Constrained Band Selection for Hyperspectral Imagery

Chein-I Chang, *Senior Member, IEEE,* and Su Wang, *Student Member, IEEE*

***Abstract*—Constrained energy minimization (CEM) has shown effective in hyperspectral target detection. It linearly constrains a desired target signature while minimizing interfering effects caused by other unknown signatures. This paper explores this idea for band selection and develops a new approach to band selection, referred to as constrained band selection (CBS) for hyperspectral imagery. It interprets a band image as a desired target signature vector while considering other band images as unknown signature vectors. As a result, the proposed CBS using the concept of the CEM to linearly constrain a band image, while also minimizing band correlation or dependence provided by other band images, is referred to as CEM-CBS. Four different criteria referred to as Band Correlation Minimization (BCM), Band Correlation Constraint (BCC), Band Dependence Constraint (BDC), and Band Dependence Minimization (BDM) are derived for CEM-CBS.. Since dimensionality resulting from conversion of a band image to a vector may be huge, the CEM-CBS is further reinterpreted as linearly constrained minimum variance (LCMV)-based CBS by constraining a band image as a matrix where the same four criteria, BCM, BCC, BDC, and BDM, can be also used for LCMV-CBS. In order to determine the number of bands required to select p, a recently developed concept, called virtual dimensionality, is used to estimate the p. Once the p is determined, a set of p desired bands can be selected by the CEM/LCMV-CBS. Finally, experiments are conducted to substantiate the proposed CEM/LCMV-CBS four criteria, BCM, BCC, BDC, and BDM, in comparison with variance-based band selection, information divergence-based band selection, and uniform band selection.**

***Index Terms*—Band correlation constraint (BCC), band correlation minimization (BCM), band dependence constraint (BDC), band dependence minimization (BDM), constrained band selection (CBS), constrained energy minimization (CEM), linearly constrained minimum variance (LCMV), virtual dimensionality (VD).**

I. Introduction

A hyperspectral image is an image cube with each image pixel represented by a column vector where each of column components is a pixel imaged by a particular spectral channel. As a result, data volume to be processed for a hyperspectral image is generally huge and enormous. Its computational complexity is also expected to be very high. In order to mitigate this problem, several approaches have been investigated by looking into how to remove information redundancy resulting from highly correlated bands. One common practice is data dimensionality reduction (DR) which implements a transform to reduce data dimensions in accordance with a certain criterion such as data variance performed by the principal components analysis (PCA). Two major issues arise from such DR approaches. One is the number of data dimensions required for DR to avoid significant loss of information. The other is that, since the data after DR have been transformed and are, therefore, no longer original data, some crucial and critical information may have been compromised and distorted.

An alternative to DR is band selection, which selects appropriate bands from the original set of spectral bands that can well represent original data. Compared to DR, the band selection has an advantage of preserving original information from the data. However, it also suffers from two similar issues encountered in the DR. One is also how many bands needed to be selected in order to preserve necessary information. The other is the criterion to be used for band selection. This paper investigates these two issues for band selection.

Since the first issue is very difficult and challenging, most approaches developed for band selection, such as [1]–[5], have devoted to the second issue while ignoring the first issue. This paper addresses this first issue by a new concept of virtual dimensionality (VD), which was recently introduced in [6] to estimate the number of spectrally distinct signatures in data. If we interpret that one signal source can be only accommodated by one single and separate dimension, it requires at least the same number of dimensions as the VD to accommodate distinct signal sources. With this interpretation, the VD has shown to be an effective measure in estimating the number of dimensions required to be retained for DR [7]–[9]. By the same token, we can use the VD to resolve the issue of number of bands required to be retained for band selection. As for the second issue, this paper develops a new approach, called constrained band selection (CBS), which is completely different from commonly used variance-based [4] or information theoretic criteria-based band-selection methods [1]–[3]. It can be considered as a constrained band correlation/dependence minimization approach, which linearly constrains a band while minimizing the correlation or dependency of this particular band with other bands in a hyperspectral image. The idea can be traced back to the linearly constrained minimum variance (LCMV) developed by Frost in [10] for adaptive beamforming in passive array processing, which makes use of a specific gain to constrain an finite impulse response (FIR) filter to look for

Manuscript received July 1, 2005; revised December 11, 2005.

C.-I Chang is with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD 21250 USA, and also with the Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan, R.O.C. (e-mail: cchang@umbc.edu).

S. Wang is with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250 USA.

Digital Object Identifier 10.1109/TGRS.2006.864389

signals coming from a particular direction from an array of sensors. When the gain is constrained to a single direction with unity, the resulting FIR filter is called the minimum variance distortionless response (MVDR) beamformer. One early application of the MVDR was the constrained energy minimization (CEM) [6], [11], [12], which interpreted an array of sensors as a bank of spectral channels and the desired signal direction as a desired target signature. The rationale of the CBS presented in this paper is also derived from the LCMV and CEM. Instead of linearly constraining the signal direction of interest, as is done in the LCMV and the desired target signature, as is done in the CEM, the CBS linearly constrains a particular band of interest while minimizing the band correlation or band dependence resulting from other bands in the sense of least-square error (LSE). Accordingly, the larger the LSE, the higher the band correlation/dependence of the particular band image with/on other band images. As a result, two versions of the CBS can be further developed, referred to as CEM-based CBS and LCMV-based CBS, respectively. When they are used for band selection, bands to be selected are based on the LSE generated by either the CEM-CBS or the LCMV-CBS.

There are two ways to implement the proposed CEM/LCMV-CBS. One is to convert a band image to a band image vector and the data sample correlation matrix used in the CEM is then replaced with a correlation matrix obtained by averaging the individual band image correlation matrix. Since a band image is generally large, the size of its converted vector will be huge and may further cause a computational problem. In order to cope with this difficulty, an alternative approach is to use the LCMV [6], [10], [12] to alleviate this dilemma. Rather than constraining a band image as a vector, the LCMV-CBS constrains a band image as a matrix where a constraint vector is imposed on each column of the band image matrix. As a result, the computational complexity can be cut by three quarters, and, thus, reduced substantially. The experimental results show that, while both the proposed LCMV-CBS and CEM-CBS perform very similarly, the LCMV-CBS does offer a significant advantage over the CEM-CBS in computation. In order to evaluate the performance of our proposed CEM/LCMV-CBS, extensive experiments are conducted to demonstrate that the CEM/LCMV-CBS, indeed, performs relatively well, and also effectively, compared to uniform band selection and band-selection methods in [1]–[4].

The remainder of this paper is organized as follows. Section II develops a new approach to band selection, called constrained band selection, which is based on the CEM. Section III presents an alternative to implement the CEM-CBS, called LCMV-CBS, which eases computational problem encountered in the CEM-CBS. Section IV conducts experimental study in three applications to evaluate the proposed CBS for performance analysis. Section V concludes some remarks.

II. CEM-CBS

In this section, we develop a new approach to band selection, referred to as CEM-based constrained band selection (CEM-CBS) which is derived from the concept of the constrained energy minimization (CEM) [6], [11], [12].

A. CEM-CBS

Let $\{\mathbf{B}_l\}_{l=1}^{L}$ be the set of all band images in a hyperspectral image cube where L is the total number of bands. Assume that the size of all the band images is $M \times N$. Each band image $\mathbf{B}_l$ can be represented by a column vector of dimension MN, denoted by $\mathbf{b}_l$. Let $\mathbf{w}_l$ be an MN-dimensional column vector that is used to specify a finite impulse response (FIR) filter designed for the band image vector $\mathbf{b}_l$ and y_l be the filter output specified by

$$y_l = \mathbf{w}_l^T \mathbf{b}_l. \tag{1}$$

The averaged least squares filter output is given by

$$\begin{aligned}\left(\frac{1}{L}\right)\sum_{l=1}^{L} y_l^2 &= \left(\frac{1}{L}\right)\sum_{l=1}^{L}\left(\mathbf{w}_l^T\mathbf{b}_l\right)\left(\mathbf{w}_l^T\mathbf{b}_l\right)^T \\ &= \mathbf{w}_l^T\left(\left(\frac{1}{L}\right)\sum_{l=1}^{L}\mathbf{b}_l\mathbf{b}_l^T\right)\mathbf{w}_l. \end{aligned}\tag{2}$$

Let $\mathbf{Q} = (1/L)\sum_{l=1}^{L}\mathbf{b}_l\mathbf{b}_l^T$ denote the band image correlation matrix. A similar optimization problem to the constrained energy minimization (CEM) can be obtained for a constrained band-selection problem as follows:

$$\min_{\mathbf{w}_l}\left\{\mathbf{w}_l^T\mathbf{Q}\mathbf{w}_l\right\} \text{ subject to } \mathbf{b}_l^T\mathbf{w}_l = 1. \tag{3}$$

The solution to (3), $\mathbf{w}_l^{\text{CEM-CBS}}$ is given by

$$\mathbf{w}_l^{\text{CEM-CBS}} = \left(\mathbf{b}_l^T\mathbf{Q}^{-1}\mathbf{b}_l\right)^{-1}\mathbf{Q}^{-1}\mathbf{b}_l. \tag{4}$$

Alternatively, we can exclude the band image $\mathbf{b}_l$ from the band image correlation matrix $\mathbf{Q}$ and further define $\widetilde{\mathbf{Q}} = (1(L-1))\sum_{j=1,j\neq l}^{L}\mathbf{b}_j\mathbf{b}_j^T$ as the band image dependence matrix. Replacing $\mathbf{Q}$ in (3) with $\widetilde{\mathbf{Q}}$ results in a similar constrained band-selection problem

$$\min_{\mathbf{w}_l}\left\{\mathbf{w}_l^T\widetilde{\mathbf{Q}}\mathbf{w}_l\right\} \text{ subject to } \mathbf{b}_l^T\mathbf{w}_l = 1. \tag{5}$$

The solution to (5), $\widetilde{\mathbf{w}}_l^{\text{CEM-CBS}}$ is the same as the one in (4) with the $\mathbf{Q}$ replaced by $\widetilde{\mathbf{Q}}$, which is given by

$$\widetilde{\mathbf{w}}_l^{\text{CEM-CBS}} = \left(\mathbf{b}_l^T\widetilde{\mathbf{Q}}^{-1}\mathbf{b}_l\right)^{-1}\widetilde{\mathbf{Q}}^{-1}\mathbf{b}_l. \tag{6}$$

B. Criteria for Band Selection

In the previous section, two versions of the CEM-CBS were developed and described by (3) and (5). Their solutions, $\mathbf{w}_l^{\text{CEM-CBS}}$ and $\widetilde{\mathbf{w}}_l^{\text{CEM-CBS}}$ specified by (4) and (6) can then be used as criteria for band selection. In this section, four different criteria are proposed for the CBS.

1) Band Correlation Minimization (BCM): According to (3), the LSE resulting from the $\mathbf{w}_l^{\text{CEM-CBS}}$ is $\rho_l = \left(\mathbf{w}_l^{\text{CEM-CBS}}\right)^T\mathbf{Q}\mathbf{w}_l^{\text{CEM-CBS}}$, which represents minimal correlation of the band image vector $\mathbf{b}_l$ with the entire hyperspectral image in the least-square sense by constraining the band image $\mathbf{b}_l$. This implies that the larger the ρ_l, the higher correlated the band image $\mathbf{b}_l$ with the hyperspectral image, thus, the more important the band $\mathbf{b}_l$. For example, a noisy band image generally has little correlation with the entire hyperspectral image, and, thus, has smaller CEM-LSE. The use of $\{\rho_l\}_{l=1}^{L}$ to select the band images $\{\mathbf{B}_l\}_{l=1}^{L}$ is called BCM.

2) Band Dependence Minimization (BDM): Similar to BCM, the LSE resulting from $\tilde{\mathbf{w}}_l^{\text{CEM-CBS}}$ is $\tilde{\rho}_l = \left(\tilde{\mathbf{w}}_l^{\text{CEM-CBS}}\right)^T \tilde{\mathbf{Q}}\tilde{\mathbf{w}}_l^{\text{CEM-CBS}}$, which indicates the dependence of the band image $\mathbf{b}_l$ on all band images other than $\mathbf{b}_l$. The larger the $\tilde{\rho}_l$, the more dependent the band image $\mathbf{b}_l$ on other band images, thus, the more significant the band. Therefore, using $\{\tilde{\rho}_l\}_{l=1}^{L}$ for band selection is called BDM.

3) Band Correlation Constraint (BCC): The BCM and BDM are criteria to measure the LSE resulting from CEM. As an alternative, we can measure the degree of a band image vector $\mathbf{b}_k$ deviated from the constraint imposed by the CEM, $\mathbf{b}_l^T\mathbf{w}_l = 1$, which is $\mathbf{b}_k^T\mathbf{w}_l$. With this interpretation, a new criterion can be defined by calculating the sum of the deviations of all bands $\{\mathbf{b}_k\}_{k=1,k\neq l}^{L}$, $\left(\mathbf{w}_l^{\text{CEM-CBS}}\right)^T\mathbf{b}_k$ other than $\mathbf{b}_l$ from the CEM-imposed constraint $\mathbf{b}_l^T\mathbf{w}_l = 1$ via $\mathbf{w}_l^{\text{CEM-CBS}}$ as follows:

$$\begin{aligned}\eta_l &= \sum_{k=1,k\neq l}^{L}\left(\mathbf{w}_l^{\text{CEM-CBS}}\right)^T\mathbf{b}_k \\ &= \left[\sum_{k=1}^{L}\left(\mathbf{w}_l^{\text{CEM-CBS}}\right)^T\mathbf{b}_k\right]-1 \\ &= \left(\mathbf{w}_l^{\text{CEM-CBS}}\right)^T\left[\sum_{k=1}^{L}\mathbf{b}_k\right]-1. \end{aligned} \tag{7}$$

It should be noted that the one in (7) is the result of $\mathbf{b}_l^T\mathbf{w}_l = 1$ form (3). Additionally, due to the constraint $\mathbf{b}_l^T\mathbf{w}_l = 1$ in (3), a band image vector $\mathbf{b}_k$ has less correlation with band image vector $\mathbf{b}_l$ if its band constraint $\mathbf{w}_l^T\mathbf{b}_k$ is far away from 1. In other words, the closer the $\mathbf{w}_l^T\mathbf{b}_k$ to 1, the higher the correlation of $\mathbf{b}_k$ to $\mathbf{b}_l$. In light of (7), η_l represents the degree of the band constraint of a particular band $\mathbf{b}_l$ on band correlation to the entire hyperspectral image. The use of $\{\eta_l\}_{l=1}^{L}$ for band selection is called BCC.

4) Band Dependence Constraint (BDC): Analogous with BCC, a criterion can be also defined via $\tilde{\mathbf{w}}_l^{\text{CEM-CBS}}$ by

$$\begin{aligned}\tilde{\eta}_l &= \sum_{k=1,k\neq l}^{L}\left(\tilde{\mathbf{w}}_l^{\text{CEM-CBS}}\right)^T\mathbf{b}_k \\ &= \sum_{k=1}^{L}\left(\tilde{\mathbf{w}}_l^{\text{CEM-CBS}}\right)^T\mathbf{b}_k - 1 \\ &= \left(\tilde{\mathbf{w}}_l^{\text{CEM-CBS}}\right)^T\left[\sum_{k=1}^{L}\mathbf{b}_k\right]-1. \end{aligned} \tag{8}$$

The use of $\{\tilde{\eta}_l\}_{l=1}^{L}$ to measure the degree of the band constraint of a particular band $\mathbf{b}_l$ on band dependence on all other band images $\{\mathbf{b}_k\}_{k=1,k\neq l}^{L}$ is called BDC.

III. LCMV-CBS

One disadvantage of the CEM-CBS is the enormous size of vectors converted from band images that causes tremendous computing time. For example, it requires a vector with 4×10^4 dimensions to represent a band image with size of 200×200 pixels. In order to mitigate this dilemma, a linearly constrained minimum variance LCMV-CBS is developed in this section. Instead of constraining a band image as a vector, the LCMV-CBS constrains a band image as an image matrix without vector conversion. More specifically, assume that $\mathbf{r}_{l,C_1}, \mathbf{r}_{l,C_2}, \ldots, \mathbf{r}_{l,C_N}$ are N columns of the lth band image $\mathbf{B}_l$, which has M rows and N columns and each column $\mathbf{r}_{l,Cn}$ is represented by the M-dimensional column vector $\mathbf{r}_{l,C_n} = (r_{l,1n}, r_{l,2n}, \ldots, r_{l,Mn})$. In this case, the lth band image $\mathbf{B}_l$ can be further expressed by a matrix given by

$$\begin{aligned}\mathbf{B}_l &= \begin{bmatrix} r_{l,11} & r_{l,12} & \cdots & r_{l,1(N-1)} & r_{l,1N} \\ r_{l,21} & r_{l,22} & \ddots & r_{l,2(N-1)} & r_{l,2N} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ r_{l,(M-1)1} & \ddots & \ddots & r_{l,(M-1)(N-1)} & r_{l,(M-1)N} \\ r_{l,M1} & r_{l,M2} & \cdots & r_{l,M(N-1)} & r_{l,MN} \end{bmatrix} \\ &= \left[\mathbf{r}_{l,C_1}\mathbf{r}_{l,C_2}\cdots\mathbf{r}_{l,C_N}\right]. \end{aligned} \tag{9}$$

Like the CEM, the goal is to design a constrained FIR linear filter with an M-dimensional weight column vector $\mathbf{v}_l = (v_{l,1}, v_{l,2}, \ldots v_{l,M})^T$ specified by a set of M filter coefficients $\{v_{l,1}, v_{l,2}, \ldots, v_{l,M}\}$ that minimizes the filter output energy subject to the following simultaneous N multiple constraints, $\mathbf{r}_{l,C_n}^T\mathbf{v}_l = \sum_{m=1}^{M} r_{l,nm}v_{l,m} = 1, 1 \leq n \leq N$ which is equivalent to

$$\mathbf{B}_l^T\mathbf{v}_l = \mathbf{1}_N \tag{10}$$

where $\mathbf{1}_N$ is an N-dimensional column vector with all 1s in its N components. It should be noted that since the weight vector $\mathbf{v}_l$ is (10) is used to constrain column vector of a band image, its dimensionality is M compared to the MN-dimensional weight vector $\mathbf{w}_l$ used in the CEM-based CBS that constrains a band image as a vector with dimensionality MN. By virtue of the N multiple constraints in (10), the CEM-CBS problem described by (3) can be rederived as the following optimization problem:

$$\min_{\mathbf{v}_l}\left\{\mathbf{v}_l^T\mathbf{v}_l\right\} \text{ subject to } \mathbf{B}_l^T\mathbf{v}_l = \mathbf{1}_N \tag{11}$$

where $\mathbf{\Sigma} = (1/L)\sum_{l=1}^{L}\mathbf{B}_l\mathbf{B}_l^T$ is the sample band correlation matrix, The problem described by (11) is referred to as LCMV-based CBS problem. The solution to (11) can be solved as

$$\mathbf{v}_l^{\text{LCMV-CBS}} = \mathbf{\Sigma}^{-1}\mathbf{B}_l\left(\mathbf{B}_l^T\mathbf{\Sigma}^{-1}\mathbf{B}_l\right)^{-1}\mathbf{1}_N \tag{12}$$

and

$$\tau_l = \left(\mathbf{v}_l^{\text{LCMV-CBS}}\right)^T\mathbf{\Sigma}\mathbf{v}_l^{\text{LCMV-CBS}} \text{ for LCMV-BCM} \tag{13}$$

plays the same role that ρ_l does for the CEM-BCM. Similar derivations to CEM-BDM can be also obtained for $\tilde{\mathbf{v}}_l^{\text{LCMV-CBS}} = \tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{B}_l\left(\mathbf{B}_l^T\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{B}_l\right)^{-1}\mathbf{1}_N$ and

$$\tilde{\tau}_l = \left(\tilde{\mathbf{v}}_l^{\text{LCMV-CBS}}\right)^T\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{v}}_l^{\text{LCMV-CBS}} \text{ for LCMV-BDM.} \quad (14)$$

Analogous with the CEM-based band correlation/dependence constraint criteria (BCC/BDC) developed in Section II

$$\zeta_l = \sum_{k=1,k\neq l}^{L}\mathbf{1}_N^T\left(\mathbf{B}_k^T\left(\mathbf{v}_l^{\text{LCMV-CBS}}\right)\right) \text{ for LCMV-BCC} \quad (15)$$

and

$$\tilde{\zeta}_l = \sum_{k=1,k\neq l}^{L}\mathbf{1}_N^T\left(\mathbf{B}_k^T\left(\tilde{\mathbf{v}}_l^{\text{LCMV-CBS}}\right)\right) \text{ for LCMV-BDC} \quad (16)$$

can be also derived for an LCMV-based band correlation/dependence constraint criteria by replacing band image vector $\mathbf{b}_l$ and CEM-based weight vectors with band image $\mathbf{B}_l$ and LCMV-based weight vectors respectively.

It should be noted that for the CEM-CBS, the dimensions of the band image correlation/dependence matrices $\mathbf{Q}$ and $\tilde{\mathbf{Q}}$ are both $MN \times MN$, while for the LCMV-CBS, the dimensions of the band image correlation/dependence matrix $\boldsymbol{\Sigma}$ and $\tilde{\boldsymbol{\Sigma}}$ are $M \times M$. That is exactly the reason why we develop the LCMC-CBS to dramatically reduce the computational complexity caused by the CEM-CBS.

One comment is worthwhile. The difference between the CEM-CBS and the LCMV-CBS lies in the fact that the former method considers a band image as a column vector via a scalar constraint, while the latter constraints each of column vector of a band image through a vector constraint. In other words, if we impose each of column in $\mathbf{B}_l$, $\mathbf{r}_{l,C_n}$ for $1 \leq n \leq N$ with the same scalar constraint $1/N$, i.e., $\mathbf{r}_{l,C_n}\mathbf{v}_l = \sum_{m=1}^{M} r_{l,mn}v_{l,m} = 1/N$, we further concatenate all the column vectors $\{\mathbf{r}_{l,C_n}\}_{n=1}^{N}$ of the lth band image $\mathbf{B}_l$ to a vector to represent the $\mathbf{B}_l$ as an MN-dimensional column vector $\mathbf{b}_l = \left[\mathbf{r}_{l,C_1}^T, \mathbf{r}_{l,C_2}^T, \ldots, \mathbf{r}_{l,C_N}^T\right]^T$ and the weight vectors $\mathbf{v}_l N$ times to an MN-dimensional weight vector $[\underbrace{\mathbf{v}_l^T, \mathbf{v}_l^T, \ldots, \mathbf{v}_l^T}_{N}]^T$, then

$$\mathbf{b}_l^T[\underbrace{\mathbf{v}_l^T, \mathbf{v}_l^T, \cdots, \mathbf{v}_l^T}_{N}] = \sum_{n=1}^{N}\sum_{m=1}^{M} r_{l,mn}v_{l,m} = \sum_{n=1}^{N}\frac{1}{N} = 1$$

which is exactly the same constraint $\mathbf{b}_l^T\mathbf{w}_l = 1$ imposed in (3). The advantage of the LCMV-CBS over the CEM-CBS is significant reduction of computational complexity. The only difference between these two approaches is that the weight vector, $\mathbf{v}_l$ generated by the LCMV-CBS has dimensionality of M as opposed to the dimensionality of MN produced by the CEM-CBS for its weight vector $\mathbf{w}_l$. Despite the fact that the LCMV-CBS and the CEM-CBS share the same design concept, both perform quite differently. According to our conducted experimental results, the CEM-CBS would only perform better than the LCMV-CBS when the number of selected bands is small such as fewer than ten bands, in which case the CEM-CBS is practically implementable and preferred to the LCMV-CBS. Otherwise, both performed very similarly in which case the LCMV-CBS is always preferable.

IV. Experiments

In this section, two real hyperspectral image data were used for experiments. Three different applications—target detection, mixed-pixel classification, and endmember extraction—were used to evaluate the CEM-CBS and LCMV-CBS in comparison with uniform band selection (UBS) and eigen-based band selection (EBS) [4], [5] for performance evaluation. Since the band selection considered in this paper is unsupervised, only the minimum variance PCA (MVPCA) in [4] and mutual information-based information divergence (ID) derived from [1]–[3] were used for comparison. In particular, the ID to be used in this paper is defined as follows.

Assume that the $\mathbf{p}_l$ is the image histogram of the lth band image, $\mathbf{B}_l$ normalized as a probability distribution, and $\mathbf{g}_l$ is its associated Gaussian distribution with mean and variance determined by sample mean and sample variance of the $\mathbf{B}_l$. The criterion of interest is to measure how much far away from a Gaussian distribution for a given band image, that is, the discrepancy between $\mathbf{p}_l$ and $\mathbf{g}_l$ defined by

$$\mathrm{D}(\mathbf{p}_l;\mathbf{g}_l) = \sum_i p_{li}\log\left(\frac{p_{li}}{g_{li}}\right) + \sum_i g_{li}\log\left(\frac{g_{li}}{p_{li}}\right) \quad (17)$$

which is called information divergence [13]. According to (17), the higher the value of $\mathrm{D}(\mathbf{p}_l;\mathbf{g}_l)$ in (17), the greater deviation of $\mathbf{p}_l$ from the Gaussian distribution, $\mathbf{g}_l$. This implies that the ID is used to measure non-Gaussianity of a band.

Additionally, the algorithms and images used for applications were selected to illustrate advantages and disadvantages of band selection in various applications. Other algorithms or images can be also used to evaluate our proposed CBS in these applications.

A. HYDICE Data

The first image data to be studied is an image scene shown in Fig. 1(a), which has a size of 64×64 pixel vectors with 15 panels in the scene and the ground truth map in Fig. 1(b). It was acquired by 210 spectral bands with a spectral coverage from 0.4 to 2.5 μm. Low signal/high noise bands (bands 1–3 and bands 202–210) and water vapor absorption bands (bands 101–112 and bands 137–153) were removed. So, a total of 169 bands were used in experiments. The spatial resolution is 1.56 m and spectral resolution is 10 nm.

Within the scene in Fig. 1(a), there is a large grass field background, and a forest on the left edge. Each element in this matrix is a square panel and denoted by p_{ij} with rows indexed by i and columns indexed by j =1, 2, 3. For each row i =1, 2, . . ., 5, there are three panels p_{i1}, p_{i2}, p_{i3}, painted by the same paint but with three different sizes. The sizes of the panels in the first, second and third columns are 3×3 m, 2×2 m, and 1×1 m, respectively. Since the size of the panels in the third column is 1×1 m, they cannot be seen visually from Fig. 1(a)

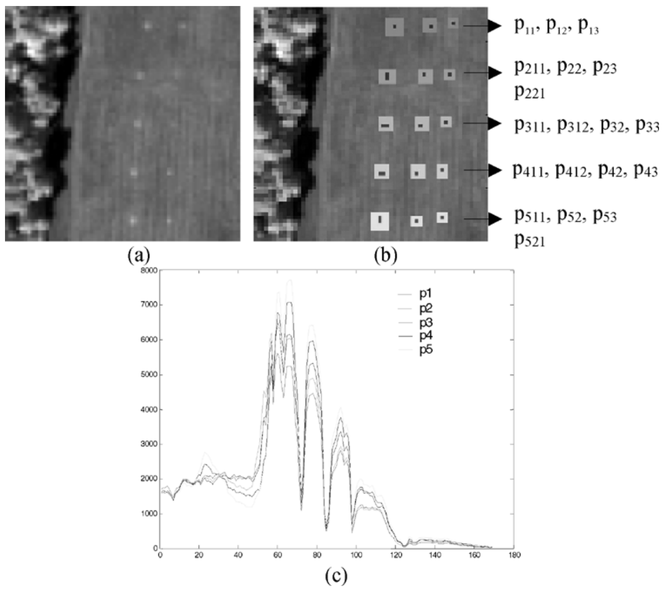


Fig. 1. (a) HYDICE panel scene which contains 15 panels. (b) Ground truth map of spatial locations of the 15 panels. (c) Spectral signatures of $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$, and $\mathbf{p}_5$.

TABLE I
VD ESTIMATES FOR THE HYDICE SCENE IN FIG. 1(a) BY THE HFC METHOD WITH VARIOUS FALSE ALARM PROBABILITIES

	$P_F = 10^{-1}$	$P_F = 10^{-2}$	$P_F = 10^{-3}$	$P_F = 10^{-4}$	$P_F = 10^{-5}$
VD	14	11	9	9	7

TABLE II
COMPARISON BETWEEN SELECTED BANDS USING DIFFERENT TECHNIQUES

BP criteria	9 selected bands
CEM-CBS BCM/BDM	18/85/112/10/109/74/115/114/96
CEM-CBS BCC/BDC	115/96/84/86/105/13/103/87/24
LCMV-CBS BCM/BDM	125/169/168/164/146/124/165/128/167
LCMV-CBS BCC/BDC	125/169/124/164/167/161/163/168/146
Uniform	1/19/37/55/73/91/109/127/145
MVPCA	60/61/67/66/65/59/57/68/62
ID	154/157/156/153/150/158/145/164/163

due to the fact that its size is less than the 1.56-m pixel resolution. For each column $j = 1, 2, 3$, the five panels have the same size but with five different paints. However, it should be noted that the panels in rows two and three were made by the same material with two different paints. Similarly, it is also the case for panels in rows four and five. Nevertheless, they were still considered as different panels, but our experiments will demonstrate that detecting panels in row five (row four) may also have effect on detection of panels in row two (row three). The 1.56-m spatial resolution of the image scene suggests that most of the 15 panels are one pixel in size, except the panels in the first column along with the second, third, fourth, and fifth rows, which are two-pixel panels, denoted by $\mathrm{p}_{211}, \mathrm{p}_{221}, \mathrm{p}_{311}, \mathrm{p}_{312}, \mathrm{p}_{411}, \mathrm{p}_{412}, \mathrm{p}_{511}, \mathrm{p}_{521}$. Since the size of the panels in the third column is 1×1 m, they cannot be seen visually from Fig. 1(a) due to the fact that its size is less than the 1.56-m pixel resolution. Fig. 1(b) shows the precise spatial locations of these 15 panels where red pixels (R pixels) are the panel center pixels and the pixels in yellow (Y pixels) are panel pixels mixed with the background. Fig. 1(c) plots the five panel spectral signatures $\mathbf{p}_i$ for $i = 1, 2, \ldots, 5$ obtained by averaging R pixels in the 3×3 m and 2×2 m panels in row i in Fig. 1(b). It should be noted the R pixels in the 1×1 m panels are not included because they are not pure pixels, mainly due to that fact that the spatial resolution of the R pixels in the 1×1 m panels is 1 m smaller than the pixel resolution is 1.56 m. These panel signatures along with the R pixels in the 3×3 m and 2×2 m panels were used as required prior target knowledge for the following comparative studies.

First, we need to determine how many bands needed for band selection. Let p denote the number of bands required for band selection. The VD was used to estimate the p where the method developed by Harsanyi *et al.* in [6] and [14], referred to as HFC method was used to calculate the VD for the HYDICE image scene in Fig. 1(a). The values of VD with various false alarm probabilities of $\mathrm{P_F}$ are tabulated in Table I.

For our experiments, VD was chosen to be 9. The selection of $p = 9$ is empirical based on the false alarm fixed at probabilities $\mathrm{P_F} = 10^{-3}, 10^{-4}$. As noted, the value of VD varies with the false alarm probability $\mathrm{P_F}$. This makes sense. With no availability of prior knowledge, the false alarm probability via the Neyman–Pearson detection theory [15] is probably one of most effective criteria to determine number of signals detected in the data. As interpreted in the introduction, one dimension can be only used to accommodate one single signal source. Therefore, it requires at least the VD-determined dimensions to separate distinct signal sources. So, as a matter of fact, the p should not be less than the VD.

1) Target Detection: In order to evaluate the impact of CBS on target detection, the commonly used CEM was implemented for performance analysis. The selection of the CEM was purely based on our preference which only uses partial target knowledge, desired target signature. It can be used to compare results for unsupervised mixed-pixel classification in the following section, Section IV-A2, where target knowledge is not provided *a priori*, rather obtained directly from the image scene. Table II tabulates the nine bands selected by the LCMV-CBS and CEM-CBS with nine highest band correlations along with the nine bands selected by uniform band and MVPCA in [4] and the ID defined by (17) where a backslash "/" is used to separate two selected bands. Since the criteria BCM and BDM yielded the identical bands for both the LCMV-CBS and the CEM-CBS, they are included in the same row in Table II. Similarly, it was true for BCC and BDC which also produced the identical bands for the LCMV-CBS and the CEM-CBS.

Fig. 2(a) shows the 15-panel detection results by the CEM using the complete set of full bands, 169 bands, whereas Fig. 2(b)–(h) shows detection results of the 15 panels by the CEM using nine bands selected by various criteria, BCM/BDM CEM-CBS, BCC/BDC CEM-CBS, BCM/BDM LCMV-CBS, BCC/BDC LCMV-CBS, uniform, MVPCA, and ID where the

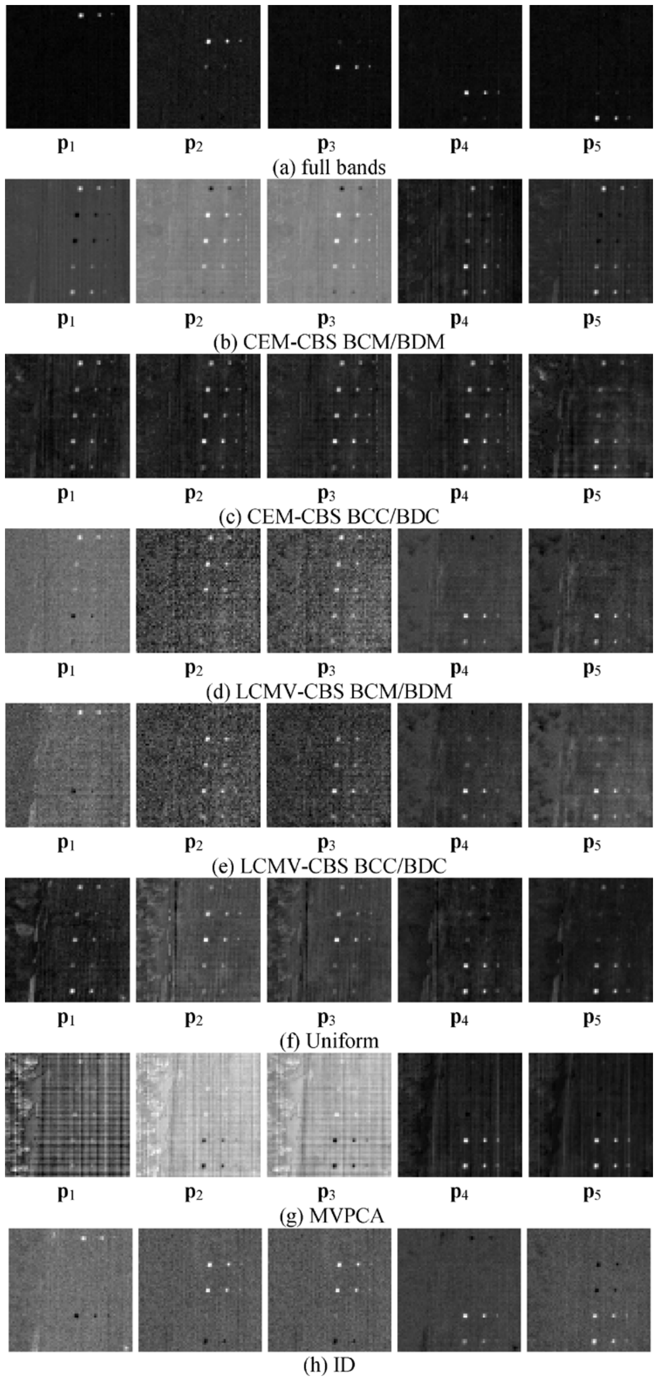


Fig. 2. Detection results of 15 panels by the CEM with nine selected bands.

desired target signature used in the CEM was selected from five panel signatures, $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$, and $\mathbf{p}_5$ in Fig. 1(c).

As demonstrated in Fig. 2(b)–(h), the ID seemed to perform better than other band-selection algorithms. However, this was due to the fact that the desired target signature used in the CEM was assumed to be known *a priori*. The results would be different and demonstrated in the following section where the unsupervised mixed-pixel classification is considered as an application. Furthermore, if we compare the results in Fig. 2(b)–(h) to that in Fig. 2(a), the CEM using full bands performed significantly better than band-selection-based techniques in Fig. 2(b)–(h). This may be due to three reasons. One is targets to be detected. Since the 15 panels are small targets, their detection requires as many as bands to provide sufficient sample spectral correlation to capture their subtle spectral properties. A second reason is the algorithm used for detection. The CEM only used one signature for detection while viewing other eight signatures as interfering signatures. As a result, it did not take advantage of the knowledge provided by these eight signatures. Instead, it used sample spectral correlation provided by the entire image to suppress these eight signatures. In this case, the sample correlation plays a key role in its detection. Like the first reason, such small panels can be only captured by subtle sample spectral correlation. Band selection may result in loss of some crucial information. A third reason is that the VD only provides an estimate of the least number of bands to be selected. It does not imply that the p must be the VD. In this particular experiment, $p = 9$ was not sufficiently large for the CEM to detect small panels. The results were significantly improved with all the 15 panels detected when the p was increased to be twice the VD, 18 or greater according to the results in [6, Ch. 17]. Their results are not included here since this experiment was not designed to show target detection, but rather demonstrated the performance in target detection for all the band-selection methods that were compared fairly under the same conditions. Nevertheless, as will be also demonstrated in the next section, if the application is mixed-pixel classification and the algorithm to be used is the fully constrained least-squares (FCLS) method developed in [6] and [16], the performance can be significantly improved using the same number of bands selected, $p = 9$.

2) Unsupervised Mixed-Pixel Classification: In the previous section, target detection was considered as an application for band selection, where the CEM was used to perform partially supervised target detection with only desired target signature required. In this section, unsupervised mixed-pixel classification is considered as another application for band selection. Since it must be performed unsupervisedly, an unsupervised algorithm which was proved to be effective, the automatic target generation process (ATGP) developed in [6], [17] was used to generate nine target pixels directly from the original image with full bands shown in Fig. 3(a) and nine target pixels from nine-band selected images shown in Fig. 3(b)–(h) to represent the required target knowledge for unsupervised classification with those target pixels corresponding to R panel pixels marked by green triangles. It should be noted that only nine target pixels were generated by the ATGP since a nine-band image has only nine dimensions that can be used for orthogonality projection.

An interesting finding from Fig. 3 is that the nine target pixels generated by the ATGP with full bands and nine selected bands by six different band-selection methods were quite different. Table III tabulates those ATGP-generated target pixels that were found to be actually R panel pixels in Fig. 1(b). It is interesting to note that only the LCMV-CBS produced four R panel pixels that were generated by the ATGP based on the nine selected bands and the LCMV-CBS with BCM/BDM was the only one finding an R panel pixel in the second column p_{22}. The MVPCA was the worst which could only find one panel pixel p_{521}.

For each band-selection method, the nine ATGP-generated target pixels were used to form the desired signature matrix for

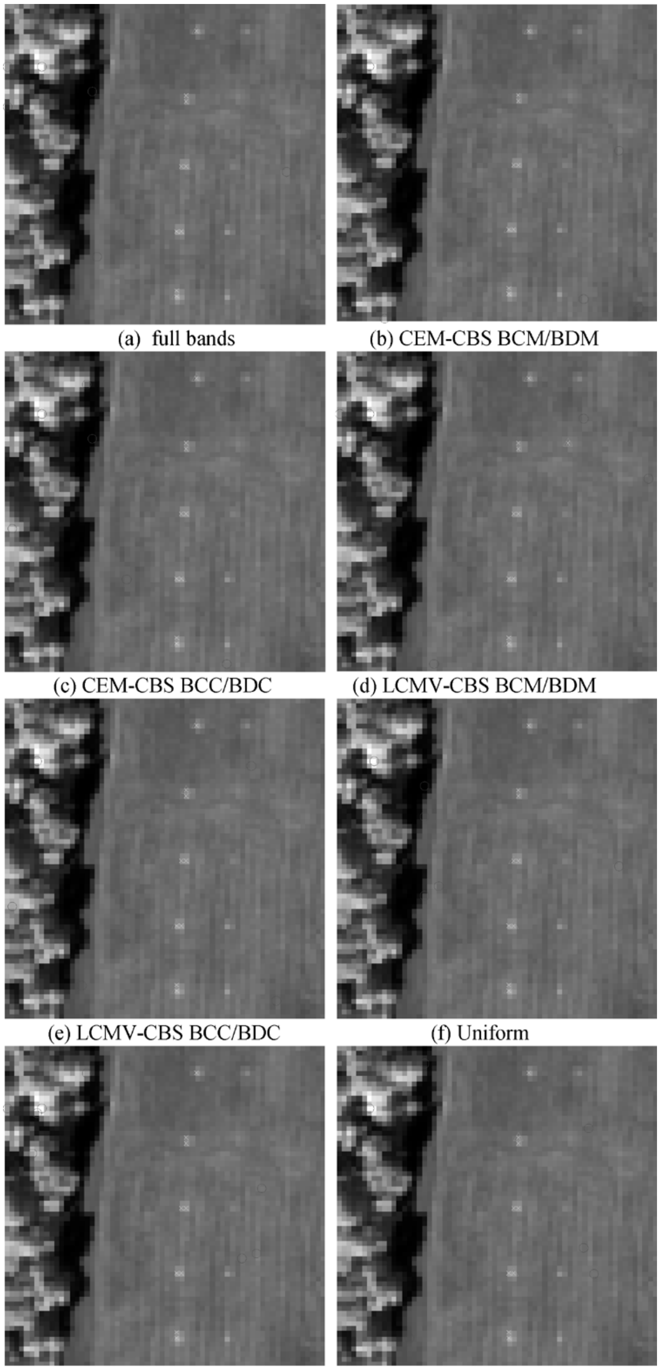


Fig. 3. Nine ATGP-generated target pixels.

the FCLS developed in [6] and [16] to classify the 15 panels. The FCLS used in the experiments was preferred to the orthogonal subspace projection (OSP) in [18] because the target signature knowledge provided by single pixels was sensitive to spectral variability unless fully abundance constraints imposed on the OSP which is the FCLS. Fig. 4(a)–(h) shows the FCLS-classification results of the 15 panels obtained by full bands and various band-selection methods, CEM-CBS, LCMV-CBS, uniform band selection, MVPCA, and ID, where the same bands selected in Table II by each band-selection method were used for the FCLS method. It is very obvious that the classification performance was determined by the R panel pixels found by the ATGP.

TABLE III
ATGP-Generated R Panel Pixels Using Full Bands and Six Different Band Selection Methods

	ATGP-generated R panel pixels
Full bands	P_{11}, P_{312}, P_{521}
CEM-CBS BCM/BDM	P_{11}, P_{411}
CEM-CBS BCC/BDC	P_{11}, P_{412}, P_{521}
LCMV-CBS BCM/BDM	P_{11}, P_{22}, P_{411}, P_{521}
LCMV-CBS BCC/BDC	P_{11}, P_{211}, P_{411}, P_{521}.
Uniform	P_{311}, P_{412}, P_{521}
MVPCA	P_{521}
ID	P_{211}, P_{412}, P_{521}

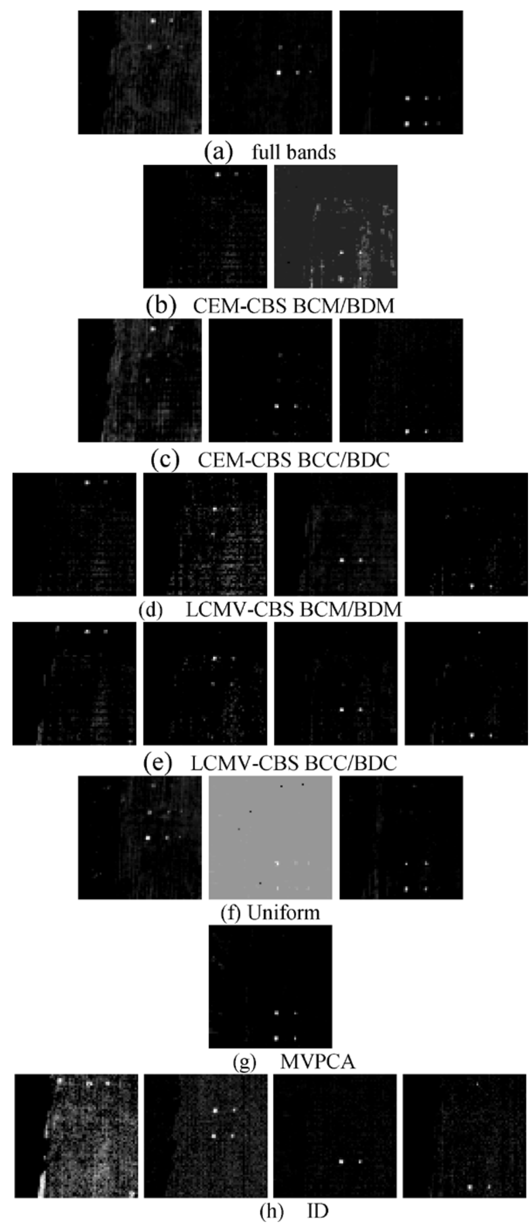


Fig. 4. Mixed-pixel classification with endmembers resulted from ATGP obtained by various band-selection methods using nine bands in Table III.

According to Table III, the LCMV-CBS found four R panel pixels among the nine ATGP-generated target pixels compared

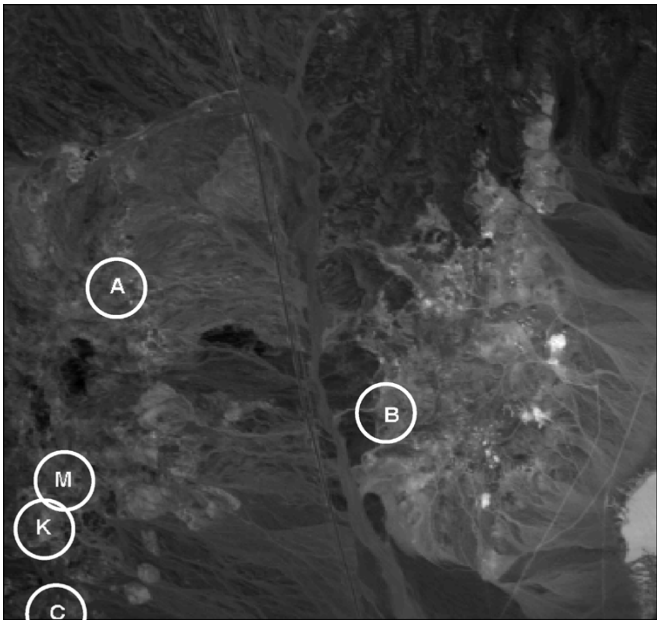


Fig. 5. Spatial positions of five pure pixels corresponding to minerals: (A) alunite, (B) buddingtonite, (C) calcite, (K) kaolinite, and (M) muscovite.

TABLE IV
VD ESTIMATES FOR THE AVIRIS SCENE IN FIG. 5 WITH VARIOUS FALSE ALARM PROBABILITIES

	$P_F = 10^{-1}$	$P_F = 10^{-2}$	$P_F = 10^{-3}$	$P_F = 10^{-4}$	$P_F = 10^{-5}$
VD	34	30	24	22	20

TABLE V
COMPARISON BETWEEN SELECTED BANDS USING DIFFERENT TECHNIQUES

criteria	22 selected bands
LCMV-CBS BCM/BDM	26/117/48/37/189/64/1/185/10/172/ 47/4/60/28/165/17/5/2/151/158/3/94
LCMV-CBS BCC/BDC	185/37/2/3/5/64/8/9/6/7/10/165/ 4/11/12/ 14/151/13/28/15/16/153
Uniform	1/9/17/25/33/41/49/57/65/73/81/89/97/ 105/113/121/129/137/145/153/161/169
MVPCA	87/85/88/86/89/84/91/80/78/90/92/83/82/79/ 93/81/ 98/99/97/189/77/76
ID	6/2/124/117/11/123/12/96/125/9/13/15/10/126/ 122/14/7 /16/5/165/112/1

to other methods which could only find three or fewer R panel pixels. Therefore, it is not surprising to see that the LCMV-CBS yielded the best performance and it even performed better than those produced by using full bands. Additionally, the FCLS-based mixed-pixel classification performance in Fig. 4 was improved significantly compared to Fig. 3 produced by the CEM which used the known panel signatures as the desired signatures. The experiments in Fig. 4 showed otherwise where the LCMV-CBS performed better than the ID. So, this experiment demonstrated that the same nine selected bands used for target detection in Fig. 3 produced significantly better performance for unsupervised mixed-pixel classification shown in Fig. 4. Most interestingly, according to Fig. 3, the ID performed better than the LCMV-CBS in the CEM detection. The above two experiments showed that with the same number of selected bands different applications may be very likely to result in different performance.

A remark on this example is worthwhile. The results in Fig. 4 were only used to evaluate performance of seven band-selection methods with fair comparison, in which case, the same ATGP and FCLS were implemented exactly the same way for full bands and the bands selected by each of band-selection methods.

B. AVIRIS Cuprite Data

Another real image is a well-known Airborne Visible/InfraRed imaging spectrometer (AVIRIS) Cuprite image scene shown in Fig. 5 which has been used to study endmember extraction extensively. It is available at website [19] and was collected by 224 spectral bands with 10-nm spectral resolution over the Cuprite mining site, Nevada, in 1997. The image in Fig. 5 has size of 350×350 pixels and is well understood mineralogically where bands 1–3, 105–115, and 150–170 have been removed prior to the analysis due to water absorption and low SNR in those bands. As a result, a total of 189 bands were used for experiments. The ground truth also provides the spatial locations of the five minerals—(A) alunite, (B) buddingtonite, (C) calcite, (K) kaolinite, and (M) muscovite—circled and labeled by A, B, C, K, and M, respectively, which can be used to verify endmembers extracted by an endmember extraction algorithm. Since the size of a band image is huge with 350×350 pixels, only the LCMV-CBS was implemented to select bands for endmember extraction to avoid the intensive computational complexity required by the CEM-CBS.

The VD estimated for this image scene was tabulated in Table IV with various false alarm probabilities P_F and tabulated in Table IV. For our experiments, $\mathrm{VD} = 22$ was chosen with $P_F = 10^{-4}$.

It has been shown in [20] that the value of the VD could be also used as the number of endmembers required to be generated. Therefore, in our experiments, both the number of bands to be selected and the number of endmembers to be extracted are set to 22. As noted previously, the VD only provides the least number of bands for band selection, and is not necessary to be exact the number of bands to be selected.

Table V tabulates 22 bands selected by the LCMV-CBS in accordance with 22 highest band correlations along with the uniformly selected 22 bands and the MVPCA-selected 22 bands where a backslash "/" is used to separate two selected bands. It should be noted that the value of the VD was only used as an estimate and not necessarily accurate which is almost impossible to know for real data.

For the purpose of endmember extraction, the widely used N-finder algorithm (N-FINDR) developed by Winter in [21] was used to extract 22 endmembers directly from the image scene. It should be noted that there is no particular reason to select the N-FINDR to perform endmember extraction. Any other endmember extraction algorithm such as pixel purity index (PPI) [22] can be also used for experiments as well.

Fig. 6(a) shows the 22 endmembers extracted by the N-FINDR using full bands with Fig. 6(b)–(f) showing the 22 endmembers extracted by the N-FINDR using 22 selected

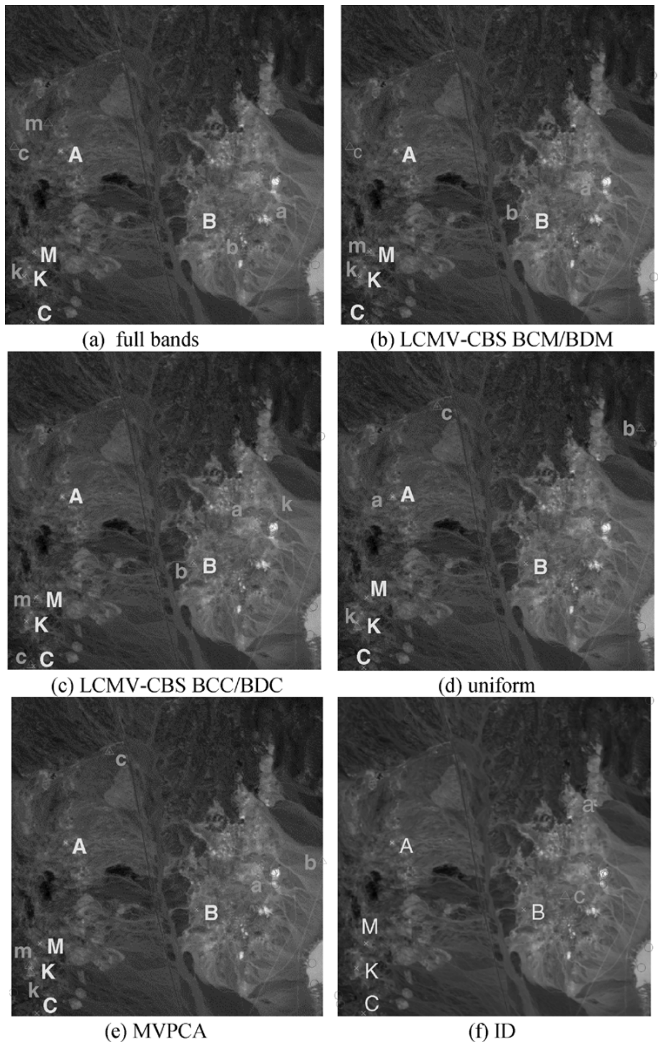


Fig. 6. Twenty-two endmembers extracted by N-FINDR using bands in Table V.

bands tabulated in Table V for comparison where the 22 N-FINDR-extracted endmember pixels are marked by red open circles with pixels marked by the lower cases of "a,b,c,k,m" with green triangles indicating that these pixels were found in correspondence to the five ground truth mineral endmembers marked by the upper cases of "A,B,C,K,M" with yellow crosses "x." It should be noted that the N-FINDR-found endmembers "a,b,c,k,m" in Fig. 6 were compared against the ground truth endmember pixels, "A,B,C,K,M" in Fig. 5 by the spectral angle mapper (SAM) [6] with their values tabulated in Table VI(A)–(F) where the coordinates included in the brackets for both "a,b,c,k,m" and "A,B,C,K,M" indicate the locations in the image scene.

As we can see from the above experiments, the LCMV-CBS performed very well to identify all the five mineral endmembers. Interestingly, the ID which produced good results for HYDICE data yielded the worst performance for Cuprite data in Fig. 6(f) and Table VI(F) where the N-FINDR could only find two mineral signatures, A and C. This result is the complete opposite of the result in CEM-based target detection in Fig. 3 where the ID produced the best result.

TABLE VI
(A) SPECTRAL SIMILARLY VALUES MEASURED BY SAM BETWEEN FOUND ENDMEMBERS AND THE GROUND TRUTH ENDMEMBERS FOR FULL BANDS. (B) SPECTRAL SIMILARLY VALUES MEASURED BY SAM BETWEEN FOUND ENDMEMBERS AND THE GROUND TRUTH ENDMEMBERS FOR LCMV-CBS BCM/BDM. (C) SPECTRAL SIMILARLY VALUES MEASURED BY SAM BETWEEN FOUND ENDMEMBERS AND THE GROUND TRUTH ENDMEMBERS FOR LCMV-CBS BCC/ BDC

	A (62,161)	B (209,234)	C (30,347)	K (22,298)	M (33,271)
a (270,273)	0.0575	0.1581	0.2045	0.0869	0.1381
b (265,267)	0.1608	0.0685	0.0803	0.1613	0.0884
c (11,156)	0.2115	0.0967	0.0350	0.2136	0.1170
k (23,298)	0.1098	0.1524	0.1918	0.0613	0.1079
m (48,130)	0.1786	0.0881	0.0681	0.1664	0.0643

(A)

	A (62,161)	B (209,234)	C (30,347)	K (22,298)	M (33,271)
a (273,188)	0.0567	0.1991	0.2427	0.0982	0.1742
b (205,228)	0.1247	0.0396	0.1123	0.1452	0.0866
c (11,156)	0.2115	0.0967	0.0350	0.2136	0.1167
k (22,299)	0.1076	0.1789	0.2177	0.0220	0.1305
m (33,272)	0.1519	0.0964	0.1040	0.1413	0.0264

(B)

	A (62,161)	B (209,234)	C (30,347)	K (22,298)	M (33,271)
a (267,164)	0.0493	0.1404	0.1813	0.0870	0.1230
b (208,235)	0.1459	0.0281	0.1110	0.1637	0.0968
c (27,346)	0.2208	0.1106	0.0466	0.2327	0.1330
k (296,165)	0.0727	0.1385	0.1844	0.0647	0.1124
m (34,274)	0.1681	0.0859	0.0898	0.1594	0.0441

(C)

According to Table VI, it is worth noting that none of ground truth mineral pixels were found by the N-FINDR even though full bands were used. Nevertheless, the spectral similarity values produced by the SAM in Table VI should provide information about closeness between the N-FINDR found mineral pixel and the ground truth mineral pixels. Using the values in Table VI as a measure of effectiveness, the LCMV-CBS with BCM/BDM in Table VI(B) was the best among all band-selection methods and LCMV-CBS was also the best on the average and even better than the one produced by the full bands.

Three concluding remarks are noteworthy.

1) On some occasions, an appropriate band selection may be more effective than one with full bands as demonstrated by the endmember extraction experiments. This could be due to the fact that spectral information from real data may be distorted by unknown signal sources and

TABLE VI (*Continued.*)
(D) SPECTRAL SIMILARLY VALUES MEASURED BY SAM BETWEEN FOUND ENDMEMBERS AND THE GROUND TRUTH ENDMEMBERS FOR UNIFORM BAND SELECTION. (E) SPECTRAL SIMILARLY VALUES MEASURED BY SAM BETWEEN FOUND ENDMEMBERS AND THE GROUND TRUTH ENDMEMBERS FOR MVPCA. (F) SPECTRAL SIMILAR VALUES MEASURED BY SAM BETWEEN FOUND ENDMEMBERS AND THE GROUND TRUTH ENDMEMBERS FOR ID

	A (62,161)	B (209,234)	C (30,347)	K (22,298)	M (33,271)
a (61,161)	0.0172	0.1645	0.2115	0.0962	0.1476
b (336,86)	0.1525	0.0711	0.0792	0.1752	0.0978
c (111,62)	0.1783	0.1002	0.0725	0.1938	0.0974
k (23,303)	0.0984	0.1671	0.2100	0.0304	0.1276

(D)

	A (62,161)	B (209,234)	C (30,347)	K (22,298)	M (33,271)
a (278,191)	0.0582	0.1956	0.2385	0.1101	0.1765
b (350,181)	0.1344	0.0597	0.0817	0.1476	0.0692
c (111,61)	0.1873	0.0934	0.0550	0.196	0.0944
k (23,303)	0.0984	0.1671	0.2100	0.0304	0.1276
m (21,295)	0.1282	0.1036	0.1355	0.0956	0.0656

(E)

	A (62,161)	B (209,234)	C (30,347)	K (22,298)	M (33,271)
a (263,119)	0.0348	0.1564	0.2007	0.0953	0.1392
c (255,219)	0.1657	0.0729	0.0567	0.1823	0.0933

(F)

too much such contaminated information may eventually hinder data analysis. As a consequence, a judicious selection of appropriate bands can avoid such dilemma and further improve performance.

2) Moreover, the applications presented in this section only serve as an illustrative purpose to demonstrate the utility of the CBS. The performance of CBS is also determined by the algorithms used for applications. Therefore, when a comparative analysis was conducted, the same algorithm was applied to various band-selection methods to make sure that all conditions were held to the same situation so that the only difference is selected bands used for analysis.
3) It should be noted that, in order to make a fair comparison among different band-selection criteria, the same algorithm must be implemented in conjunction with the VD which determines the same number of bands to be selected for all the criteria. Only the same algorithm coupled with the same VD-determined number of bands can ensure that the difference in performance only comes from a specific band-selection criterion. In this section, we demonstrated effectiveness of band-selection criteria with the number of bands determined by the VD in three different applications where three algorithms, the CEM for target detection, the UFCLS for unsupervised mixed-pixel classification, and the N-FINDR for endmember extraction were used for performance evaluation. According to our conducted experiments, the CBS outperformed other band-selection criteria in all the case.
4) Our experimental results showed that the CBS performed robustly and uniformly well compared to the other band-selection methods. Although only three applications, target detection, mixed-pixel classification and endmember extraction were considered in this paper, some other applications, such as mixed-pixel quantification, data compression were also investigated for the CBS. Similar conclusions can be also drawn. So, in order to avoid replication, their results are not included.

V. CONCLUSION

The CEM/LCMV has enjoyed success in hyperspectral target detection and classification. Its applications to hyperspectral data exploitation has been yet to be explored. This paper presents another new application of the CEM/LCMV in band selection. Seemingly, it is difficult to make a connection between the CEM/LCMV and band selection. However, if we interpret the desired target signature used in the CEM/LCMV as a particular band to be selected, the CEM/LCMV that was originally minimizes interfering effects resulting from all other signal sources is equivalent to the CEM/LCMV-CBS that minimizes band correlation/dependence resulting from all other bands. With this interpretation it is natural to expand the ability of the CEM/LCMV in target detection and classification to that in band selection. This paper presents two versions of such an approach to band selection, CEM-based CBS and LCMV-based CBS for band selection. In order to resolve the issue in determination of the number of bands required to be selected, p, a recently developed concept of VD is implemented in conjunction with the proposed CBS to estimate the p. However, it should be noted that the VD only provides a reasonable guideline for a band-selection method in determining a least number of bands required to be retained but does not necessarily imply that the number of bands to be selected must be equal to the VD. Experiments demonstrate that the CBS is indeed a promising, robust and effective band-selection technique. We believe that our proposed CBS method will become a very useful band-selection technique once people realize its potential in hyperspectral data exploitation. As a final concluding remark, we would like to point out that the VD is completely determined by the image data to be processed and is independent of applications. Therefore, the VD is a versatile technique and can be used for various applications [23] and the band selection is only one of its applications.

REFERENCES

[1] P. W. Mausel, W. J. Kramber, and J. K. Lee, "Optimal band selection for supervised classification of multispectral data," *Photogramm. Eng. Remote Sens.*, vol. 56, no. 1, pp. 55–60, Jan. 1990.
[2] C. Conese and F. Maselli, "Selection of optimum bands from TM scenes thorugh mutual information analysis," *ISPRS Photogramm. Eng. Remote Sens.*, vol. 48, no. 3, pp. 2–11, Jan. 1993.

[3] S. D. Stearns, B. E. Wilson, and J. R. Peterson, “Dimesnionality reduction by optimal band selection for pixel classification of hyperspectral imagery,” in *Proc. Applications of Digital Image Processing XVI*, vol. 2028, 1993, pp. 118–127.

[4] C.-I Chang, Q. Du, T. S. 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.

[5] R. Huang and M. He, “Band selection based feature weighting for classification of hyperspectral data,” *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 2, pp. 156–159, Apr. 2005.

[6] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. New York: Plenum, 2003.

[7] B. Ramakrishna, J. Wang, A. Plaza, and C.-I Chang, “Spectral/spatial hyperspectral image compression in conjunction with virtual dimensionality,” presented at the SPIE Symp. Defense and Security Conf. Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI, vol. 5806, Orlando, FL, Mar./Apr. 28–1, 2005.

[8] J. Wang and C.-I Chang, “Dimensionality reduction by independent component analysis for hyperspectral image analysis,” presented at the IEEE Int. Geoscience Remote Sensing Symp., Seoul, Korea, Jul. 25–29, 2005.

[9] C.-I Chang and A. Plaza, “Fast iterative algorithm for implementation of pixel purity index,” *IEEE Geosci. Remote Sens. Lett.*, vol. 3, pp. 63–67, Jan. 2006.

[10] O. L. Frost III, “An algorithm for linearly constrained adaptive array processing,” *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.

[11] J. C. Harsanyi, *Detection and Classification of Subpixel Spectral Signatures in Hyperspectral Image Sequences*. Baltimore, MD: Dept. Elect. Eng., Univ. Maryland Baltimore County, 1993.

[12] C.-I Chang, “Target signature-constrained mixed pixel classification for hyperspectral imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 5, pp. 1065–1081, May 2002.

[13] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[14] J. C. Harsanyi, W. Farrand, and C.-I Chang, “Detection of subpixel spectral signatures in hyperspectral image sequences,” in *Proc. Amer. Soc. Photogrammetry Remote Sensing Annu. Meeting*, Reno, NV, 1994, pp. 236–247.

[15] H. V. Poor, *An Introduction to Detection and Estimation Theory*, 2nd ed. New York: Springer-Verlag, 1994.

[16] D. Heinz and C.-I Chang, “Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery,” *IEEE Trans. Geoscience Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

[17] H. Ren and C.-I Chang, “Automatic spectral target recognition in hyperspectral imagery,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1232–1249, Oct. 2003.

[18] J. Harsanyi and C.-I Chang, “Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach,” *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779–785, Jul. 1994.

[19] [Online]. Available: http://speclab.cr.usgs.gov/cuprite.html

[20] J. Wang and C.-I Chang, “Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1586–1600, Jun 2006.

[21] M. E. Winter, “N-finder: An algorithm for fast autonomous spectral endmember determination in hyperspectral data,” in *Image Spectr. V*, vol. 3753, 1999, pp. 266–277.

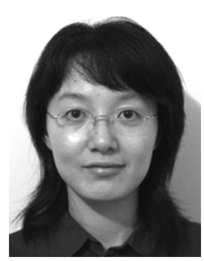
[22] J. W. Boardman, F. A. Kruse, and R. O. Green, “Mapping target signatures via partial unmixing of AVIRIS data,” presented at the Summaries JPL Airborne Earth Science Workshop, Pasadena, CA, 1995.

[23] C.-I Chang, “Utility of virtual dimensionality in hyperspectral data exploitation,” presented at the SPIE Defense and Security Symp., Algorithms and Techniques for Multspectral, Hyperspectral, and Ultraspectral Imagery XII (OR52), Orlando, FL, Apr. 17–21, 2006.

Chein-I Chang (S’81–M’87–SM’92) received the B.S. degree from Soochow University, Taipei, Taiwan, R.O.C., in 1973, the M.S. degree from the Institute of Mathematics, National Tsing Hua University, Hsinchu, Taiwan, in 1975, and the M.A. degree from the State University of New York, Stony Brook, in 1977, all in mathematics. He received the M.S. and M.S.E.E. degrees from the University of Illinois at Urbana-Champaign, in 1982 and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 1987.

He is currently with Department of Electrical Engineering as a Distinguished Lecture Chair, National Chung Hsing University, Taichung, Taiwan. He has been with the University of Maryland Baltimore County (UMBC), Baltimore, since 1987, as a Visiting Assistant Professor from January 1987 to August 1987, Assistant Professor from 1987 to 1993, Associate Professor from 1993 to 2001, and Professor in the Department of Computer Science and Electrical Engineering since 2001. He was a Visiting Research Specialist in the Institute of Information Engineering at the National Cheng Kung University, Tainan, Taiwan, from 1994 to 1995. He has three patents on automatic pattern recognition and several pending patents on image processing techniques for hyperspectral imaging and detection of microcalcifications. His research interests include multispectral/hyperspectral image processing, automatic target recognition, medical imaging, information theory and coding, signal detection and estimation, and neural networks. He is the author of a book *Hyperspectral Imaging: Techniques for Spectral Detection and Classification* (Kluwer, 2003). He is on the editorial board and was the Guest Editor of a special issue on telemedicine and applications of the *Journal of High Speed Networks*.

Dr. Chang received a National Research Council Senior Research Associateship Award from 2002 to 2003 at the U.S. Army Soldier and Biological Chemical Command, Edgewood Chemical and Biological Center, Aberdeen Proving Ground, MD. He is an Associate Editor in the area of hyperspectral signal processing for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He is a Fellow of SPIE and a member of Phi Kappa Phi and Eta Kappa Nu.



Su Wang (S’04) received the B.S. and M.S. degrees from the Nanjing University of Aeronautics and Astronautics, Nanjing, China. She is currently pursuing the Ph.D. degree in electrical engineering at the University of Maryland Baltimore County.