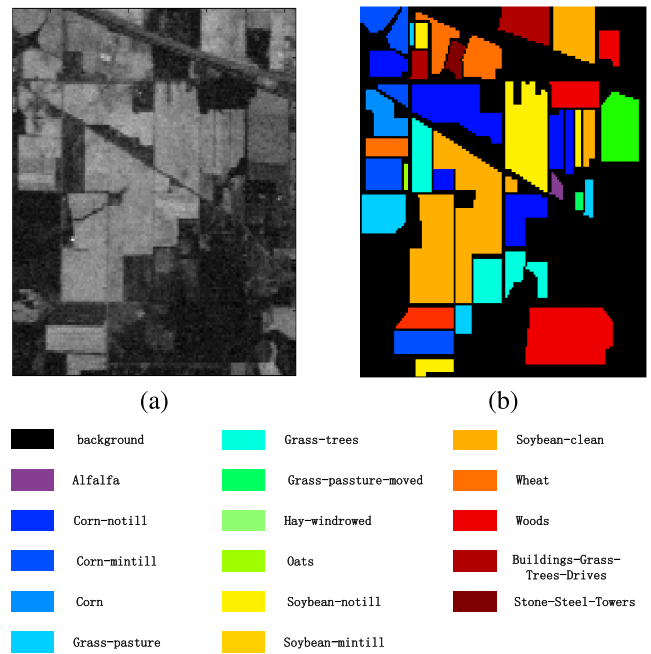


FIGURE 4. Sample image in the Indian Pines Scene dataset. (a) Grayscale image. (b) Corresponding pseudo color ground truth.

C. CLASSIFIER DESCRIPTION

In order to evaluate the classification performance of each band selection method, we resort to the multi-class classifiers with the labeled images and selected bands. Following previous works [15], [58], both K-nearest neighborhood (KNN) and support vector machine (SVM) are used to compare the classification performance of the selected image bands obtained by various band selected methods. We use the KNN and SVM methods proposed in [48] and [59]. The classification process of hyperspectral images are introduced briefly as follows.

1) K-NEAREST NEIGHBORHOOD (KNN) [59]

For a testing sample (a pixel in this paper), the K nearest neighbor classification algorithm is to find K training samples which are closest to the testing sample. The category of the testing sample is determined by a majority-voting scheme, using the categories of the K training samples. When sufficient training samples are available, the KNN algorithm can achieve higher performance than vanilla supervised classifiers.

2) SUPPORT VECTOR MACHINE (SVM) [48], [57]

SVM is a popular supervised learning model that can analyze data, identify patterns, and be used for classification and regression analysis. In our experiment, a set of one-vs-all SVMs based on the polynomial and RBF kernels have been tested. One of the advantages of SVMs is that it is not limited by the number of samples and the dimensionality of samples.

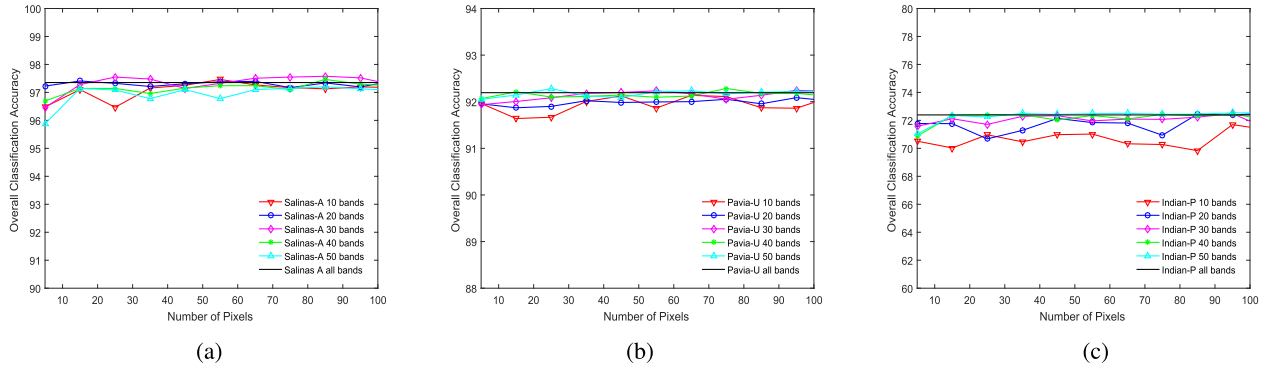


FIGURE 5. Illustration of the effect on the overall classification accuracy (OCA) with different numbers of sampling pixels. The classifier is KNN, and K is set to 6, 9 and 12 respectively. Sparse level K_0 is set to 6. (a) Salinas-A scene dataset. (b) Pavia university scene dataset. (c) Indian pines scene dataset

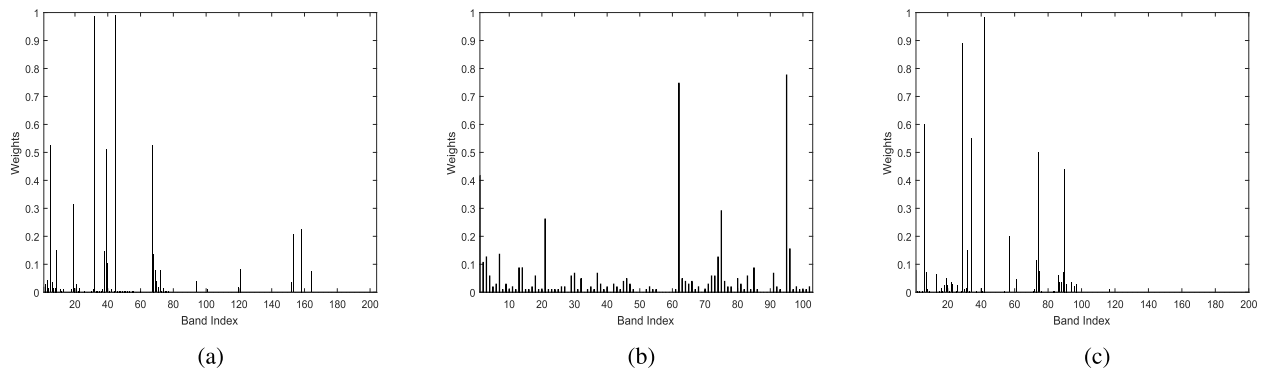


FIGURE 6. The histograms of coefficients on the three datasets. The distribution of the band contribution is nonuniform. Most weights of bands are very low. (a) Salinas-A scene dataset. (b) Pavia university scene dataset. (c) Indian pines scene dataset.

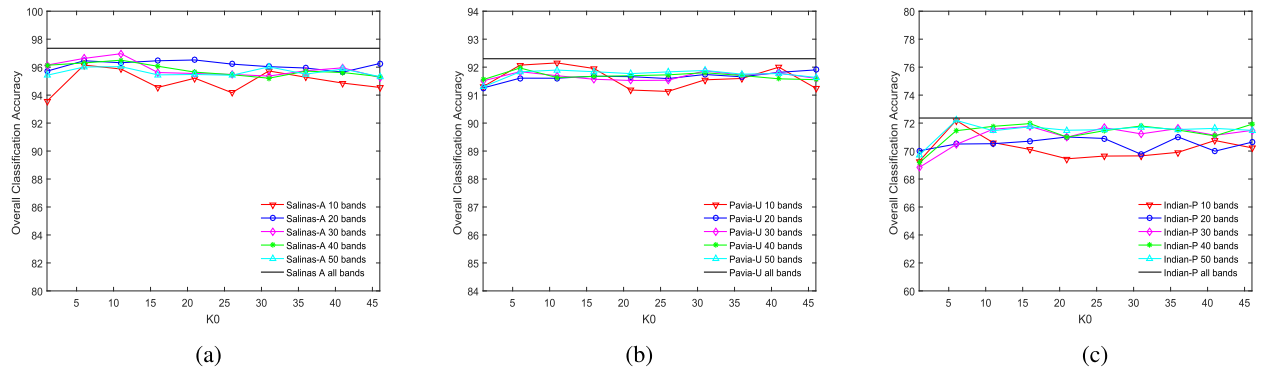


FIGURE 7. The overall classification accuracy with different sparsity levels. We use the KNN classifier on the three datasets. (a) Salinas-A scene dataset. (b) Pavia university scene dataset. (c) Indian pines scene dataset.

D. SAMPLE NUMBER ANALYSIS

In the proposed band selection method, several dictionaries should be created during the optimization process, meanwhile these dictionaries must be overcompleted. Since the number of pixels in hyperspectral images is often greater than the number of bands, if all the pixels are involved in the optimization, the dictionaries are not overcompleted definitely. In order to ensure the completeness of the dictionaries, we select a part of the pixels in the entire image as processing objects,

omitting the rest before the optimization. The 2-6 steps of Algorithm 1 show the optimization process. Therefore, what we concern in this subsection is the effect of the sample number on the band selection results, reflected on the OCA metric.

To analyze the effect, we perform ablation studies on three datasets. Specifically, we randomly pick out 20 samples from each class to form the training set. Using our proposed method, we select 10, 20, 30, 40 and 50 bands from all

TABLE 1. Overall classification accuracy (OCA) with the specific number of pixels on three datasets.

Number of pixels	15	25	35	45
Salinas-A(%)	97.1	96.47	97.16	97.25
Pavia-U(%)	92.15	92.04	92.20	92.08
Indian-P(%)	70.03	70.96	70.46	70.98
Number of pixels	55	65	75	85
Salinas-A(%)	97.46	97.27	97.17	97.13
Pavia-U(%)	92.23	92.56	92.07	93.28
Indian-P(%)	71.02	70.32	70.545	70.27

bands for the experiments. Thus, each dataset has five band selection results. To investigate the effect of different numbers of samples on the classification accuracy, we change the sample number from 5 to 100. Figure 5 illustrates the classification results of the KNN classifiers (For Salinas-A, $K = 6$; For Pavia-U, $K = 9$; For Indian-P, $K = 12$.) using the OCA metric. From Figure 5, it can be seen that the classification accuracy almost remains unchanged with the increase of sample number, and only minor fluctuations (the value is not greater than 2%) appear at several specific locations. On the Salinas-A dataset, the changes in the number of pixels have little effect on the classification results. On the other two datasets, the overall trend of the result curve is still horizontal. In addition, on the Salinas-A and Pavia-U datasets, the classification accuracy of the band selection is very close to the classification accuracy of the whole bands during changing the number of sampling points. The variance of classification accuracy on the Indian-P dataset is larger than the other two datasets. The main reason may be that there are more classes, and samples are more irregularly distributed in the Indian-Pines dataset.

Table 1 provides more accurate classification results with specific sampling settings when 10 bands are selected. From the results in Figure 5 and Table 1, we can draw the conclusion that the number of sampling points has little effect on the results of band selection. Thus, a small number of sampling points can be selected for reducing the computation. In the following experiments, we empirically set the number of pixels to be 50 ($N = 50$) for classification evaluation, without further parameter tuning.

E. SELECTED BAND ANALYSIS

In Algorithm 1, the coefficient matrix \tilde{X} is the key support of band selection. To figure out which kind of bands our method has selected, we draw bar diagrams which illustrate the ratio of the count of non-zero entries to the count of all entries in one column of coefficient matrix \tilde{X} . The calculation method is given in Algorithm 1 and Equation (9). This ratio represents the contribution of one band in the hyperspectral image to the whole image, or the proportion of bands associated with this band in the image. More specifically, in the coefficient matrix \tilde{X} , the higher ratio of non-zero entries in the i -th column means the larger contribution of the i -th band to the whole image. Figure 6 illustrates the bar diagrams in which 50 sampling pixels are selected randomly and the sparsity

TABLE 2. The TOP 5 Selected bands and their weights.

Salinas-A	Band Index	32	45	3	153	39
	Weight	0.985	0.980	0.525	0.485	0.446
Pavia-U	Band Index	91	59	3	1	87
	Weight	0.670	0.456	0.408	0.359	0.330
Indian-P	Band Index	42	29	35	6	1
	Weight	0.995	0.905	0.695	0.670	0.590

level K_0 is 6 in solving the sparse optimization problem. In Figure 6, we can see that the contribution of each band to the original hyperspectral image is very different. In the Salinas-A dataset, the weights of two specific bands are close to one, and weights of most bands are below 0.5. In the Pavia-U dataset, the weight of only one band is close to 0.7, and weights of almost all bands are below 0.45. In the Indian-P dataset, most weights of the bands are zeros. Table 2 lists the top 5 weights of all bands in the three datasets. The weight distribution is not uniform over the bands. These experimental results further indicate that there are many redundancies in the hyperspectral bands and only a few bands are useful. Therefore, it is reasonable to select appropriate bands for specific applications.

F. SPARSE LEVEL ANALYSIS

In solving the coefficient matrix \tilde{X} , the sparsity level K_0 should be set appropriately. To investigate its impact, we evaluate a set of choices of K_0 on the three datasets, where we use the KNN classifier and K is set to be 6, 9, 12, respectively. Figure 7 shows the classification accuracy curves when the sparsity level K_0 changes from 1 to 50 on the three datasets. There are several fundamental observations in Figure 7: 1) for the Salinas-A dataset, when K_0 is less than 3, the classification accuracy is relatively low. When K_0 is greater than 3, the classification performance first increases, then drops down slightly. The classification performance of the model with 10 bands fluctuates more drastically than other models. This indicates that the model with very few bands is sensitive to the sparse level K_0 . The same trend can be observed in other two datasets. 2) for the Pavia-U dataset, the classification performance with different sparse levels and selected bands remains almost unchanged (no more than 2% fluctuation), even they vary locally. This result shows that our method is very robust in this dataset under different sparse levels. 3) for the Indian-P dataset, the models with more bands are better in the classification performance. The models are more fluctuated on this dataset under different K_0 . Based on above facts, we set $K_0 = 6$ in the following experiments.

G. COMPARISON WITH OTHER METHODS

In order to fairly evaluate the performance of different band selection algorithms, KNN and SVM classifiers are adopted to classify the hyperspectral image after the band selection. The value of K in KNN is set to 6, 9 and 12 for the Salinas-A, Pavia-U and Indian-P datasets, respectively. The parameters of the SVM classifier is obtained by the K -fold cross

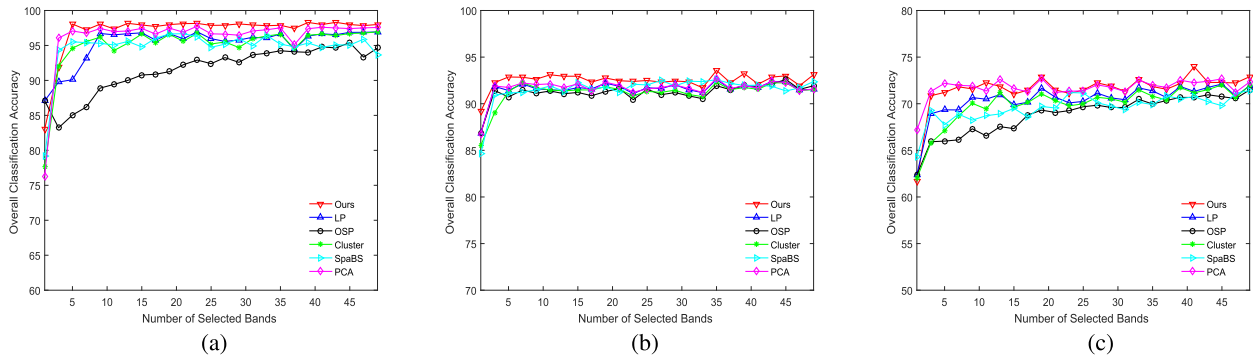


FIGURE 8. The classification results of different band selection methods using the KNN classifier on the three datasets. (a) Salinas-A scene dataset. (b) Pavia university scene dataset. (c) Indian pines scene dataset.

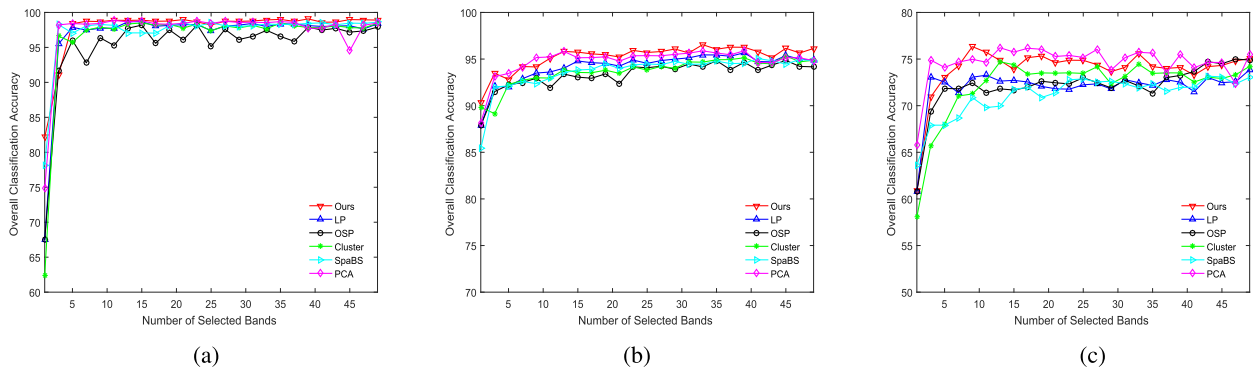


FIGURE 9. The classification results of different band selection methods using the SVM classifier on the three datasets. (a) Salinas-A scene dataset. (b) Pavia university scene dataset. (c) Indian pines scene dataset.

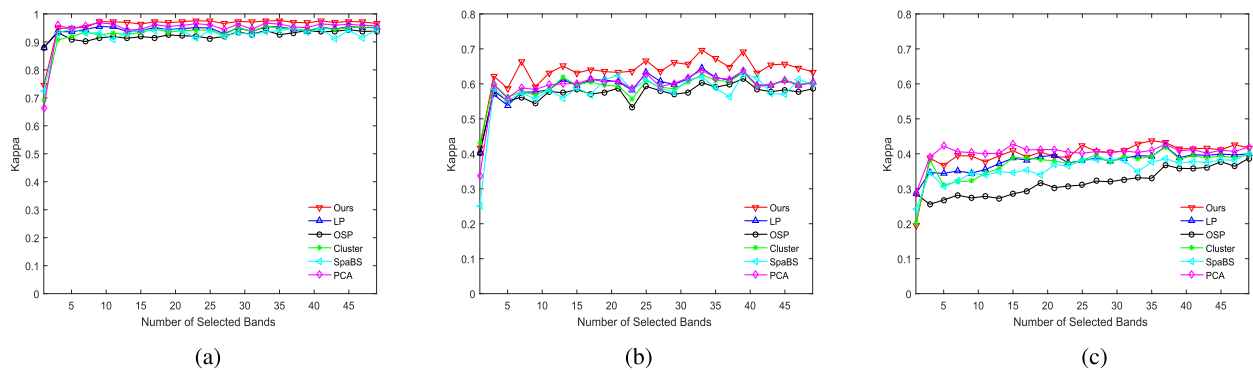


FIGURE 10. The Kappa coefficients of different band selection methods using the KNN classifier on the three datasets. (a) Salinas-A scene dataset. (b) Pavia university scene dataset. (c) Indian pines scene dataset.

validation. The compared methods have been described in section III. A.

1) OVERALL CLASSIFICATION PERFORMANCE

For the classification performance evaluation, 20 samples are selected randomly in each classes. For our method, we set $K_0 = 6$. In order to reduce the randomness, we perform all the compared methods with 10 trials and use the averaged results. We report the results with different selected band number n , ranging from 1 to 50. The Figure 8 and Figure 9 show the classification performance of the proposed method and compared algorithms on the three datasets. From the comparison results, we can see that our proposed method consistently

outperforms other methods by a large margin, especially on the Salinas-A and Pavia-U datasets. More specifically, the OSP method achieves the lowest recognition rate. The LP, SpaBS and cluster based methods have similar classification results, which are inferior to our method. Even though the PCA method is superior to other methods, it is not better than our proposed method.

2) KAPPA COEFFICIENT

To further verify the advantages of our proposed method, we also calculate the Kappa coefficient, which measures the agreement between two raters who each classify N items into C mutually exclusive categories. The results with the KNN

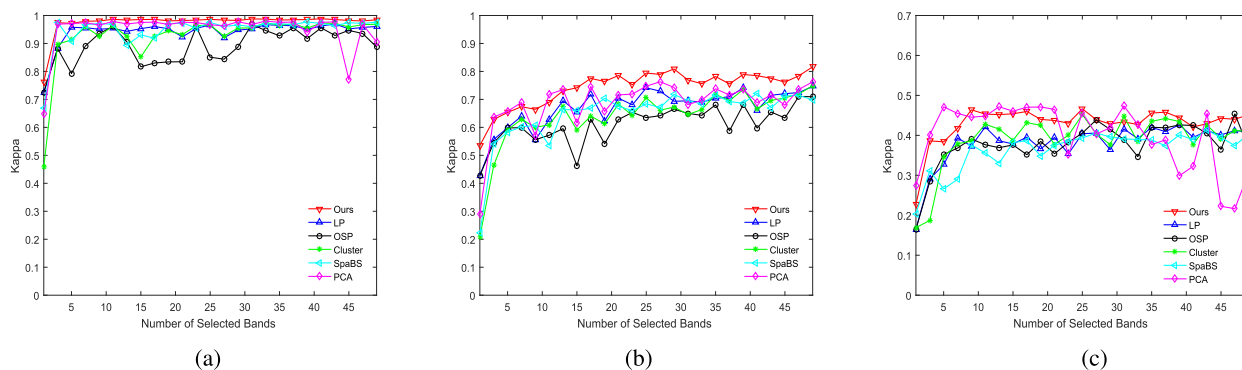


FIGURE 11. The Kappa coefficients of different band selection methods using the SVM classifier on the three datasets. (a) Salinas-A scene dataset. (b) Pavia university scene dataset. (c) Indian pines scene dataset.

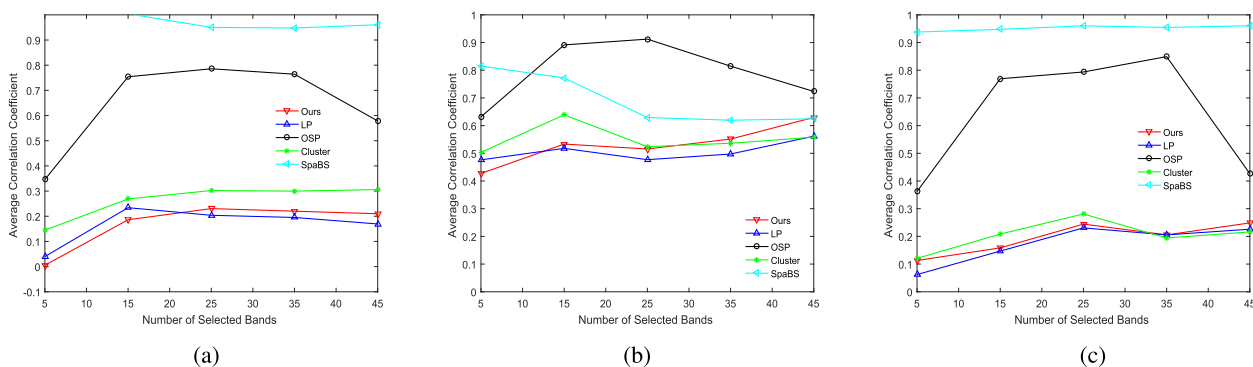


FIGURE 12. Comparison of average correlation coefficients with different band selection methods on three datasets. (a) Salinas-A scene dataset. (b) Pavia university scene dataset. (c) Indian pines scene dataset.

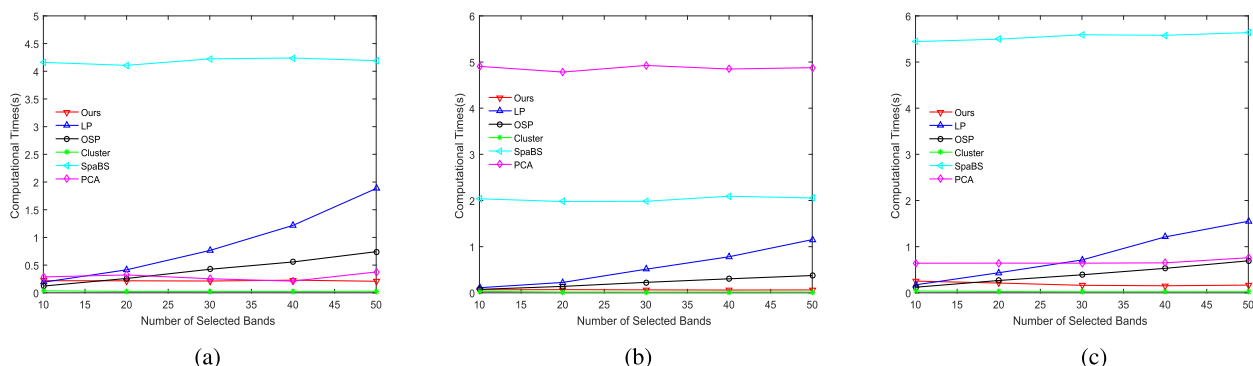


FIGURE 13. Comparison of computational time with different band selection methods on three datasets. (a) Salinas-A scene dataset. (b) Pavia university scene dataset. (c) Indian pines scene dataset.

and SVM classifiers are shown in Figure 10 and Figure 11, respectively. From the results, we can see that our method also achieve better performance with the Kappa coefficient measure. In summary, the proposed method is better than most of compared methods in terms of both classification accuracy and Kappa coefficient. Our method is inferior to the PCA method in several cases on the Indian-P dataset, however, the performance is still very comparative.

3) BAND CORRELATION ANALYSIS

To measure the separation of selected bands, we calculate the correlation coefficient of bands selected by different methods. Generally, the lower the correlation among spectral bands is,

the better the subset of selected bands is. Figure 12 illustrates the average correlation coefficient with the compared methods other than PCA. On the three data subsets, our method ranks first or second in this metric, which is in agreement with previous results shown in Figures 8-11. While, SpaBS's performance is a little bad. This further confirms that our method is able to select the most informative and separative bands for the real-world applications.

4) COMPUTATIONAL TIME

In order to assess the efficiency of the proposed method, the computational time of each method is also provided. For fair comparison, all compared methods are implemented in

MATLAB2015, and are tested on an Intel Core i5-2400 CPU with 12-GB random access memory. Figure 13 illustrates the runtime of different band selection methods with varied numbers of bands. From the results, we can see that the cluster based method takes the least time. Although our method is not the fastest, it is faster than most of the compared methods. In addition, our method is not significantly affected by the number of bands. While the runtime of LP and OSP methods increases quickly as the number of bands ascends. Because of the large number of iterations in the SpaBS method, the time efficiency is also poor. Therefore, our method has a great advantage in time efficiency.

IV. CONCLUSION

In this work, we propose a novel band selection method based on the multi-dictionary sparse representation. The proposed method is fully unsupervised and consists of three main components: 1) creating multi-dictionaries; 2) optimizing sparse coefficients; 3) computing and sorting weights of desire bands. Our proposed method not only reduces the dimension of spectral bands, but also preserves the original information of bands. Extensive experimental results show that the proposed method is very effective and outperforms other competitors. Of course, the method can be improved. For example, because of the large number of dictionaries, the amount of calculation increases. In the future, how to design an effective framework to reduce the number of dictionaries and the amount of calculation is the next work.

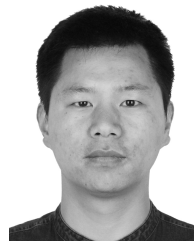
REFERENCES

- [1] C. I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. New York, NY, USA: Plenum Publishing Co., 2003.
- [2] X. Jia and J. A. Richards, "Efficient maximum likelihood classification for imaging spectrometer data sets," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 2, pp. 274–281, Mar. 1994.
- [3] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 1, pp. 55–63, Jan. 1968.
- [4] R. Ernest, *Adaptive Control Processes: A Guided Tour*. Princeton, NJ, USA: Princeton Univ. Press, 1961.
- [5] H. Lu, L. Zhang, and S. Serikawa, "Maximum local energy: An effective approach for multisensor image fusion in beyond wavelet transform domain," *Comput. Math. Appl.*, vol. 64, no. 5, pp. 996–1003, 2012.
- [6] W. Zhao, H. Lu, and D. Wang, "Multisensor image fusion and enhancement in spectral total variation domain," *IEEE Trans. Multimedia*, vol. 20, no. 4, pp. 866–879, Apr. 2018.
- [7] L. He, X. Xu, H. Lu, Y. Yang, F. Shen, and H. T. Shen, "Unsupervised cross-modal retrieval through adversarial learning," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2017, pp. 1153–1158.
- [8] C. Bo, H. Lu, and D. Wang, "Hyperspectral image classification via JCR and SVM models with decision fusion," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 2, pp. 177–181, Feb. 2016.
- [9] P. Li, D. Wang, L. Wang, and H. Lu, "Deep visual tracking: Review and experimental comparison," *Pattern Recognit.*, vol. 76, pp. 323–338, Apr. 2018.
- [10] X. Jia, B.-C. Kuo, and M. M. Crawford, "Feature mining for hyperspectral image classification," *Proc. IEEE*, vol. 101, no. 3, pp. 676–697, Mar. 2013.
- [11] V. Zubko, Y. J. Kaufman, R. I. Burg, and J. V. Martins, "Principal component analysis of remote sensing of aerosols over oceans," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 730–745, Mar. 2007.
- [12] T. V. Bandos, L. Bruzzone, and G. Camps-Valls, "Classification of hyperspectral images with regularized linear discriminant analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 862–873, Mar. 2009.
- [13] B.-C. Kuo and D. A. Landgrebe, "Nonparametric weighted feature extraction for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 5, pp. 1096–1105, May 2004.
- [14] P. Bajcsy and P. Groves, "Methodology for hyperspectral band selection," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 7, pp. 793–802, 2004.
- [15] C.-I. Chang, Q. Du, T.-L. Sun, and M. L. G. Althouse, "A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 6, pp. 2631–2641, Nov. 1999.
- [16] L. Bruzzone, F. Roli, and S. B. Serpico, "An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 6, pp. 1318–1321, Nov. 1995.
- [17] R. Huang and M. He, "Band selection based on feature weighting for classification of hyperspectral data," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 2, pp. 156–159, Apr. 2005.
- [18] S. De Backer, P. Kempeneers, W. Debruyne, and P. Scheunders, "A band selection technique for spectral classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 3, pp. 319–323, Jul. 2005.
- [19] M. Riedmann and E. J. Milton, "Supervised band selection for optimal use of data from airborne hyperspectral sensors," in *Proc. IGARSS IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2003, pp. 1770–1772.
- [20] W. Wei, Q. Du, and N. H. Younan, "Fast supervised hyperspectral band selection using graphics processing unit," *J. Appl. Remote Sens.*, vol. 6, no. 1, p. 061504, 2012.
- [21] H. Yang, Q. Du, H. Su, and Y. Sheng, "An efficient method for supervised hyperspectral band selection," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 1, pp. 138–142, Jan. 2011.
- [22] X. Cao, T. Xiong, and L. Jiao, "Supervised band selection using local spatial information for hyperspectral image," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 329–333, Mar. 2016.
- [23] L. Wang, X. Jia, and Y. Zhang, "A novel geometry-based feature-selection technique for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 4, no. 1, pp. 171–175, Jan. 2007.
- [24] Q. Du and H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 564–568, Oct. 2008.
- [25] H. Yang, Q. Du, and G. Chen, "Unsupervised hyperspectral band selection using graphics processing units," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 3, pp. 660–668, Sep. 2011.
- [26] C. Sui, Y. Tian, and Y. Xu, "An unsupervised band selection method based on overall accuracy prediction," in *Proc. 22nd Int. Conf. Pattern Recognit.*, 2014, pp. 3756–3761.
- [27] C. Sui, Y. Tian, Y. Xu, and Y. Xie, "Unsupervised band selection by integrating the overall accuracy and redundancy," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 185–189, Jan. 2015.
- [28] S. Li and H. Qi, "Sparse representation based band selection for hyperspectral images," in *Proc. 18th IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 2693–2696.
- [29] Q. Du, J. M. Bioucas-Dias, and A. Plaza, "Hyperspectral band selection using a collaborative sparse model," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2012, pp. 3054–3057.
- [30] H.-C. Chien, C.-H. Lai, and K.-H. Liu, "Unsupervised band selection based on group-based sparse representation," in *Proc. Asian Conf. Comput. Vis.*, 2016, pp. 389–401.
- [31] K. Sun, X. Geng, and L. Ji, "A new sparsity-based band selection method for target detection of hyperspectral image," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 329–333, Feb. 2015.
- [32] W. Sun, M. Jiang, W. Li, and Y. Liu, "A symmetric sparse representation based band selection method for hyperspectral imagery classification," *Remote Sens.*, vol. 8, no. 3, p. 238, 2016.
- [33] H. Lu, Y. Li, M. Chen, H. Kim, and S. Serikawa, "Brain intelligence: Go beyond artificial intelligence," *Mobile Netw. Appl.*, vol. 23, no. 2, pp. 368–375, 2018.
- [34] X. Xu, L. He, H. Lu, L. Gao, and Y. Ji, "Deep adversarial metric learning for cross-modal retrieval," in *World Wide Web*, Apr. 2018, pp. 1–16.
- [35] J.-L. Starck, F. Murtagh, and J. M. Fadili, *Parse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
- [36] M. Elad and M. Aharon, "Image denoising via learned dictionaries and sparse representation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2006, pp. 895–900.
- [37] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.

- [38] J. Zhang, D. Zhao, and W. Gao, "Group-based sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3336–3351, Aug. 2014.
- [39] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma, "Toward a practical face recognition system: Robust alignment and illumination by sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp. 372–386, Feb. 2012.
- [40] C.-Y. Lu, H. Min, J. Gui, L. Zhu, and Y.-K. Lei, "Face recognition via weighted sparse representation," *J. Vis. Commun. Image Represent.*, vol. 24, no. 2, pp. 111–116, Feb. 2013.
- [41] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, and S. Yan, "Sparse representation for computer vision and pattern recognition," *Proc. IEEE*, vol. 98, no. 6, pp. 1031–1044, Jun. 2010.
- [42] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3973–3985, Oct. 2011.
- [43] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification via kernel sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 217–231, Jan. 2013.
- [44] L. He, W. Ruan, and Y. Li, "Sparse representation based hyperspectral imagery classification via expanded dictionary," in *Proc. 4th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens.*, Jun. 2012, pp. 1–4.
- [45] C. Bo, H. Lu, and D. Wang, "Spectral-spatial k-nearest neighbor approach for hyperspectral image classification," *Multimedia Tools Appl.*, vol. 77, no. 9, pp. 10419–10436, 2018.
- [46] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Sparse representation for target detection in hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 629–640, Jun. 2011.
- [47] M. Pal, "Random forest classifier for remote sensing classification," *Int. J. Remote Sens.*, vol. 26, no. 1, pp. 217–222, 2007.
- [48] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, p. 27, 2011.
- [49] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Mach. Learn.*, vol. 29, no. 2, pp. 131–163, Nov. 1997.
- [50] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Asilomar Conf. Signals, Syst. Comput.*, 1993, pp. 40–44.
- [51] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [52] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.
- [53] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
- [54] A. Martinez-Uso, F. Pla, J. M. Sotoca, and P. Garcia-Sevilla, "Clustering-based multispectral band selection using mutual information," in *Proc. 18th Int. Conf. Pattern Recognit.*, 2006, pp. 760–763.
- [55] *Hyperspectral Remote Sensing Scenes*. Accessed: Mar. 8, 2017. [Online]. Available: <http://www.ehu.eus/ccwintco/index.php/Hyperspectral-Remote-Sensing-Scenes>
- [56] X. Ma, J. Geng, and H. Wang, "Hyperspectral image classification via contextual deep learning," *EURASIP J. Image Video Process.*, vol. 2015, no. 1, p. 20, 2015.
- [57] J. A. Gualtieri and R. F. Crompt, "Support vector machines for hyperspectral remote sensing classification," in *Proc. Aipr Workshop, Adv. Comput.-Assist. Recognit.*, 1998, pp. 1–28.
- [58] Y. Yuan, G. Zhu, and Q. Wang, "Hyperspectral band selection by multi-task sparsity pursuit," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 631–644, Feb. 2015.
- [59] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.



FEI LI received the B.S. degree from the University of Petroleum China in 1999 and the M.S. degree from Nankai University, Tianjin, China, in 2005. She is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China. Her research interests are in deep learning hyperspectral image processing.



PINGPING ZHANG received the B.E. degree in mathematics and applied mathematics from Henan Normal University, Xinxiang, China, in 2012. He is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering, Dalian University of Technology, Dalian, China. His research interests are in deep learning, saliency detection, object tracking, and semantic segmentation.



HUCHUAN LU (SM'12) received the M.Sc. degree in signal and information processing and the Ph.D. degree in system engineering from the Dalian University of Technology (DUT), China, in 1998 and 2008, respectively. He has been a Faculty Member since 1998 and a Professor since 2012 with the School of Information and Communication Engineering, DUT. His research interests are in the areas of computer vision and pattern recognition. In recent years, he focuses on visual tracking, saliency detection, and semantic segmentation. He serves as an Associate Editor for the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: Part B.

• • •