A New Band Selection Method for Hyperspectral Image Based on Data Quality

Kang Sun, Xiurui Geng, Luyan Ji, and Yun Lu

***Abstract*—Most unsupervised band selection methods take the information of bands into account, but few of them pay attention to the quality of bands. In this paper, by combining idea of noise-adjusted principal components (NAPCs) with a state-of-art band selection method [maximum determinant of covariance matrix (MDCM)], we define a new index *Q* to quantitatively measure the quality of the hyperspectral data cube. Both signal-to-noise ratios (SNRs) and correlation of bands are simultaneously considered in *Q*. Based on the new index defined in this article, we propose an unsupervised band selection method called minimum noise band selection (MNBS). Taking the quality (*Q*) of the data cube as selection criterion, MNBS tries to find the bands with both high SNRs and low correlation (high *Q*). The subset selection method, sequential backward selection (SBS), is used in MNBS to improve the search efficiency. Some comparative experiments based on simulated as well as real hyperspectral data are conducted to evaluate the performance of MNBS in this study. The experimental results show that the bands selected by MNBS are always more effective than those selected by other methods in terms of classification.**

***Index Terms*—Band selection, dimensionality reduction, hyperspectral data, noise-adjusted principal component (NAPC), noise fraction.**

I. INTRODUCTION

HYPERSPECTRAL remote sensing has enabled many elaborate image or spectral analysis by providing a large number of spectral bands in the past few years. However, these bands usually have high correlation due to the continuous and narrow spectral samplings. Therefore, extracting effective information or removing redundancy in highly correlated hyperspectral data has received more and more attention.

Most researchers try to generate "new bands" by combining the original bands with different weights together, which is known as "feature extraction." Those generated new bands are expected to contain as large information as possible and have little correlation at the same time. A lot of methods have been reported with regard to feature extraction, such as [1]–[5]. The main disadvantage of these methods is that the bands obtained by them are lacking in physical meaning.

Some other researchers resort to band selection instead of band transformation to perform dimensionality reduction. Different from feature extraction, band selection aims to select a set of qualified or representative bands from larger/full bands set, while no new bands are generated. Bands selected by these methods are of physical meaning as they are exactly the subset of original bands. In addition, band selection will benefit data acquisition, data transportation, and data storage a lot. Many methods in terms of band selection have been reported in recent years. These methods can be roughly classified into two types: supervised and unsupervised. The former generally requires some prior knowledge or training samples, such as [6]–[13]. However, because the prior information is often not available in practice, the supervised techniques are not suitable for hyperspectral band selection [14]. Therefore, this paper focuses on the latter, i.e., unsupervised band selection methods.

Many unsupervised band selection methods involve bands prioritization, most of which are based on information evaluation. For instance, [15] uses information entropy as band selection criterion, [16] and [17] are based on spectral derivative, and Chang [18] proposes to select the bands that have the farthest deviation from Gaussian distribution, which is measured by information divergence (ID). However, these prioritization-based methods have not taken bands' correlation into consideration. As a result, the bands obtained by these methods are highly correlated in general. Some others have considered bands' correlation, e.g., a new similarity measure, maximal information compression index (MICI), which possesses properties such as symmetry, sensitivity to scaling, and invariance to rotation, is proposed in [19] for feature selection. In [20], in order to find the bands with large variance and low correlation, the bands whose covariance matrix has the maximum determinant are selected. The readers are referred to [21]–[26] to read more about unsupervised band selection.

Most of these unsupervised take bands' information into account. Nevertheless, few of them pay attention to the quality of bands. Generally, bands with high level of noise can result in large variances. Therefore, those "noisy" bands will have higher "selection weights" if the quality of the data is neglected, which probably leads to an unreasonable result.

In this paper, we attempt to select the bands with good quality rather than large information. In order to obtain a quantitative selection criterion, we extend the definition of image quality

Manuscript received November 25, 2013; revised February 18, 2014; accepted April 17, 2014. Date of publication May 07, 2014; date of current version August 01, 2014. This work was supported by the National High Technology Research and Development Program (863 Program) of China under Grant 2013AA122804.

K. Sun and X. Geng are with the Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China (e-mail: sunkang-1234@ 163.com; gengxr@sina.com).

L. Ji was with the Key Laboratory of Technology in Geo-spatial Information Processing and Application System, Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China. She is now with the Centre for Earth System Science, Tsinghua University, Beijing 100084, China (e-mail: jiluyan@mail.tsinghua.edu.cn).

Y. Lu is with the School of Materials Science and Engineering, Anhui University of Technology, Anhui 243032, China (e-mail: luyunrosa@sina.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2014.2320299

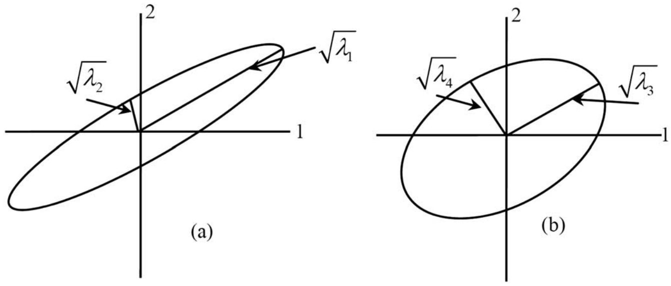


Fig. 1. Different measure criteria for information. (a) The sum of square root of variances (information) $\sqrt{\lambda_1}+\sqrt{\lambda_2} > \sqrt{\lambda_3}+\sqrt{\lambda_4}$, with large total information but high correlation. (b) The product of square root of variances, $\sqrt{\lambda_1}\sqrt{\lambda_2} < \sqrt{\lambda_3}\sqrt{\lambda_4}$, with small total information but low correlation.

which is generally measured by signal-to-noise ratio (SNR) to hyperspectral data cube quality. More specifically, by combining the idea proposed in noise-adjusted principal components (NAPCs) [27] with that proposed in maximum determinant of covariance matrix (MDCM) [20], we define a new index Q to measure the quality of the hyperspectral data cube. High Q implies good cube quality which means high SNR and low correlation of bands.

Based on the newly defined index Q, we propose a band selection method in this paper, called minimum noise band selection (MNBS). MNBS aims at selecting bands with good quality (high Q) or with high SNR and low correlation. In order to save computing time, the subset selection strategy, sequential backward selection (SBS) [28], is adopted in MNBS. SBS is a "top-down" method that begins with full bands set followed by removing bands successively. The results of experiments show that MNBS can find the bands with high quality effectively and is very robust to noise. As a result, the classifications based on the bands selected by MNBS are of high accuracy.

The remainder of this paper is organized as follows. Section II gives a brief instruction of MDCM. Section III elaborates the presented method, namely, MNBS, which includes the definition of hyperspectral data quality (Q), the explanation, and fast implementation of MNBS. The evaluation experiments that are conducted on hyperspectral data sets are described in Section IV. Eventually, Section V shows some concluding remarks.

II. MDCM Band Selection

MDCM [20] is an unsupervised BS method which attempts to find the bands that have large information and low correlation at the same time. Fig. 1 shows the example two-dimensionally (2-D). Fig. 1(a) has larger variance (information) than Fig. 1(b), i.e., $\sqrt{\lambda_1}+\sqrt{\lambda_2} > \sqrt{\lambda_3}+\sqrt{\lambda_4}$. However, the bands in Fig. 1(a) are more highly correlated than those in Fig. 1(b), where $\sqrt{\lambda_1}\sqrt{\lambda_2} < \sqrt{\lambda_3}\sqrt{\lambda_4}$, which means the ellipse in Fig. 1(b) has a larger area (volume in high dimensional space) than that in Fig. 1(a). It can be easily proved that the volume of the ellipsoid equals the determinant of the covariance matrix except for a constant factor. In order to discourage the selection of highly correlated bands, the determinant of bands' covariance matrix is adopted as the BS criterion in MDCM.

In fact, MDCM is equivalent to finding the bands of meeting the following mathematical expression

$$[\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n] = \arg\max_{\mathbf{X}_n} \det\left(\frac{1}{M}\mathbf{X}_n\mathbf{X}_n^T\right) \quad (1)$$

where $\mathbf{X}_n = [\mathbf{x}_{a_1},\mathbf{x}_{a_2},\ldots,\mathbf{x}_{a_n}]$ represents arbitrary n bands in hyperspectral data $\mathbf{X}$, $[\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n]$ is the finally selected bands set, and M is the number of pixels.

III. Methodology

In this section, we give detailed discussion about MNBS, including the quality of the hyperspectral data cube and the implementation of MNBS. Assume we are selecting n bands from centralized data with L bands $\mathbf{X} = [\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_L]^T$, where $\mathbf{x}_i = [x_{1i}, x_{2i},\ldots,x_{Mi}]^T$ is the column vector constructed by the ith band of $\mathbf{X}$ and M is the number of pixels that consists of ns samples and nl lines ($M = ns * nl$).

A. Quality of Hyperspectral Data Cube

In this section, we expect to define an index Q to quantitatively measure the quality of the hyperspectral data cube. As is well known, SNR is a commonly used index to measure image quality which is often defined as $\mathrm{SNR} = 10\log_{10}(E(\mathbf{x}^T\mathbf{x})/E(\mathbf{n}^T\mathbf{n}))$, where $\mathbf{x}$ denotes signals and $\mathbf{n}$ denotes noises. However, the conventional SNR is often used to measure the quality for a single band rather than a data cube. For a data cube, the correlation of bands should be also considered. Obviously, high SNRs and low correlation of bands means a high cube quality. Therefore, the newly defined index Q is expected to not only eliminate the impact of noise, but also take correlation into account.

In order to eliminate the impact of noise, we perform noise whitening to the original data, which is derived from the idea proposed in NAPCs. The noise term needs to be separated from the data in advance to perform noise whitening. In this study, the noise estimation method proposed in [3] and [29] is employed, which estimates the noise statistics from the data via a shift difference. This method exploits the fact that in most remotely sensed data, any spectral vector in the image is strongly correlated with the neighboring pixels, while the noise signal shows only weak spatial correlations. The method can be understood to create a noise term $\mathbf{N}$ whose dimensions are ns-1, nl-1, and L, allowing for the extra column and row associated with the shifting. Specifically, the noise of the kth band is the average of the row shift difference and the column shift difference of the band is as follows:

$$\mathbf{N}(i,j,k) = \frac{[\mathbf{x}_k(i,j)-\mathbf{x}_k(i-1,j)] + [\mathbf{x}_k(i,j)-\mathbf{x}_k(i,j-1)]}{2}. \quad (2)$$

The following step is noise whitening that aims at normalizing the level and removing the correlation of the noise in each band. From statistical perspective, the covariance matrix of noise $\mathbf{N}$ is an identity matrix after noise whitening. Therefore, it can be easily obtained that the noise-whitened data is

$$\mathbf{Y} = \mathbf{F}_{\mathbf{N}}^T \mathbf{X} \tag{3}$$

where $\mathbf{F}_{\mathbf{N}} = \mathbf{U} * \boldsymbol{\Delta}^{-\frac{1}{2}}$ is the whitening matrix for $\mathbf{N}$, $\mathbf{U}$, $\boldsymbol{\Delta}$ are eigenvector matrix and eigenvalue matrix of the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{N}}$ for noise [30], respectively. The noise in the transformed data set has unit variance and no correlation with each other. Therefore, as for the noise-whitened bands, their variances are equivalent to SNRs and their correlation can be regarded as the correlation of "signals."

It has been proved in [20] that the bands whose covariance matrix has the maximum determinant are of the largest information and the lowest correlation, as shown in (1). Inspired by MDCM, we defined the quality Q of the hyperspectral data cube $\mathbf{X}$ as the determinant of the covariance matrix for the noise-whitened data $\mathbf{Y}$ as follows:

$$Q(\mathbf{X}) = \det\left(\frac{1}{M}\mathbf{Y}\mathbf{Y}^T\right) = \det\left(\frac{1}{M}\mathbf{F}_{\mathbf{N}}^T \mathbf{X}\mathbf{X}^T \mathbf{F}_{\mathbf{N}}\right) \tag{4}$$

where $\det(\bullet)$ is the determinant operator. It can be seen that the index Q associates the quality (SNRs) with penalty of high correlation of the data cube. Therefore, bands with high Q are of high SNRs and low correlation.

B. Minimum Noise Band Selection

MNBS attempts to select n bands with largest Q from $\mathbf{X}$, i.e.

$$[\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n] = \underset{\mathbf{x}_{a_1}, \mathbf{x}_{a_2}, \ldots, \mathbf{x}_{a_n}}{\arg\max}\ Q(\mathbf{x}_{a_1}, \mathbf{x}_{a_2}, \ldots, \mathbf{x}_{a_n}) \tag{5}$$

where $[\mathbf{x}_{a_1}, \mathbf{x}_{a_2}, \ldots, \mathbf{x}_{a_n}]$ is the arbitrary n bands from $\mathbf{X}$ and $[\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n]$ is the final desired bands.

There are a total of $\binom{L}{n}$ candidates for L bands, which means the exhaustive search is infeasible for hyperspectral data. Therefore, MNBS resorts to subset selection method which yields suboptimal results but is of high efficiency. The subset selection method SBS [28] is adopted in this paper. SBS is a "top-down" search method that regards the full bands set as the initial candidate set, followed by removing bands successively from the candidate set. More specifically, all the bands are considered as the initial candidate set at first, then each band in that set is removed tentatively followed by calculating data quality Q for the corresponding left bands. Bands with the highest Q are kept as the new candidate set in the next iteration. Thus, one band has been removed. The redundant bands are removed one by one similarly until the number of left bands is desired. Since Q is used as the criterion for bands removal in each iteration, the bands selected by MNBS are of good quality rather than large information.

It should be pointed out that although the data quality is defined as the determinant of the covariance matrix for $\mathbf{Y}$, MNBS is not equivalent to performing MDCM to the noise-whitened data. In fact, performing unsupervised band selection method to noise-whitened data is unreasonable since any noise-whitened band is a linear combination of the original bands. More specifically, the ith band of noise-whitened data $\mathbf{Y}$ is the summation of the original bands with different weights; therefore, it neither correspond to the ith band of $\mathbf{X}$ nor to any band of $\mathbf{X}$. The correct way is all the filtration or selection of the bands should be performed to original bands set $\mathbf{X}$ while the criterion function are compared on the corresponding noise-whitened bands set $\mathbf{Y}$.

C. Further Justification and Fast Implementation for MNBS

The computational complexity for MNBS is very high due to the frequent noise-whitening for the data. However, it seems that the repetitious operation of noise-whitening may not be necessary. In fact, a tactful transformation to the index $Q(\mathbf{X})$ can reduce the computational complexity greatly. According to the basic property of determinant

$$\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det(\mathbf{A}) * \det(\mathbf{B}) \tag{6}$$

(4) can be transformed into

$$Q(\mathbf{X}) = \det\left(\frac{1}{M}\mathbf{F}_{\mathbf{N}}^T \mathbf{X}\mathbf{X}^T \mathbf{F}_{\mathbf{N}}\right) = \det\left(\frac{1}{M}\mathbf{F}_{\mathbf{N}}\mathbf{F}_{\mathbf{N}}^T \mathbf{X}\mathbf{X}^T\right). \tag{7}$$

Considering $\mathbf{F}_{\mathbf{N}} = \mathbf{U}\Delta^{-\frac{1}{2}}$, then $\mathbf{F}_{\mathbf{N}}\mathbf{F}_{\mathbf{N}}^T = \mathbf{U}\Delta^{-\frac{1}{2}}\Delta^{-\frac{1}{2}}\mathbf{U}^T = \Sigma_{\mathbf{N}}^{-1}$. Therefore, (7) becomes

$$\begin{aligned} Q(\mathbf{X}) &= \det\left(\frac{1}{M}\mathbf{F}_{\mathbf{N}}\mathbf{F}_{\mathbf{N}}^T \mathbf{X}\mathbf{X}^T\right) = \det\left(\frac{1}{M}\Sigma_{\mathbf{N}}^{-1}\mathbf{X}\mathbf{X}^T\right) \\ &= \det\left(\Sigma_{\mathbf{N}}^{-1}\Sigma_{\mathbf{X}}\right) = \frac{\det(\Sigma_{\mathbf{X}})}{\det(\Sigma_{\mathbf{N}})} \end{aligned} \tag{8}$$

where $\boldsymbol{\Sigma}_{\mathbf{X}}$ and $\boldsymbol{\Sigma}_{\mathbf{N}}$ are the covariance matrices for original data and noise, respectively.

It can be seen that the numerator of $Q(\mathbf{X})$ is the determinant of covariance matrix for original data which measures the *total information* while the denominator indicates the *noise information*. As a result, high correlation of bands will result in small numerator and low SNR will cause large denominator. Therefore, to maximize $Q(\mathbf{X})$ can be expected to obtain the bands with high SNR as well as low correlation.

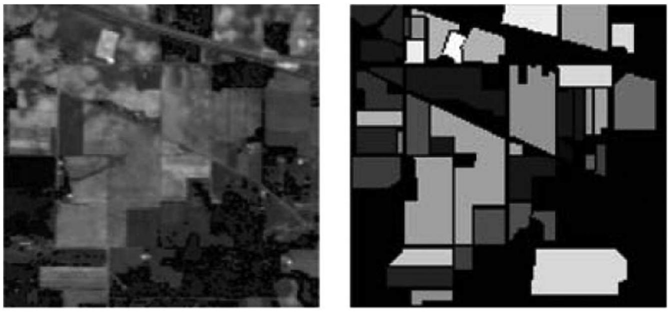
Furthermore, through the transformation, MNBS is based only on the covariance matrices of original data and noise and no whitening is involved. Thus, the computational complexity of MNBS has been greatly reduced. The pseudo code for MNBS is presented as follows:

Algorithm 1: The MNBS Algorithm

Input: Observations $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_L]^T$ and the number of selected bands n, where $\mathbf{x}_i$ is the column vector constructed by the i-th band.

1. Noise estimation. Estimate noise $\mathbf{N}$ based on (2) or other methods.

2. Set the indicating variable $k = L$, calculate the covariance matrices $\boldsymbol{\Sigma}_{\mathbf{X}}$ and $\boldsymbol{\Sigma}_{\mathbf{N}}$ for original data and noise respectively.



(a) (b)

Fig. 2. (a) The 18th band of the data, (b) Ground truth of the data.

%% Main loop

While $k > n$

3. Remove each row and column from $\mathbf{\Sigma_X}$ and $\mathbf{\Sigma_N}$ tentatively which results in k sub-matrices denoted as $\mathbf{\Sigma_X}^{(i)}$ and $\mathbf{\Sigma_N}^{(i)}$. More specifically, $\mathbf{\Sigma_X}^{(i)}$ is obtained by removing the i-th row and i-th column of $\mathbf{\Sigma_X}$ while $\mathbf{\Sigma_N}^{(i)}$ is the sub-matrix of $\mathbf{\Sigma_N}$.

4. Calculate the index for each band according to (8).

for $i = 1 : k$

$\mathbf{Q}(i) = \frac{\det\left(\mathbf{\Sigma_X}^{(i)}\right)}{\det\left(\mathbf{\Sigma_N}^{(i)}\right)}$

endfor

5. Determine the band to be removed. $index = \arg\max_i \mathbf{Q}(i)$.

6. Update the covariance matrices by $\Sigma_\mathbf{X} = \Sigma_\mathbf{X}^{(index)}$ and $\Sigma_\mathbf{N} = \Sigma_\mathbf{N}^{(index)}$

7. $k \leftarrow k - 1$

endwhile

8. Output: The left bands $[\mathbf{x}_{a_1}, \mathbf{x}_{a_2}, \ldots, \mathbf{x}_{a_n}]$ are the final selection.

IV. Experiments

In this section, some experiments are conducted to test the performance of MNBS. Meanwhile, two other unsupervised band selection methods, namely, MDCM and minimum variance principal components analysis (MVPCA) [22] are compared with MNBS in terms of accuracy. Two real hyperspectral data sets acquired by airborne visible infrared imaging spectrometer (AVIRIS) were used in these experiments.

A. Experiments on Indian Pines Data Set

The data used in this experiment were acquired by AVIRIS in 1992, which contains 145×145 pixels and 220 bands with a wavelength ranges from 400 to 2500 nm (Fig. 2). Generally, the “noisy” bands (1–3, 103–112, 148–165, and 217–220) should be manually removed in advance [31]. But in our experiments, all the bands are purposefully kept as natural test for these unsupervised band selection methods. The number of bands to be selected was set to 15 according to [32].

1) Experiment 1: Add Noise Bands: In this experiment, three simulated noise bands generated by Gaussian distribution are artificially added to the data. The mean value of noise is set as the average value of the data cube while the variance of noise equals to that of the band (band 29) with maximum variance. These noise bands are inserted as *51*-th, *101*-th, and *151*-th band of the new data, respectively. As a result, the new data set has a total of 223 bands. The bands selected by MDCM, MVPCA, and MNBS are listed in Table I.

TABLE I
Bands Selected by MDCM, MVPCA, and MNBS

Method	MDCM	MVPCA	MNBS
Bands	1, 18, 20, 29, 34, 35, 39, ***51***, 58, 62, 76, 89, 90, ***101***, ***151***	22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 42, ***51***, ***101***, ***151***	9, 17, 18, 28, 37, 53, 59, 65, 76, 83, 85, 122, 135, 171, 184

TABLE II
Results for Different SNRs

Method		MDCM	MVPCA	MNBS
Bands	∞ dB	1, 18, 20, 23, 29, 32, 34, 35, 39, 57, 61, 62, 76, 88, 89	21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 41, 42	9, 17, 18, 28, 37, 52, 58, 64, 75, 82, 84, 120, 133, 168, 181
	30 dB	1, 18, 20, 23, 29, 32, 34, 35, 39, 57, 61, 62, 76, 88, 89	21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 41, 42	9, 17, 18, 28, 37, 52, 58, 64, 75, 82, 84, 120, 133, 168, 181
	20 dB	1, ***17***, 18, 20, 23, 29, 32, 34, 35, 40, 57, 61, 75, 88, 89	21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 41, 42	9, ***16***, 18, 28, ***38***, ***51***, 58, 64, 75, 82, 84, 120, 133, 168, 181
	15 dB	1, ***17***, 18, 20, 23, 29, 34, 35, ***37***, ***52***, 57, 61, 75, 88, 89	21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 41, 42	9, ***16***, 18, 28, ***38***, ***51***, 58, 64, 75, 82, 84, 120, 133, 168, 181
	10 dB	1, ***17***, 18, 20, 29, 34, 35, ***37***, 39, ***52***, 57, 61, 75, 88, 89	***17***, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 42, ***52***	9, ***16***, 18, 28, ***38***, ***51***, 58, 64, 75, 82, 84, 120, 133, 168, 181

The results listed in Table I show that MNBS can remove the artificial “noise bands” (band 51, 101, and 151) successfully. This is because the noise bands have little spatial correlation; thus, they can be estimated by shift-difference method. By noise-whitening, these bands are heavily suppressed. As a result, these noise bands have few chances to be selected in MNBS.

On the contrary, both MDCM and MVPCA are preferred to these noise bands rather than those with high quality. MDCM is especially sensitive to noise, as it has selected a natural noise band (band 1). It can be ascribed to the fact that both methods are based on bands’ information and consider little about data quality.

2) Experiment 2: Add Noise to Bands: In this test, we add Gaussian noise with different levels ($\mathrm{SNR} = 30$, 20, 15 and 10 dB) to three real bands, respectively. In order to test the effectiveness of MNBS, the noises were added to 17th, 37th, and 52th bands, because these three bands were originally selected by MNBS but were neglected by MDCM and MVPCA. We conducted the test to study that if these bands can be removed by MNBS when some noises were added. It should be noted that different from experiment 1, there are a total of 220 bands this time. The results are listed in Table II.

It can be seen from Table II that the artificially added noises did not influence the results for all the compared methods when the noises are of low level ($\mathrm{SNR} = 30$ dB). However, MNBS can recognize and remove these noise bands (bands 17, 37, and 52) accurately when SNR is equal to or less than 20 dB. In addition, the substituted bands (bands 16, 38, and 51) are adjacent to the

(a) (b) (c)

Fig. 3. Bands with Gaussian noise (15 dB): (a) band 17; (b) band 37; and (c) band 52.

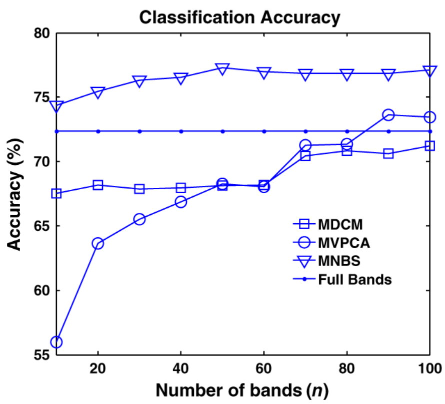


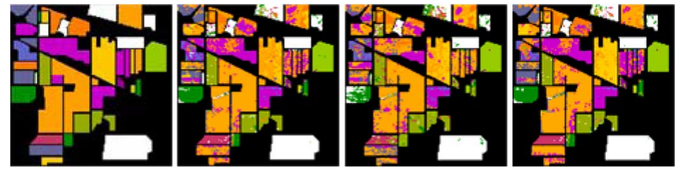
Fig. 4. SVM classification accuracy based on bands selected by different methods with different numbers.

noise bands. These results demonstrate that MNBS can keep away from the noise bands and is quite robust. The removal of the noise bands is reasonable as these bands are heavily corrupted by noise and thus have limited useful information (see Fig. 3).

Both MDCM and MVPCA selected whole or part of the noise bands due to their negligence of the quality of image when SNR is equal to or less than 20 dB. The addition of the random noise significantly increases the variance of bands and therefore, increases the "weights" of the noise bands when the variance is considered as the band selection criterion. In addition, it is again shown that MVPCA is more robust to noise compared with MDCM (Table II). In fact, MDCM selected a natural noise band (band 1) even if no noise is added (as shown in Tables I and II). From the authors' point of view, this is because MDCM takes both the variance and correlation as the criterion. There is a little correlation of noises; thus, the same levels of noise have more weights in MDCM than those in MVPCA.

3) Experiment 3: Accuracy Test: In this experiment, the accuracy of bands selected by MDCM, MVPCA, and MNBS are compared. The selected bands are used for supervised classification [support vector machine (SVM)] and the classification accuracy is used as the comparison criterion. Sixteen different ground-classes' training samples (20%) are randomly chosen according to ground truth [Fig. 2(b)]. The SVM tool embedded in ENVI is used for the classification with radial basis function (RBF) as the kernel. The classification accuracy (%) for different methods is plotted in Fig. 4 and the classification results based on 15 bands selected by these methods are shown in Fig. 5.

From Figs. 4 and 5, we can see that the classification result based on bands selected by MNBS is the most accurate while that



(a) (b) (c) (d)

Fig. 5. SVM classification results based on 15 bands selected by different methods. (a) Ground truth. (b) MDCM. (c) MVPCA. (d) MNBS.

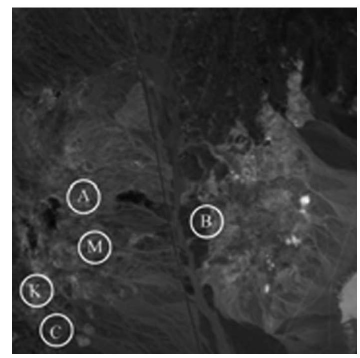


Fig. 6. Five minerals: Alunite (A); Buddingtonite (B); Calcite (C); Kaolinite (K); and Muscovite (M) are circled in Cuprite scene.

based on MVPCA has the lowest accuracy. Moreover, the classification results from bands selected by MNBS always outperform those from full bands. The effectiveness of MNBS can be attributed to the fact that MNBS considers not only the information and correlation but also the noises of band. On the other hand, MDCM is based on bands' information and correlation while MVPCA only takes bands' information into account. Therefore, classification results based on the bands obtained by these two methods are less accurate.

B. Experiments on Cuprite Scene

In this section, another well-known AVIRIS image which acquired over Cuprite mining site, Nevada is used to further compare the three unsupervised band selection methods. Different from previous experiments, the noise bands and water absorption bands (1–3, 105–115, and 150–170) [18] are all removed. Therefore, a total of 189 bands are used in the experiments. Again, the classification accuracy is used to measure the effectiveness of the bands selected by these methods. Five minerals, namely, Alunite (A), Buddingtonite (B), Calcite (C), Kaolinite (K), and Muscovite (M) are classified by supervised methods, constrained linear discriminant analysis (CLDA) [33]–[35]. Their approximate spatial locations of these five mineral are shown in Fig. 6.

1) Experiment 4: Accuracy Test: To evaluate the band selection performance, the classification results based on full bands and bands selected by the methods are taken into comparison. In addition, the training samples of these classes are artificially selected by ENVI ROI tools. The supervised classification results derived from different band sets are shown in Fig. 7.

From the classification maps in Fig. 7, it can be seen that the classification based on the bands selected by MNBS is visually

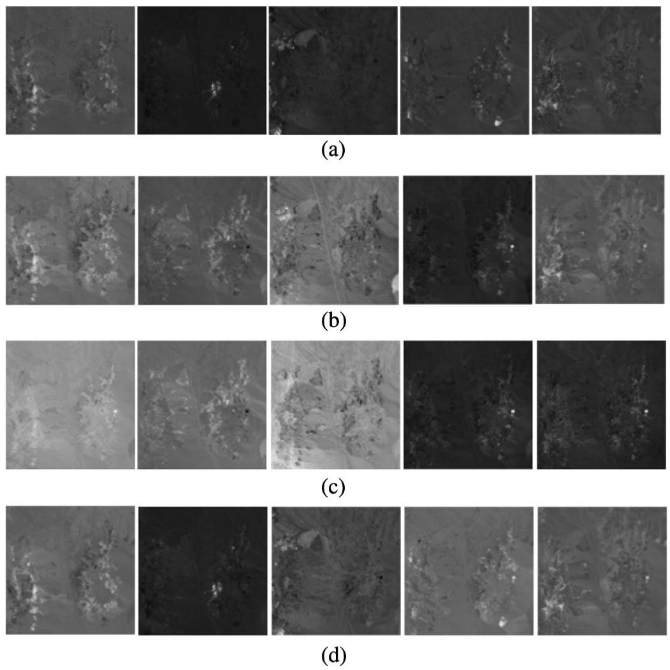


Fig. 7. CLDA classification maps of (from left to right) A, B, C, K, and M: Classification maps of bands (a) using full 189 bands; (b) using 20 bands selected by MDCM (band 1, 2, 3, 4, 5, 6, 9, 10, 16, 22, 29, 30, 36, 39, 53, 74, 94, 123, 154, 159); (c) using 20 bands selected by MVPCA (bands 74, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 98, 189); and (d) using 20 bands selected by MNBS (bands 5, 10, 15, 23, 29, 38, 53, 74, 80, 93, 107, 124, 134, 137, 149, 154, 157, 161, 166, 172).

close to that based on full bands. But the background suppression of the MNBS-based classification is a little worse. On the other hand, both MDCM- and MVPCA-based classifications are deviating from the full bands' result, although it is shown that MDCM is better than MVPCA in terms of accuracy.

As no pixel-level ground truth maps are available, we use the correlation coefficient (CC) of the classification maps from full bands and those from the selected band to quantitatively compare the accuracy of the methods mentioned above [36]. Fig. 8 plots the CCs with different numbers of selected bands.

Two conclusions can be drawn from Fig. 8: first, the "classification accuracy" (measured by the correlation between the result from full bands and that from the selected bands, not real classification accuracy since no pixel-level ground truth is available) is increasing with the number of selected bands; second, MNBS is the most accurate band selection method in terms of classification accuracy while the performance of MVPCA is the worst. The reason is MVPCA considers only the variance, MDCM considers variance and correlation while MNBS considers variance, correlation, and noise as the band selection criterion. As a result, the bands selected by MNBS are the most effective for classification.

2) Experiment 5: Efficiency Evaluation: The computational time of these methods is also compared (Table III). From Table III, we can see that MVPCA is the most fast since it is a band prioritization method. The improvement proposed in Section III-C is very effective and it speeds up MNBS about 21 times. The speed of MNBS is acceptable, although it is shown that it is the lowest among these methods.

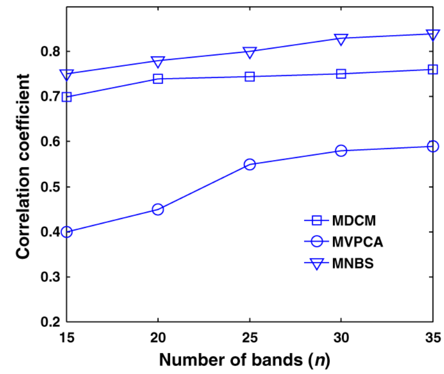


Fig. 8. CCs between the classification maps from the full bands and those from the selected bands.

TABLE III
COMPUTATION TIME OF THE METHODS

Method	MDCM	MVPCA	MNBS (improved)	MNBS (original)
Time (s)	10.93	2.21	14.08	292.41

V. CONCLUSION

In this paper, an unsupervised band selection method called MNBS has been proposed. MNBS adopts a newly defined hyperspectral data cube quality Q as the selection criterion. Since Q integrates NAPCs with MDCM, it takes both SNRs and correlation of bands into consideration. Therefore, bands selected by MNBS are of good cube qualities, namely, high SNRs and low correlation. The experiments show that MNBS can remove the bands with low qualities and therefore obtain high classification accuracy. However, MNBS is a little time-consuming although an improvement has been proposed. Thus, a more efficient version of MNBS is expected to be studied in the future.

REFERENCES

[1] I. Jolliffe, *Principal Component Analysis*. Hoboken, NJ, USA: Wiley, 2005.
[2] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Hoboken, NJ, USA: Wiley, 2004.
[3] A. A. Green *et al.*, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 1, pp. 65–74, Jan. 1988.
[4] C. Lee and D. A. Landgrebe, "Feature extraction based on decision boundaries," *IEEE Trans.Pattern Anal. Mach. Intell.*, vol. 15, no. 4, pp. 388–400, Apr. 1993.
[5] B.-C. Kuo, "Nonparametric weighted feature extraction for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 5, pp. 1096–1105, May 2004.
[6] I. Thomas *et al.*, "A review of multi-channel indices of class separability," *Int. J. Remote Sens.*, vol. 8, no. 3, pp. 331–350, 1987.
[7] 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.
[8] T.-M. Tu *et al.*, "A fast two-stage classification method for high-dimensional remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 1, pp. 182–191, Jan. 1998.
[9] L. Bruzzone and S. B. Serpico, "A technique for feature selection in multiclass problems," *Int. J. Remote Sens.*, vol. 21, no. 3, pp. 549–563, 2000.
[10] K. Nirmal, "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1552–1565, Jul. 2004.

[11] Z. Guo *et al.*, "A hypergraph based semi-supervised band selection method for hyperspectral image classification," in *Proc. IEEE Int. Conf. Image Process. (ICIP'13)*, Melbourne, Australia, 2013, pp. 3137–3141.
[12] P. S. Thenkabail *et al.*, "Selection of hyperspectral narrowbands (HNBs) and composition of hyperspectral two band vegetation indices (HVIs) for biophysical characterization and discrimination of crop types using field reflectance and hyperion/EO-1 data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 427–439, Apr. 2013.
[13] B.-C. Kuo *et al.*, "A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 1, pp. 317–326, Jan. 2014.
[14] S. Jia *et al.*, "Unsupervised band selection for hyperspectral imagery classification without manual band removal," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 531–543, Apr. 2012.
[15] J. C. Russ, *The Image Processsing Handbook*, 6th ed. Boca Raton, FL, USA: CRC Press, 2010.
[16] D. J. Wiersma and D. A. Landgrebe, "Analytical design of multispectral sensors," *IEEE Trans. Geosci. Remote Sens.*, vol. GE-18, no. 2, pp. 180–189, Apr. 1980.
[17] J. C. Price, "Band selection procedure for multispectral scanners," *Appl. Opt.*, vol. 33, no. 15, pp. 3281–3288, 1994.
[18] C.-I. Chang and S. Wang, "Constrained band selection for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1575–1585, Jun. 2006.
[19] P. Mitra, C. Murthy, and S. K. Pal, "Unsupervised feature selection using feature similarity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 301–312, Mar. 2002.
[20] C. Sheffield, "Selecting band combinations from multispectral data," *Photogramm. Eng. Remote Sens.*, vol. 51, no. 6, pp. 681–687, 1985.
[21] P. S. Chacvez, G. L. Berlin, and B. Sowers, "Statistical method for selecting landsat MSS ratios," *J. Appl. Photogr. Eng.*, vol. 1, no. 8, pp. 23–30, 1982.
[22] C.-I. Chang *et al.*, "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.
[23] P. Bajcsy and P. Groves, "Methodology for hyperspectral band selection," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 7, pp. 793–802, 2004.
[24] J. M. Sotoca, F. Pla, and J. S. Sánchez, "Band selection in multispectral images by minimization of dependent information," *IEEE Trans. Syst. Man Cybern. C: Appl. Rev.*, vol. 37, no. 2, pp. 258–267, Mar. 2007.
[25] P. Latorre-Carmona *et al.*, "Effect of denoising in band selection for regression tasks in hyperspectral datasets," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 473–481, Apr. 2013.
[26] K. Tan *et al.*, "Hyperspectral image classification using band selection and morphological profiles," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 1, pp. 40–48, Jan. 2014.
[27] J. B. Lee, A. S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote sensing data by a noise-adjusted principal components transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 3, pp. 295–304, May 1990.
[28] T. Marill and D. M. Green, "On the effectiveness of receptors in recognition systems," *IEEE Trans. Inf. Theory*, vol. 9, no. 1, pp. 11–17, Jan. 1963.
[29] P. Switzer and A. A. Green, "Min/max autocorrelation factors for multivariate spatial imagery," Tech. Rep. 6, Dept. Statist., Stanford Univ., 1984, pp. 1–14.
[30] P. Comon, "Independent component analysis: A new concept?," *Signal Process.*, vol. 36, no. 287–314, 1994.
[31] R. Archibald and G. Fann, "Feature selection and classification of hyperspectral images with support vector machines," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 674–677, Oct. 2007.
[32] Y. Qian, F. Yao, and S. Jia, "Band selection for hyperspectral imagery using affinity propagation," *IET Comput. Vis.*, vol. 3, no. 4, pp. 213–222, 2009.
[33] 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.
[34] 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.
[35] Q. Du and C.-I. Chang, "A linear constrained distance-based discriminant analysis for hyperspectral image classification," *Pattern Recog.*, vol. 34, no. 2, pp. 361–373, 2001.
[36] R. V. Platt and A. F. Goetz, "A comparison of AVIRIS and Landsat for land use classification at the urban fringe," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 7, pp. 813–819, 2004.

Kang Sun received the M.S. degree in agricultural science from Shenyang Agriculture University, Shenyang, China, in 2010. He is currently pursuing the Ph.D. degree in signal processing at the Institute of Electronics, Chinese Academy of Sciences, Beijing, China, and the University of Chinese Academy of Sciences, Beijing, China.

His research interests include on feature extraction, spectral unmixing, and intelligence computation in hyperspectral remote sensing data analysis.

Xiurui Geng received the Ph.D. degree in hyperspectral remote sensing from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2005.

Currently, he is an Associate Professor with the Key Laboratory of Technology in Geo-Spatial Information Process and Application Systems, Institute of Electronics, Chinese Academy of Sciences. His research interests include hyperspectral data understanding and algorithm development in feature extraction, mixed pixel analysis, image matching, target detection and classification by introducing new algebra, geometry, analysis, and statistics viewpoints.

Luyan Ji received the M.S. degree in electrical engineering from the University of Science and Technology, Hefei, China, in 2008, and the Ph.D. degree in signal processing from the Institute of Electronics, Chinese Academic of Sciences, Beijing, China, in 2013.

Currently, she is a Postdoctoral Research Fellow with the Centre for Earth System Science, Tsinghua University, Beijing, China. Her research interests include issues related to the spectral–temporal characteristics of soil water-holding capacity, the study of spectral mixing analysis, and feature extraction in hyperspectral remote sensing application.

Yun Lu received the Ph.D. degree in materials science and engineering from the School of Nano and New Materials Engineering, Changwon National University, Changwon, South Korea, in 2007.

Currently, she is an Associate Professor with the School of Materials Science and Engineering, Anhui University of Technology, Anhui, China. Her research interests include computational materials science, multiscale material design, and surface modification of materials.