Constrained-Target Band Selection for Multiple-Target Detection

Yulei Wang, Lin Wang, Chunyan Yu, *Member, IEEE*, Enyu Zhao, Meiping Song, Chia-Hsien Wen, and Chein-I Chang, *Life Fellow, IEEE*

***Abstract*— This paper develops a new approach to band selection for multiple-target detection, called constrained-target band selection (CTBS). Its idea is derived from the concept of constrained energy minimization (CEM) by constraining a target of interest, while minimizing the variance resulting from the background (BKG). By taking advantage of CEM, the variance produced by a target of interest can be further used as a measure of prioritizing bands as well as a means of selecting bands for this particular target. As a result, two CTBS-based band prioritization (BP) criteria, called minimal variance-based BP (MinV-BP) and maximal variance-based BP (MaxV-BP), and two CTBS-based BS methods, called sequential forward CTBS (SF-CTBS) and sequential backward CTBS (SB-CTBS), can be derived for multiple-target detection. Since the bands selected by CTBS vary with targets of interest used to constrain CEM, in order for CTBS to be applied to multiple targets, a new fusion technique, called band fusion selection (BFS), is further developed for CTBS to integrate bands selected by different targets so that CTBS can work for all targets. Unlike most BS methods for target detection which generally simultaneously select a fixed set of bands for all targets of interest, the ideas of constraining multiple-target detection and using BFS are novelty of this paper. Experimental results show that CTBS performs well for multiple-target detection.**

***Index Terms*— Band fusion selection (BFS), band prioritization (BP), band selection (BS), constrained energy minimization (CEM), constrained-target band selection (CTBS), virtual dimensionality (VD).**

Manuscript received February 13, 2018; revised May 6, 2018, October 6, 2018, and February 9, 2019; accepted March 1, 2019. Date of publication April 11, 2019; date of current version July 22, 2019. The work of Y. Wang was supported in part by the National Nature Science Foundation of China under Grant 61801075 and in part by the Fundamental Research Funds for the Central Universities under Grant 3132016331. The work of L. Wang was supported by the Fundamental Research Funds for Central Universities under Grant JB150508. The work of C. Yu was supported by the National Nature Science Foundation of Liaoning Province under Grant 20170540095. The work of E. Zhao was supported by the National Nature Science Foundation of China under Grant 41801231. The work of M. Song was supported in part by the National Nature Science Foundation of China under Grant 61601077 and in part by the Fundamental Research Funds for the Central Universities under Grant 3132017124. The work of C.-I Chang was supported by the Fundamental Research Funds for Central Universities under Grant 3132016331 and Grant ZD20180073. *(Corresponding author: Lin Wang.)*

Y. Wang is with the Information and Technology College, Dalian Maritime University, Dalian 116026, China, and also with the State Key Laboratory of Integrated Services Networks, Xi'an 710000, China (e-mail: wangyulei.wyl@163.com).

L. Wang is with the School of Physics and Optoelectronic Engineering, Xidian University, Xi'an 710126, China (e-mail: lwang@mail.xidian.edu.cn).

C. Yu and E. Zhao are with the Center of Hyperspectral Imaging in Remote Sensing, Information and Technology College, Dalian Maritime University, Dalian 116026, China (e-mail: yuchunyan1997@126.com; 421009432@qq.com).

M. Song is with the Center of Hyperspectral Imaging in Remote Sensing, Information and Technology College, Dalian Maritime University, Dalian 116026, China, and also with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710216, China (e-mail: smping@163.com).

C.-H. Wen is with the Department of Computer Science and Information Management, Providence University, Taichung 02912, Taiwan (e-mail: chwen@pu.edu.tw).

C.-I Chang is with the Center of Hyperspectral Imaging in Remote Sensing, Information and Technology College, Dalian Maritime University, Dalian 116026, China, with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, MD 21250 USA, and also with the Department of Computer Science and Information Management, Providence University, Taichung 02912, Taiwan (e-mail: cchang@umbc.edu).

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/TGRS.2019.2904264

I. Introduction

BAND selection (BS) has received considerable interests in recent years [1]–[18] and will continue to be a major area in the future hyperspectral data exploration. Technically speaking, BS can be performed in either a supervised or an unsupervised fashion. Regarding supervised BS, it generally requires certain prior knowledge such as training samples, specific applications. As for unsupervised BS, there is no *a priori* knowledge available to be used for BS. In this case, BS has no direct link to any specific application and must rely on data characteristics and statistics [2], [3]. On the other hand, BS can also be implemented by band prioritization (BP) which selects band according to priorities ranked by a BP criterion [2], [3] or by a searching strategy which designs algorithms to search for bands either one after another sequentially [8] or band groups/clusters/subsets simultaneously [9]–[13].

This paper develops a rather new specific application-based supervised BS approach, to be called constrained-target band selection (CTBS) for multiple-target detection where only the knowledge of targets of interest is assumed to be known *a priori*, and no other prior knowledge such as background (BKG) information is required. Due to significant refined spectral solution, hyperspectral imaging sensors can uncover many subtle and unknown signal sources which are generally embedded in BKG and cannot be identified by *a priori* knowledge or visual inspection. Accordingly, how to deal with BKG issue is very challenging and crucial to hyperspectral data exploitation. This is clearly witnessed by hyperspectral image classification where the BKG issue is generally not included in evaluation of classification performance. Interestingly, to deal with the BKG issue for classification a recent work in [19] develops a one-class oriented approach which used cluster space representation to model BKG as to achieve better classification. Such one-class classification shares a similar

idea with target detection where the class of interest to be classified is considered as one hypothesis corresponding to the desired signal to be detected versus BKG as an uninteresting signal corresponding to another hypothesis to represent noise. Compared to the one-class classification in [19] a well-known subpixel target detector, called constrained energy minimization (CEM) developed in [20]–[23], makes use of sample correlation matrix $\mathbf{R}$ to suppress BKG rather than modeling BKG which requires tuning several parameters that have significant impact on classification. As a result, CEM is much easily implemented than one-class classification since it only needs the knowledge of the target of interest to be detected and nothing else.

As a matter of fact, the development of CEM arose from this need because it only requires the knowledge of the target signature to be detected as the desired target signature $\mathbf{d}$, while suppressing all other signal sources, specifically, including BKG signatures. Its idea is to design a target detector specified by $\mathbf{d}$ using $\mathbf{R}^{-1}$, which is the inverse of the global sample correlation matrix $\mathbf{R} = (1/N)\sum_{i=1}^{N} \mathbf{r}_i \mathbf{r}_i^T$ formed by all data sample vectors, $\{\mathbf{r}_i\}_{i=1}^{N}$, to suppress unknown BKG by minimizing the variance $(\mathbf{d}^T\mathbf{R}^{-1}\mathbf{d})^{-1}$. It is believed that CEM is one of a very few hyperspectral techniques that can deal with the BKG issue effectively via $\mathbf{R}^{-1}$. Besides, since BKG has so many unknown signal sources, each of which may respond to various wavelengths, it is highly desirable to select those bands that only correspond to targets of interest but have no response to signal sources of no interest. This suggests that CEM can be made more effective if not all full bands are needed to process CEM. With this in mind, CEM was previously used to design a new BS method, called constrained band selection (CBS) in [5] which constrains a selected band as a band of interest, $\mathbf{b}_l$, by replacing $\mathbf{d}$ with $\mathbf{b}_l$, while minimizing the variance effect resulting from all other bands as bands of no interest by replacing $\mathbf{R} = (1/N)\sum_{i=1}^{N} \mathbf{r}_i \mathbf{r}_i^T$ with $\tilde{\mathbf{Q}}_{\Omega-\{\mathbf{b}_l\}} = (1/L-1)\sum_{i=1,i\neq l}^{L} \mathbf{b}_i \mathbf{b}_i^T$ where L is the total number of bands and $\mathbf{b}_l$ is the lth band vector representing the lth band image $\mathbf{B}_l$. The resultant minimal variance derived from CBS is given by $(\mathbf{b}_l^T \tilde{\mathbf{Q}}_{\Omega-\{\mathbf{b}_l\}}\mathbf{b}_l)^{-1}$ which is used to calculate priority scores to rank bands. So, basically, CBS can be considered as a BP-based BS technique.

The proposed CTBS takes another twist from CBS to develop a new CEM-based BS approach for a specific application in target detection. Instead of following the idea of CBS which constrains a band of interest, CTBS constrains a target of interest, $\mathbf{d}$. Due to their use of different constraints, $\mathbf{b}_l$ for CBS and $\mathbf{d}$ for CTBS, CBS and CTBS are derived from two completely different rationales.

More specifically, assume that $\mathbf{d}$ is a target signature of interest used to detect targets specified by $\mathbf{d}$. For each single band $\mathbf{b}_l$, we can calculate $\mathrm{V}(\mathbf{b}_l) = (\mathbf{d}_{\mathbf{b}_l}^T \mathbf{R}_{\mathbf{b}_l}^{-1}\mathbf{d}_{\mathbf{b}_l})^{-1}$ which can be used as the priority score assigned to the lth band, $\mathbf{b}_l$ to rank the priority of $\mathbf{b}_l$. As a result, all the bands can be prioritized according to increasing values of $\mathrm{V}(\mathbf{b}_l)$. The smaller the $\mathrm{V}(\mathbf{b}_l)$, the more significant is the $\mathbf{b}_l$ band. The criterion using increasing values of $\mathrm{V}(\mathbf{b}_l)$ as BP to rank bands is called minimal variance-based BP (MinV-BP).

In contrast to MinV-BP, for each single band $\mathbf{b}_l$, we can also remove $\mathbf{b}_l$ from the full band set, Ω by replacing $\mathrm{V}(\mathbf{b}_l)$ with $\mathrm{V}(\Omega - \mathbf{b}_l) = (\mathbf{d}_{\Omega-\{\mathbf{b}_l\}}^T \mathbf{R}_{\Omega-\{\mathbf{b}_l\}}^{-1}\mathbf{d}_{\Omega-\{\mathbf{b}_l\}})^{-1}$. In this case, the value of $\mathrm{V}(\Omega - \mathbf{b}_l)$ can also be used as the priority score assigned to the lth band, $\mathbf{b}_l$ to rank the priority of $\mathbf{b}_l$. As a consequence, all the bands can be prioritized according to decreasing values of $\mathrm{V}(\Omega - \mathbf{b}_l)$. In other words, as an opposite to $\mathrm{V}(\mathbf{b}_l)$, the larger the $\mathrm{V}(\Omega - \mathbf{b}_l)$, the more significant is the $\mathbf{b}_l$ band. The criterion using decreasing values of $\mathrm{V}(\Omega - \mathbf{b}_l)$ as BP to rank bands is called maximal variance-based BP (MaxV-BP).

Interestingly, the two CTBS-based BP criteria defined above can be further used as band searching criteria to develop their own corresponding BS methods as follows.

Suppose that Ω_{BS} is a set of bands already being selected and $\{\mathbf{r}_{\Omega_{\mathrm{BS}}}^i\}_{i=1}^{N}$ is the data set where each of data samples, $\mathbf{r}_{\Omega_{\mathrm{BS}}}^i$ is the ith data sample vector using only bands in Ω_{BS}. Then, the sample correlation matrix using only selected bands in Ω_{BS} is formed by $\mathbf{R}_{\Omega_{\mathrm{BS}}} = (1/N)\sum_{i=1}^{N} \mathbf{r}_{\Omega_{\mathrm{BS}}}^i (\mathbf{r}_{\Omega_{\mathrm{BS}}}^i)^T$. According to MinV-BP, we can replace the single band-based $\mathrm{V}(\mathbf{b}_l)$ with the band subset Ω_{BS}-based $\mathrm{V}(\Omega_{\mathrm{BS}}) = (\mathbf{d}_{\Omega_{\mathrm{BS}}}^T \mathbf{R}_{\Omega_{\mathrm{BS}}}^{-1}\mathbf{d}_{\Omega_{\mathrm{BS}}})^{-1}$. When the band subset Ω_{BS} is augmented in a feed forward manner by one band at a time sequentially, the resulting BS method is called sequential forward CTBS (SF-CTBS). Similarly, we can also replace $\mathrm{V}(\Omega - \mathbf{b}_l)$ with $\mathrm{V}(\Omega - \Omega_{\mathrm{BS}}) = (\mathbf{d}_{\Omega-\Omega_{\mathrm{BS}}}^T \mathbf{R}_{\Omega-\Omega_{\mathrm{BS}}}^{-1}\mathbf{d}_{\Omega-\Omega_{\mathrm{BS}}})^{-1}$ and augment selected bands by one at a time sequentially in a backward manner to derive its corresponding backward version, called sequential backward CTBS (SB-CTBS).

In summary, there are two types of processes to implement CTBS. One is to use CTBS as a BP criterion to select bands one at a time sequentially. This gives rise to two versions of BP-based CTBS, MinV-BP and MaxV-BP. The other is to use CTBS as a searching strategy to select bands for BS in a feed forward manner and a backward manner sequentially, which also result in two versions, SF-CTBS and SB-CTBS.

Finally, it should be particularly noted that the design rationale of CTBS is quite different from that proposed in CBS [5]. Basically, CTBS deals with both target signatures and bands of interest all together as an entity compared to CBS which only focuses on bands of interest with nothing else. This section concludes with an interesting comment. Despite that there are recent reports on target detection-based BS, either using sparse-based CEM (SCEM) [16] or using maximum–submaximum ratio (MSR) criterion or correlation coefficient (CC) as criteria with PSO used as a search strategy [18] to find bands. They are all quite different from the work presented in this paper in many ways. First of all, we do not use data sparsity as SCEM or MSR/CC as criteria to perform BS. Second, we use virtual dimensionality (VD) [22] to determine the number of bands, while the threshold t used for the sparse model in [16] and [18] and the number of particles P in [18] were empirically chosen. Third, the search algorithms used in this paper are completely different from PSO in [18]. Finally, our BS is designed for each of multiple targets and then fuse bands selected from each target, whereas SCEM in [16] is designed to select bands for

single target and target-constrained interference-minimization filter (TCIMF) in [24] is designed to select bands for multiple targets at the same time. These are crucial elements in this paper which cannot be found in the existing literature.

II. CEM

Since CEM plays a key role in development of CTBS, this section briefly reviews its concept as follows. It is derived from linearly constrained minimum variance (LCMV) originally proposed by Frost for adaptive beamforming [25].

Suppose that a hyperspectral image is represented by a collection of image pixel vectors, denoted by $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N\}$, where $\mathbf{r}_i = (r_{i1}, r_{i2}, \ldots, r_{iL})^T$ for $1 \leq i \leq N$ is an L-dimensional pixel vector, N is the total number of pixels in the image, and L is the total number of spectral bands. Furthermore, assume that $\mathbf{d} = (d_1, d_2, \ldots, d_L)^T$ is specified by a desired signature of interest to be used for target detection. It should be noted that $\mathbf{d}$ does not have to be a data sample vector and can be a spectral signature chosen from a database or spectral library. The goal is to find a target detector that can detect data samples specified by the desired target signature $\mathbf{d}$ via a finite impulse response (FIR) linear filter with L filter coefficients, $\{w_1, w_2, \ldots, w_L\}$, denoted by an L-dimensional vector $\mathbf{w} = (w_1, w_2, \ldots, w_L)^T$ which minimizes the filter output energy subject to the constraint $\mathbf{d}^T\mathbf{w} = \mathbf{w}^T\mathbf{d} = 1$. Let y_i denote the output of the designed FIR filter resulting from the input $\mathbf{r}_i$. Then, y_i can be expressed by

$$y_i = \sum_{l=1}^{L} w_l r_{il} = (\mathbf{w})^T \mathbf{r}_i = \mathbf{r}_i^T \mathbf{w} \tag{1}$$

and the average energy of the filter output is given by

$$\begin{aligned}(1/N)\sum_{i=1}^{N} y_i^2 &= (1/N)\sum_{i=1}^{N} \left(\mathbf{r}_i^T \mathbf{w}\right)^2 \\ &= \mathbf{w}^T \left[(1/N)\sum_{i=1}^{N} \mathbf{r}_i \mathbf{r}_i^T\right] \mathbf{w} = \mathbf{w}^T \mathbf{R} \mathbf{w}\end{aligned} \tag{2}$$

where $\mathbf{R} = (1/N)\sum_{i=1}^{N} \mathbf{r}_i \mathbf{r}_i^T$ is the sample autocorrelation matrix of the image. The goal is to solve the following linearly constrained optimization problem:

$$\min_{\mathbf{w}} \{\mathbf{w}^T \mathbf{R} \mathbf{w}\} \quad \text{s.t. } \mathbf{d}^T\mathbf{w} = \mathbf{w}^T\mathbf{d} = 1 \tag{3}$$

where $\mathbf{w}^T\mathbf{R}\mathbf{w}$ can be considered as either the variance resulting from signals not passing through the filter. The optimal solution to (3) is shown in [20]–[23] to be

$$\mathbf{w}^{\mathrm{CEM}} = (\mathbf{d}^T\mathbf{R}^{-1}\mathbf{d})^{-1}\mathbf{R}^{-1}\mathbf{d} \tag{4}$$

and

$$\min_{w} \mathbf{w}\mathbf{R}^{-1}\mathbf{w} = (\mathbf{w}^{\mathrm{CEM}})^T\mathbf{R}^{-1}\mathbf{w}^{\mathrm{CEM}} = (\mathbf{d}^T\mathbf{R}^{-1}\mathbf{d})^{-1} \tag{5}$$

which is the minimal variance resulting from unwanted signal sources impinging upon an array of sensors [25], referred to as CEM error. With the optimal weight, $\mathbf{w}^{\mathrm{CEM}}$ specified by (4) a filter called CEM, denoted by $\delta^{\mathrm{CEM}}(\mathbf{r})$ was derived [20] as a target detector specified by

$$\delta^{\mathrm{CEM}}(\mathbf{r}) = (\mathbf{w}^{\mathrm{CEM}})^T\mathbf{r} = (\mathbf{d}^T\mathbf{R}^{-1}\mathbf{d})^{-1}(\mathbf{R}^{-1}\mathbf{d})^T\mathbf{r}. \tag{6}$$

III. CEM-Based CBS

Before deriving CTBS, it is worth describing how CBS was developed by CEM. Let $\Omega = \{\mathbf{B}_l\}_{l=1}^{L}$ 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 $K \times M$. Each band image $\mathbf{B}_l$ can be represented by a column vector of dimension KM, denoted by $\mathbf{b}_l$ and $\mathbf{B}_\Omega = [\mathbf{b}_1\mathbf{b}_2 \ldots \mathbf{b}_L]$ is the band matrix formed by all band image vectors, $\{\mathbf{b}_l\}_{l=1}^{L}$. Let $\mathbf{w}_l$ be an KM-dimensional column vector that is used to specify a FIR filter designed for the band image vector $\mathbf{b}_l$ and y_l be the filter output specified by

$$y_l = \mathbf{w}_l^T\mathbf{b}_l. \tag{7}$$

The averaged least-squares filter output is given by

$$(1/L)\sum_{l=1}^{L} y_l^2 = (1/L)\sum_{l=1}^{L} \left(\mathbf{w}_l^T\mathbf{b}_l\right)\left(\mathbf{w}_l^T\mathbf{b}_l\right)^T = \mathbf{w}_l^T\mathbf{Q}\mathbf{w}_l \tag{8}$$

where $\mathbf{Q} = (1/L)\sum_{l=1}^{L}\mathbf{b}_l\mathbf{b}_l^T = (1/L)\mathbf{B}_\Omega\mathbf{B}_\Omega^T$ denote the band image correlation matrix. Since CEM generally excludes the band image $\mathbf{b}_l$ from the band image correlation matrix $\mathbf{Q}$, we can further define $\tilde{\mathbf{Q}} = (1(L-1))\sum_{j=1, j\neq l}^{L}\mathbf{b}_j\mathbf{b}_j^T$ as the band image dependence matrix. Replacing $\mathbf{Q}$ in (8) with $\tilde{\mathbf{Q}}$ results in a CBS problem

$$\min_{\mathbf{w}_l} \left\{\mathbf{w}_l^T\tilde{\mathbf{Q}}\mathbf{w}_l\right\} \quad \text{s.t. } \mathbf{b}_l^T\mathbf{w}_l = 1. \tag{9}$$

The solution to (9), $\tilde{\mathbf{w}}_l^{\mathrm{CBS}}$ is given by

$$\tilde{\mathbf{w}}_l^{\mathrm{CBS}} = \left(\mathbf{b}_l^T\tilde{\mathbf{Q}}^{-1}\mathbf{b}_l\right)^{-1}\tilde{\mathbf{Q}}^{-1}\ \mathbf{b}_l. \tag{10}$$

According to (10), the LSE resulting from $\tilde{\mathbf{w}}_l^{\mathrm{CBS}}$ is

$$\begin{aligned}\tilde{\rho}_l &= \left(\tilde{\mathbf{w}}_l^{\mathrm{CBS}}\right)^T\tilde{\mathbf{Q}}\tilde{\mathbf{w}}_l^{\mathrm{CBS}} \\ &= \left(\left(\mathbf{b}_l^T\tilde{\mathbf{Q}}^{-1}\mathbf{b}_l\right)^{-1}\tilde{\mathbf{Q}}^{-1}\mathbf{b}_l\right)^T\tilde{\mathbf{Q}}\left(\mathbf{b}_l^T\tilde{\mathbf{Q}}^{-1}\ \mathbf{b}_l\right)^{-1}\tilde{\mathbf{Q}}^{-1}\mathbf{b}_l \\ &= \left(\mathbf{b}_l^T\tilde{\mathbf{Q}}^{-1}\mathbf{b}_l\right)^{-2}\mathbf{b}_l^T\tilde{\mathbf{Q}}^{-1}\mathbf{b}_l = \left(\mathbf{b}_l^T\tilde{\mathbf{Q}}^{-1}\mathbf{b}_l\right)^{-1}\end{aligned} \tag{11}$$

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 less significant the band. Therefore, $\{\tilde{\rho}_l\}_{l=1}^{L}$ can be used as priority scores to rank all the bands in $\Omega = \{\mathbf{B}_l\}_{l=1}^{L}$.

IV. CTBS

Assume that $\{\mathbf{b}_l\}_{l=1}^{L}$ is a set of band images representing a hyperspectral image cube where $\mathbf{b}_l$ is the lth spectral band represented by a column vector, $\mathbf{b}_l = (b_{l1}, b_{l2}, \ldots, b_{lN})^T$ and $\{b_{li}\}_{i=1}^{N}$ is the set of all N pixels in the lth band image, $\mathbf{b}_l$.

By taking advantage of the CEM error derived from (5), we can define the following new measure that can be used a criterion for CTBS

$$\mathrm{V}(\Omega_{\mathrm{BS}}) = \left(\mathbf{d}_{\Omega_{\mathrm{BS}}}^T\mathbf{R}_{\Omega_{\mathrm{BS}}}^{-1}\mathbf{d}_{\Omega_{\mathrm{BS}}}\right)^{-1} \tag{12}$$

which is the minimal variance specified by (2) using only those bands in $\Omega_{\rm BS}$. Most importantly, we can prove by the following theorem:

$$\{\mathrm{V}(\Omega_j)\}_{j=1}^{L} \tag{13}$$

is in fact a monotonically decreasing sequence where $\Omega_j = \{\mathbf{b}_{l_1}, \mathbf{b}_{l_2}, \ldots, \mathbf{b}_{l_j}\}$ is any j-band subset containing j bands, $\mathbf{b}_{l_1}, \mathbf{b}_{l_2}, \ldots, \mathbf{b}_{l_j}$.

Despite that V($\Omega_{\rm BS}$) in (12) and $(\mathbf{b}_l^T \tilde{\mathbf{Q}}^{-1} \mathbf{b}_l)^{-1}$ have similar forms with $\mathbf{d}_{\Omega_{\rm BS}}$ and $\mathbf{R}_{\Omega_{\rm BS}}$ in (12) corresponding to $\mathbf{b}_l$ and $\tilde{\mathbf{Q}}$ in (11), respectively, they are quite different in design rationale. In particular, the former is specified by target signature $\mathbf{d}$ and sample correlation matrix as opposed to the latter specified by a desired band $\mathbf{b}_l$ and the band image correlation matrix $\tilde{\mathbf{Q}}$. The following theorem (Theorem 1) is crucial to the development of CTBS. There is no such corresponding theorem as a counterpart in CBS [5].

Theorem 1: Let $\Omega = \{\mathbf{b}_l\}_{l=1}^{L}$ be a full band set. For any band subset $\Omega_k = \{\mathbf{b}_{l_j}\}_{j=1}^{k} \subset \Omega$ and an arbitrary band image vector $\mathbf{b}_{k+1} \notin \Omega_k$

$$\frac{1}{\mathbf{d}_{\Omega_k \cup \{\mathbf{b}_{k+1}\}}^T \mathbf{R}_{\Omega_k \cup \{\mathbf{b}_{k+1}\}}^{-1} \mathbf{d}_{\Omega_k \cup \{\mathbf{b}_{k+1}\}}} < \frac{1}{\mathbf{d}_{\Omega_k}^T \mathbf{R}_{\Omega_k}^{-1} \mathbf{d}_{\Omega_k}}. \tag{14}$$

Proof: (See Appendix).

From (14), we can also show that

$$\frac{1}{\mathbf{d}_{\Omega_1}^T \mathbf{R}_{\Omega_1}^{-1} \mathbf{d}_{\Omega_1}} < \frac{1}{\mathbf{d}_{\Omega_2}^T \mathbf{R}_{\Omega_2}^{-1} \mathbf{d}_{\Omega_2}} \quad \text{if } \Omega_2 \subset \Omega_1 \text{ and } \Omega_2 \neq \Omega_1. \tag{15}$$

Let $\Omega_1 = \Omega_2 \cup \{\mathbf{b}_j\}_{j=1}^{p}$. Equation (15) can be proved by implementing (14) p times repeatedly as follows:

$$\begin{aligned} &\frac{1}{\mathbf{d}_{\Omega_1}^T \mathbf{R}_{\Omega_1}^{-1} \mathbf{d}_{\Omega_1}} \\ &< \frac{1}{\mathbf{d}_{\Omega_2 \cup \{\mathbf{b}_1, \ldots, \mathbf{b}_{p-1}\}}^T \mathbf{R}_{\Omega_2 \cup \{\mathbf{b}_1, \ldots, \mathbf{b}_{p-1}\}}^{-1} \mathbf{d}_{\Omega_2 \cup \{\mathbf{b}_1, \ldots, \mathbf{b}_{p-1}\}}} \\ &< \cdots < \frac{1}{\mathbf{d}_{\Omega_2 \cup \{\mathbf{b}_1\}}^T \mathbf{R}_{\Omega_2 \cup \{\mathbf{b}_1\}}^{-1} \mathbf{d}_{\Omega_2 \cup \{\mathbf{b}_1\}}} < \frac{1}{\mathbf{d}_{\Omega_2}^T \mathbf{R}_{\Omega_2}^{-1} \mathbf{d}_{\Omega_2}}. \end{aligned} \tag{16}$$

Furthermore, (14) also allows us to derive the following inequalities:

$$\frac{1}{\mathbf{d}_{\Omega_1 \cup \Omega_2}^T \mathbf{R}_{\Omega_1 \cup \Omega_2}^{-1} \mathbf{d}_{\Omega_1 \cup \Omega_2}} < \frac{1}{\mathbf{d}_{\Omega_1}^T \mathbf{R}_{\Omega_1}^{-1} \mathbf{d}_{\Omega_1}}$$

and

$$\frac{1}{\mathbf{d}_{\Omega_1 \cup \Omega_2}^T \mathbf{R}_{\Omega_1 \cup \Omega_2}^{-1} \mathbf{d}_{\Omega_1 \cup \Omega_2}} < \frac{1}{\mathbf{d}_{\Omega_2}^T \mathbf{R}_{\Omega_2}^{-1} \mathbf{d}_{\Omega_2}}. \tag{17}$$

Of particular interest is the following inequality derived in [26]:

$$\frac{1}{\mathbf{d}_{\Omega}^T \mathbf{R}_{\Omega}^{-1} \mathbf{d}_{\Omega}} < \frac{1}{\mathbf{d}_{\tilde{\Omega}}^T \mathbf{R}_{\tilde{\Omega}}^{-1} \mathbf{d}_{\tilde{\Omega}}} \tag{18}$$

which can be also shown to be a special case of (15) with Ω_1 replaced by the full band set Ω and Ω_2 replaced by $\tilde{\Omega}$ which is a proper subset of Ω.

Finally, by virtue of (14), it shows that $\{\mathrm{V}(\Omega_j)\}_{j=1}^{L}$ is indeed a monotonically decreasing sequence.

V. Band Prioritization by CTBS

Since the sequence $\{\mathrm{V}(\Omega_j)\}_{j=1}^{L}$ is monotonically decreasing according to (14), we can develop two types of BP criteria, single band MinV-BP and leave-one-out MaxV-BP.

A. Single Band Minimum Variance Band Prioritization by CTBS

According to (5), $(\mathbf{d}^T \mathbf{R}^{-1} \mathbf{d})^{-1}$ is the variance resulting from CEM which can be considered as the interfering effect caused by signal sources including BKG signal sources which do not pass through the CEM filter (3). This quantity is obtained by full band set Ω, i.e., $\mathrm{V}(\Omega) = (\mathbf{d}_{\Omega}^T \mathbf{R}_{\Omega}^{-1} \mathbf{d}_{\Omega})^{-1} = (\mathbf{d}^T \mathbf{R}^{-1} \mathbf{d})^{-1}$. On the other hand, from (12), $\mathrm{V}(\Omega_{\rm BS}) = (\mathbf{d}_{\mathbf{b}_l}^T \mathbf{R}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l})^{-1}$ is calculated by using a single band $\mathbf{b}_l$ with $\Omega_{\rm BS} = \{\mathbf{b}_l\}$. So, for each single band image $\mathbf{b}_l$, we calculate the variance

$$\mathrm{V}(\mathbf{b}_l) = \left(\mathbf{d}_{\mathbf{b}_l}^T \mathbf{R}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l}\right)^{-1}. \tag{19}$$

which can be used as a criterion to measure the variance caused by data sample vectors not specified by $\mathbf{d}$ using only one single band, $\mathbf{b}_l$. Consequently, the value of $\mathrm{V}(\mathbf{b}_l)$ can be further used as a priority score of $\mathbf{b}_l$. With this interpretation, a single band-based BP (SB-BP) criterion can be derived to rank bands according to increasing values of $\mathrm{V}(\mathbf{b}_l)$ (19). That is, the smaller the $\mathrm{V}(\mathbf{b}_l)$, the higher priority is the $l^{\rm th}$ band, $\mathbf{b}_l$. Since it ranks bands by starting the minimal variance in (19) as its first selected band, it is called MinV-BP. A detailed description of implementing MinV-BP is provided in the following.

MinV-BP

1. Initial condition:
   Let $n_{\rm BS}$ be the number of bands needed to be selected, which can be determined by a criterion such as virtual dimensionality (VD) in [23, Chapter 17].
   Find the first band by
   $$\mathbf{b}_{l_1} = \arg\left\{\min_{\mathbf{b}_l \in \Omega} \mathrm{V}(\mathbf{b}_l)\right\} \tag{20}$$
2. Rank all bands for $1 \le l \le L$ according to $\{\mathrm{V}(\mathbf{b}_l)\}_{l=1}^{L}$ by (19)
   $$\mathbf{b}_{l_j} \succ \mathbf{b}_{l_k} \Leftrightarrow \mathrm{V}(\mathbf{b}_{l_j}) < \mathrm{V}(\mathbf{b}_{l_k}) \tag{21}$$
   i.e., $\mathbf{b}_{l_1} \succ \mathbf{b}_{l_2} \succ \cdots \succ \mathbf{b}_{l_L}$ where the notation "$\succ$" is used to indicate "superior to."
3. The final set of selected bands is given by the first $n_{\rm BS}$ bands by $\{\mathbf{b}_{l_1}, \mathbf{b}_{l_2}, \ldots, \mathbf{b}_{l_{n_{BS}}}\}$.

It is worth noting that $\mathbf{R}_{\mathbf{b}_l}$ is a scalar and so is $\mathbf{R}_{\mathbf{b}_l}^{-1}$. Interestingly, according to [27] $\mathbf{R}_{\mathbf{b}_l}$ should remove all data sample vectors specified by $\mathbf{d}_{\mathbf{b}_l}$ because the functionality of $\mathbf{R}_{\mathbf{b}_l}^{-1}$ is to suppress the interfering effect caused by all data sample vectors other than those specified by $\mathbf{d}_{\mathbf{b}_l}$. So, let Δ be such data set excluding from data sample vectors specified by $\mathbf{d}_{\mathbf{b}_l}$. We can define $\tilde{\mathbf{R}}_{\mathbf{b}_l} = \sum_{r_{\mathbf{b}_l} \in \Delta} r_{\mathbf{b}_l}^2$ and $(s_l^{\rm BKG})^2$ as the

BKG energy given by $(s_l^{\mathrm{BKG}})^2 = \tilde{\mathbf{R}}_{\mathbf{b}_l}$. In this case, $\mathbf{d}_{\mathbf{b}_l}^T \tilde{\mathbf{R}}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l}$ becomes the signal-to-background (SBR) ratio, i.e., $\mathrm{SBR}_l = \mathbf{d}_{\mathbf{b}_l}^T \tilde{\mathbf{R}}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l} = (d_{\mathbf{b}_l}^2 / ((s_l^{\mathrm{BKG}})^2))$. Consequently, the larger the SBR_l, the smaller is the $(\mathbf{d}_{\mathbf{b}_l}^T \tilde{\mathbf{R}}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l})^{-1}$. Since the size of the data sample vectors specified by $\mathbf{d}_{\mathbf{b}_l}$ is generally small, $\tilde{\mathbf{R}}_{\mathbf{b}_l} \approx \mathbf{R}_{\mathbf{b}_l}$. In this case, $\mathrm{V}(\mathbf{b}_l) = (\mathbf{d}_{\mathbf{b}_l}^T \mathbf{R}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l})^{-1}$ in (19) can also be represented by $\mathrm{MinV}(\mathbf{b}_l) = (\mathbf{d}_{\mathbf{b}_l}^T \tilde{\mathbf{R}}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l})^{-1} = (\mathrm{SBR}_l)^{-1}$ which is the reciprocal of SBR_l.

B. Leave-One-Out Maximum Variance Band Prioritization by CTBS

As an alternative to MinV-BP, we can also replace the full band set Ω in (5) with $\Omega_{\mathrm{BS}} = \Omega - \mathbf{b}_l$ in (12) to derive

$$\mathrm{V}(\Omega - \mathbf{b}_l) = \left(\mathbf{d}_{\Omega-\{\mathbf{b}_l\}}^T \mathbf{R}_{\Omega-\{\mathbf{b}_l\}}^{-1} \mathbf{d}_{\Omega-\{\mathbf{b}_l\}}\right)^{-1}. \tag{22}$$

Like (19) which is $(\mathrm{SBR}_l)^{-1}$ resulting from a single band $\mathbf{b}_l$, (22) can also be considered as $(\mathrm{SBR}_{\Omega-\{\mathbf{b}_l\}})^{-1}$ resulting from $(L-1)$ bands, $\Omega - \{\mathbf{b}_l\}$. Its idea is to first remove any single band, $\mathbf{b}_l$, from the full band set Ω and then to calculate variance by (22). The band which yields the maximal variance should be most significant one since its removal from Ω causes the maximal variance. With this in mind, $\mathrm{V}(\Omega - \mathbf{b}_l)$ in (22) can be used as a priority score to rank $\mathbf{b}_l$ for BP. That is, the higher the $\mathrm{V}(\Omega - \mathbf{b}_l)$, the more significant is the band $\mathbf{b}_l$. As a result, all bands can be ranked by decreasing values of $\mathrm{V}(\Omega - \mathbf{b}_l)$ as opposed to MinV-BP which ranks all bands by increasing values of $\mathrm{V}(\mathbf{b}_l)$ in (19). Because of that, it is called MaxV-BP criterion. A detailed step-by-step implementation of MaxV-BP is given as follows.

MaxV-BP

1. Initial condition:
   Let n_{BS} be the number of bands needed to be selected, which can be determined by a criterion such as virtual VD.
   $$\mathbf{b}_{l_1} = \arg\{\max_{\mathbf{b}_l \in \Omega} \mathrm{V}(\Omega - \mathbf{b}_l)\} \tag{23}$$
2. Rank all bands $1 \le l \le L$ according to $\{\mathrm{V}(\Omega - \mathbf{b}_l)\}_{l=1}^L$ by (22)
   $$\mathbf{b}_{l_j} \succ \mathbf{b}_{l_k} \Leftrightarrow \mathrm{V}(\Omega - \mathbf{b}_{l_k}) < \mathrm{V}(\Omega - \mathbf{b}_{l_j}) \tag{24}$$
   i.e., $\mathbf{b}_{l_1} \succ \mathbf{b}_{l_2} \succ \cdots \succ \mathbf{b}_{l_L}$ where the notation "$\succ$" is used to indicate "superior to."
3. The final set of selected bands is given by the first n_{BS} bands by $\{\mathbf{b}_{l_1}, \mathbf{b}_{l_2}, \ldots, \mathbf{b}_{l_{n_{BS}}}\}$.

VI. Band Selection by CTBS

In this section, the two CTBS-based BP criteria designed in Section V are further used as search strategies to derive their corresponding band selection algorithms, respectively, described in Sections VI-A and VI-B.

A. Sequential Forward-CTBS

For each single band $\mathbf{b}_l$, we can replace the full band set Ω used in the CEM-produced variance $(\mathbf{d}^T \mathbf{R}^{-1} \mathbf{d})^{-1} = (\mathbf{d}_\Omega^T \mathbf{R}_\Omega^{-1} \mathbf{d}_\Omega)^{-1}$ with the single band $\mathbf{b}_l$ to become $(\mathbf{d}_{\mathbf{b}_l}^T \mathbf{R}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l})^{-1}$ which can be used as the priority score assigned to the lth band, $\mathbf{b}_l$ to rank the priority of $\mathbf{b}_l$. As a result, the band, which yields the minimal variance

$$\mathbf{b}_{l_1}^* = \arg\left\{\min_{\mathbf{b}_l \in \Omega} \mathrm{V}(\mathbf{b}_l)\right\} = \arg\left\{\min_{\mathbf{b}_l \in \Omega} \left(\mathbf{d}_{\mathbf{b}_l}^T \mathbf{R}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l}\right)^{-1}\right\} \tag{25}$$

will be selected as the first band, denoted by $\mathbf{b}_{l_1}^*$. Then, the second band which yields the minimal variance again

$$\begin{aligned} \mathbf{b}_{l_2}^* &= \arg\left\{\min_{\mathbf{b}_l \in \Omega - \{\mathbf{b}_{l_1}\}} \mathrm{V}(\mathbf{b}_l)\right\} \\ &= \arg\left\{\min_{\mathbf{b}_l \in \Omega - \{\mathbf{b}_{l_1}\}} \left(\mathbf{d}_{\mathbf{b}_l}^T \mathbf{R}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l}\right)^{-1}\right\} \end{aligned} \tag{26}$$

is then selected as the second band, $\mathbf{b}_{l_2}^*$. The same process is repeatedly over and over again by continuously adding one new band at a time via repeatedly implementing (26). The resulting algorithm is called SF-CTBS.

Now, if we consider a band as a feature vector, SFFS developed in [28] can be used to develop an SFFS-based band selection to augment bands to be selected one at a time sequentially by (19) instead of ranking all bands by their priority scores $\{\mathrm{V}(\mathbf{b}_l)\}_{l=1}^L$ as MinV-BP does. The resulting BS is referred to as SF-CTBS.

SF-CTBS

1. Initial condition:
   Let n_{BS} be the number of bands needed to be selected, which can be determined by a criterion such as virtual VD.
   Find
   $$\begin{aligned} \mathbf{b}_{l_1} &= \arg\left\{\min_{\mathbf{b}_l \in \Omega} \left(\mathbf{d}_{\mathbf{b}_l}^T \mathbf{R}_{\mathbf{b}_l}^{-1} \mathbf{d}_{\mathbf{b}_l}\right)^{-1}\right\} \\ \Omega_1 &= \{\mathbf{b}_{l_1}\}. \end{aligned} \tag{27}$$
2. Band augmentation
   $$\mathbf{b}_{l_j} = \arg\left\{\min_{\mathbf{b}_l \in \Omega_{j-1}^c} \left(\mathbf{d}_{\Omega_{j-1} \cup \{\mathbf{b}_l\}}^T \mathbf{R}_{\Omega_{j-1} \cup \{\mathbf{b}_l\}}^{-1} \mathbf{d}_{\Omega_{j-1} \cup \{\mathbf{b}_l\}}\right)^{-1}\right\} \tag{28}$$
   where $\Omega_{j-1} = \{\mathbf{b}_{l_1}, \mathbf{b}_{l_2}, \ldots, \mathbf{b}_{l_{j-1}}\}$ and $\Omega_{j-1}^c = \Omega - \Omega_{j-1}$.
3. If $j < n_{\mathrm{BS}}$,
   $$\Omega_j = \{\mathbf{b}_{l_1}, \mathbf{b}_{l_2}, \ldots, \mathbf{b}_{l_j}\} = \Omega_{j-1} \cup \{\mathbf{b}_{l_j}\} \tag{29}$$
   and go step 2. Otherwise, BS is terminated. The final set of selected bands is given by the first selectedn_{BS} bands, i.e., $\{\mathbf{b}_{l_1}, \mathbf{b}_{l_2}, \ldots, \mathbf{b}_{l_{n_{BS}}}\}$.

It should be pointed out that SF-CTBS does not have to run through all bands as MinV-BP does. It can terminate

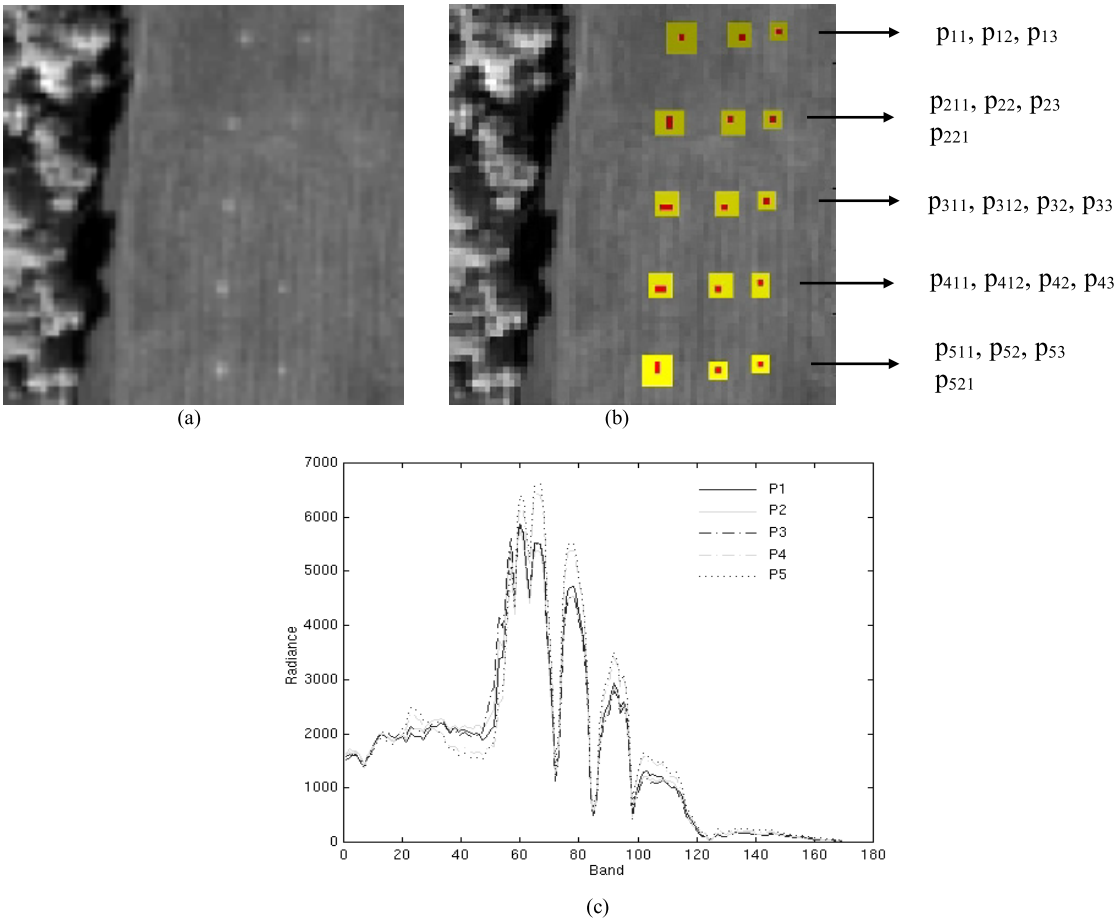


Fig. 1. (a) HYDICE panel scene which contains 15 panels. (b) Ground truth map of spatial locations of 19 R panel pixels. (c) Spectra of $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$ and $\mathbf{p}_5$.

augmentation process as long as the number of bands to selected, n_{BS} is reached.

B. Sequential Backward-CTBS

In contrast to SF-CTBS we can also develop a sequential backward search (SBS), which uses leave-one-out as a technique to select optimal feature vectors. The resulting BS is called SB-CTBS.

More specifically, for each single band, say lth band, $\mathbf{b}_l$, we consider the band subset $\Omega - \mathbf{b}_l$ by removing $\mathbf{b}_l$ from the full band set Ω and then replace the full band set Ω used in the CEM-produced variance $(\mathbf{d}^T\mathbf{R}^{-1}\mathbf{d})^{-1} = (\mathbf{d}_\Omega^T\mathbf{R}_\Omega^{-1}\mathbf{d}_\Omega)^{-1}$ with the single band $\Omega - \mathbf{b}_l$ to become $\mathrm{V}(\Omega - \mathbf{b}_l) = (\mathbf{d}_{\Omega-\{\mathbf{b}_l\}}^T\mathbf{R}_{\Omega-\{\mathbf{b}_l\}}^{-1}\mathbf{d}_{\Omega-\{\mathbf{b}_l\}})^{-1}$ which can be used as the priority score assigned to the lth band, $\mathbf{b}_l$ to rank the priority of $\mathbf{b}_l$. In particular, the band which yields the maximal variance

$$\begin{aligned}\mathbf{b}_{l_1}^* &= \arg\Big\{\max_{\mathbf{b}_l\in\Omega}\mathrm{V}(\Omega-\{\mathbf{b}_l\})\Big\}\\ &= \arg\Big\{\max_{\mathbf{b}_l\in\Omega}\left(\mathbf{d}_{\Omega-\{\mathbf{b}_l\}}^T\mathbf{R}_{\Omega-\{\mathbf{b}_l\}}^{-1}\mathbf{d}_{\Omega-\{\mathbf{b}_l\}}\right)^{-1}\Big\}\end{aligned} \tag{30}$$

will be selected as the first band, denoted by $\mathbf{b}_{l_1}^*$ to be the most significant band since the variance in (30) produced the maximal variance if $\mathbf{b}_{l_1}^*$ is removed from Ω. Now let $\Omega_1 = \Omega - \{\mathbf{b}_{l_1}\}$. Then, the second band which yields the maximal variance again

$$\begin{aligned}\mathbf{b}_{l_2}^* &= \arg\Big\{\min_{\Omega_1-\{\mathbf{b}_l\}}\mathrm{V}(\Omega_1-\mathbf{b}_l)\Big\}\\ &= \arg\Big\{\min_{\Omega_1-\{\mathbf{b}_l\}}\left(\mathbf{d}_{\Omega_1-\{\mathbf{b}_l\}}^T\mathbf{R}_{\Omega_1-\{\mathbf{b}_l\}}^{-1}\mathbf{d}_{\Omega_1-\{\mathbf{b}_l\}}\right)^{-1}\Big\}\end{aligned} \tag{31}$$

is then selected as the second band, $\mathbf{b}_{l_2}^*$. The same process is repeatedly over and over again by continuously removing selected bands one at a time from the full band set Ω. The resulting algorithm is called SB-CTBS.

It should be noted that SB-CTBS is different from SF-CTBS in two different aspects. First of all, SB-CTBS removes bands from unselected band sets to augment a desired selected band subset according to (33) compared to SF-CTBS which augments selected bands one at a time using (28). Second, SB-CTBS uses the sample correlation matrix formed by data sample vectors using all bands except the bands already selected, $\mathbf{R}_{\Omega-\Omega_j}$, while SF-CTBS only uses the selected bands in Ω_j to form the sample correlation matrix $\mathbf{R}_{\Omega j}$. Nevertheless, both SF-CTBS and SB-CTBS do not have to run through all bands as MaxV-BP does. Their augmentation processes can be terminated as long as the number of bands to selected, n_{BS} is reached.

SB-CTBS

1. Initial condition:
   Let n_{BS} be the number of bands needed to be selected, which can be determined by a criterion such as virtual VD.
   Find
   $$\mathbf{b}_{l_1} = \arg\left\{\max_{\mathbf{b}_l\in\Omega}\left(\mathbf{d}^T_{\Omega-\{\mathbf{b}_l\}}\mathbf{R}^{-1}_{\Omega-\{\mathbf{b}_l\}}\mathbf{d}_{\Omega-\{\mathbf{b}_l\}}\right)^{-1}\right\}$$
   $$\Omega_1 = \{l_1\} \tag{32}$$
2. Band Reduction
   $$\mathbf{b}_{l_j} = \arg\left\{\max_{\mathbf{b}_l\in\Omega^c_{j-1}}\left(\mathbf{d}^T_{\Omega-(\Omega_{j-1}\cup\{\mathbf{b}_l\})}\mathbf{R}^{-1}_{\Omega-(\Omega_{j-1}\cup\{\mathbf{b}_l\})}\times\mathbf{d}_{\Omega-(\Omega_{j-1}\cup\{\mathbf{b}_l\})}\right)^{-1}\right\} \tag{33}$$
   where $\Omega_{j-1} = \{\mathbf{b}_{l_1}, \mathbf{b}_{l_2}, \ldots, \mathbf{b}_{l_{j-1}}\}$ and $\Omega^c_{j-1} = \Omega - \Omega_{j-1}$.
3. If $j < n_{\mathrm{BS}}$,
   $$\Omega_j = \{\mathbf{b}_{l_1}, \mathbf{b}_{l_2}, \ldots, \mathbf{b}_{l_j}\} = \Omega_{j-1}\cup\{\mathbf{b}_{l_j}\} \tag{29}$$
   and go step 2. Otherwise, BS is terminated. The final set of selected bands is given by the first selected n_{BS} bands, i.e., $\{\mathbf{b}_{l_1}, \mathbf{b}_{l_2}, \ldots, \mathbf{b}_{l_{n_{BS}}}\}$.

VII. Fusion of Bands Selected for Multiple Targets By CTBS

As noted, CTBS is particularly designed for target detection with a desired target signature $\mathbf{d}$ being available by prior knowledge. In order for CTBS to select bands for detection of multiple targets, different target signatures must be used to select different sets of bands. Consequently, a different target signature will result in a different set of selected bands. In addition, the number of selected bands, n_{BS} will also vary with the target signature $\mathbf{d}$. For example, if there are a set of p target signatures of interest, $\{\mathbf{d}_j\}_{j=1}^p$, then CTBS must be performed p times, one for each target of interest. As a result, the number of selected bands, n_j for each $\mathbf{d}_j$ is different. So is the set of selected bands for each $\mathbf{d}_j$, n_j.

Generally, there two band fusions can be done naturally. Let $\Omega^j_{n_j} = \{\mathbf{b}^j_1, \mathbf{b}^j_2, \ldots, \mathbf{b}^j_{l_{n_j}}\}$ be band subsets generated by one of two CEM-based BS techniques, SF-CTBS and SB-CTBS for each of the desired signature $\{\mathbf{d}_j\}_{j=1}^p$. In general, there are two logic ways to fuse bands. One is to find the overlapped bands by taking the intersection of $\{\Omega^j_{n_j}\}_{j=1}^p$, $\tilde{\Omega}^p_{\mathrm{BS}} = \bigcap_{j=1}^p \Omega^j_{n_j}$. The main problem with this approach is that on many occasions $\tilde{\Omega}^p_{\mathrm{BS}}$ may be empty, i.e., $\tilde{\Omega}^p_{\mathrm{BS}} = \emptyset$. The other way is to find the joint bands by taking the union of $\{\Omega^j_{n_j}\}_{j=1}^p$, i.e., $\Omega^p_{\mathrm{BS}} = \bigcup_{j=1}^p \Omega^j_{n_j}$. The main issue arising from this approach is that Ω^p_{BS} may be too large in most cases.

A. Band Fusion for CTBS-Based Methods

In order to resolve the dilemmas described above, this section presents a novel approach to fusing different sets of bands selected for multiple targets by CTBS.

Assume that $\{\Omega^j_{n_j}\}_{j=1}^p$ are p band subsets, selected by a CTBS-based method for each of $\{\mathbf{d}_j\}_{j=1}^p$. The proposed idea is to assign each selected band $\mathbf{b}^j_{l_m}$ in $\{\Omega^j_{n_j}\}_{j=1}^p$ a band count denoted by $n(\mathbf{b}^j_{l_m})$, according to how frequently the band appears in these p band subsets. More specifically, if a band, $\mathbf{b}^j_{l_m}$, is selected for q times with $1 \le q \le p$, then its band count $n(\mathbf{b}^j_{l_m})$ indicates that the band, $\mathbf{b}^j_{l_m}$, is selected for q targets among $\{\mathbf{d}_j\}_{j=1}^p$. The higher the value of q, the more significant is the band, $\mathbf{b}^j_{l_m}$. There are two extreme cases. One is the best scenario where $q = p$ in which case the selected band works for all p targets. The other is the worst scenario where $q = 1$ in which case the selected band only works for one particular target. Such fusion process is referred to band fusion selection (BFS) and can be summarized as follows.

BFS

1. Let $\Omega^j_{n_j} = \{\mathbf{b}^j_1, \mathbf{b}^j_2, \ldots, \mathbf{b}^j_{l_{n_j}}\}$ generated by an CEM-based BS method for the desired target signature $\mathbf{d}_j$.
2. Find $\Omega^p_{\mathrm{BS}} = \bigcup_{j=1}^p \Omega^j_{n_j}$. Let $\mathbf{b}^j_{l_m}$ be a band in $\Omega^j_{n_j}$ with $1 \le l_m \le l_{n_j}$.
3. Calculate the frequency of each band $\mathbf{b}^j_{l_m}$ in Ω^p_{BS} by
   $$n(\mathbf{b}^j_{l_m}) = \sum_{k=1}^p I_{\Omega^k_{n_{\mathrm{BS}}}}(\mathbf{b}^j_{l_m}) \tag{34}$$
   where $I_{\Omega^k_{n_{\mathrm{BS}}}}(\mathbf{b}^j_{l_m})$ is an indicator function defined by
   $$I_{\Omega^k_{n_k}}(\mathbf{b}^j_{l_m}) = \begin{cases} 1; & \text{if } \mathbf{b}^j_{l_m} \in \Omega^k_{n_k} \\ 0; & \text{if } \mathbf{b}^j_{l_m} \notin \Omega^k_{n_k} \end{cases} \tag{35}$$
4. For $1 \le j \le p, 1 \le l_m \le n_j$ define the priority probability of $\mathbf{b}^j_{l_m}$ as
   $$p(\mathbf{b}^j_{l_m}) = \frac{n\{\mathbf{b}^j_{l_m}\}}{\sum_{k=1}^p \sum_{\mathbf{b}^k_{l_n}\in\Omega^k_{n_k}} n(\mathbf{b}^k_{l_n})} \tag{36}$$
5. Rank all the bands in Ω^p_{BS} according to their priority probabilities $\{p(\mathbf{b}^k_{l_n})\}_{k=1,n=1}^{p,n_k}$ calculated by (36), that is,
   $$\mathbf{b}^j_{l_m} \succ \mathbf{b}^k_{l_n} \Leftrightarrow p(\mathbf{b}^j_{l_m}) > p(\mathbf{b}^k_{l_n}) \tag{37}$$
   where "$A \succ B$" indicates "A has a higher ranking than B." It should be noted when $p(\mathbf{b}^j_{l_m}) = p(\mathbf{b}^k_{l_n})$, $\mathbf{b}^j_{l_m}$ and $\mathbf{b}^k_{l_n}$ have the same priority probability, in which case the one yielding the higher priority for a particular target signature will be selected.

B. Discussion on CTBS

As it is designed, CTBS is a BS approach to multiple target detection. Since each target has different spectral characteristics, it requires a different value of n_j. How to determine the number of bands, n_{BS}, required for BS has been a long standing issue. A general approach is to use VD to determine, n_{BS}. However, it is by no means that VD is the only means to be used for this purpose. There are also many methods available

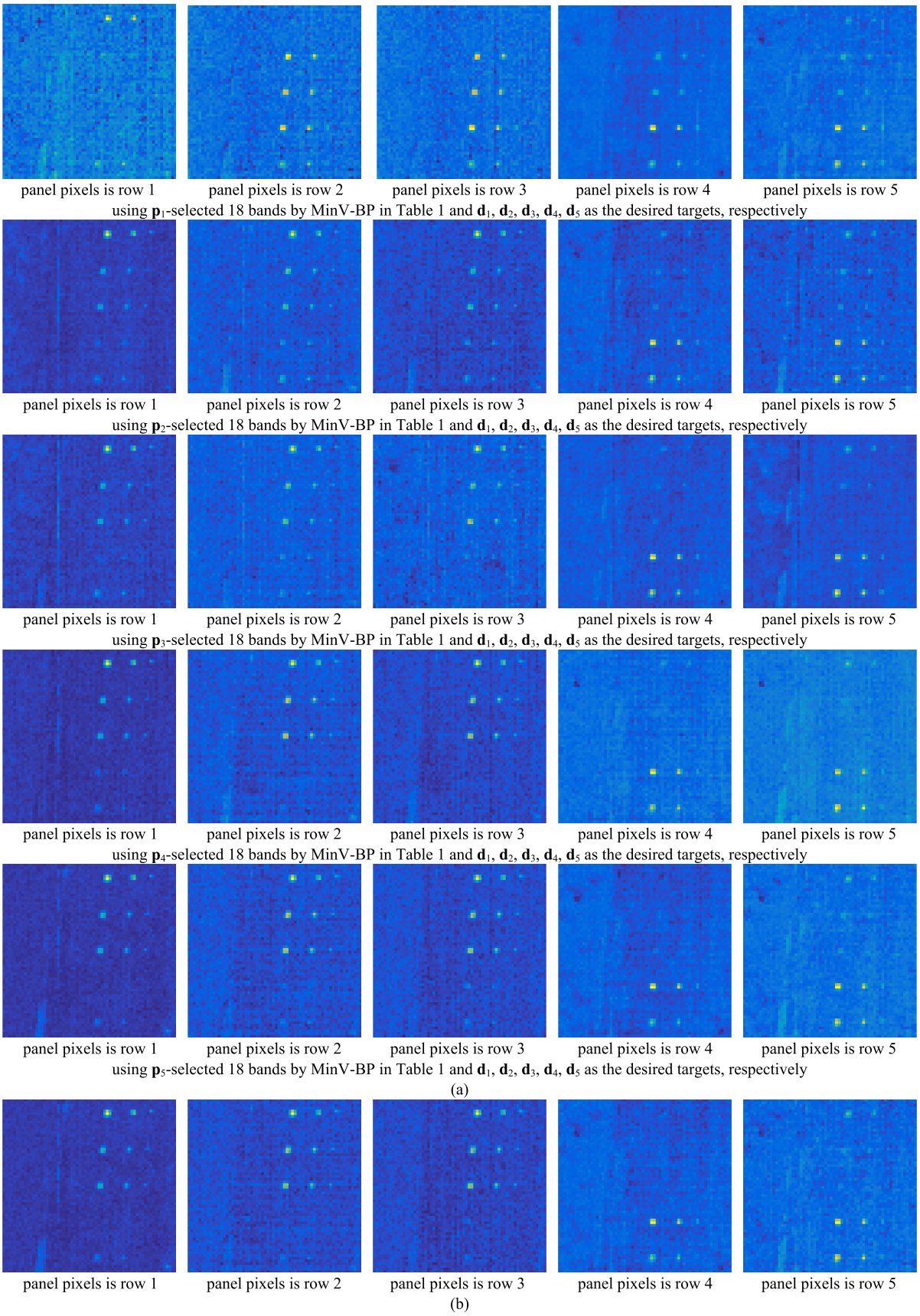


Fig. 2. CEM detection of 19 R panel pixels using 18 bands selected by (a) MinV-BP and (b) BF-MinV-BP in Table I.

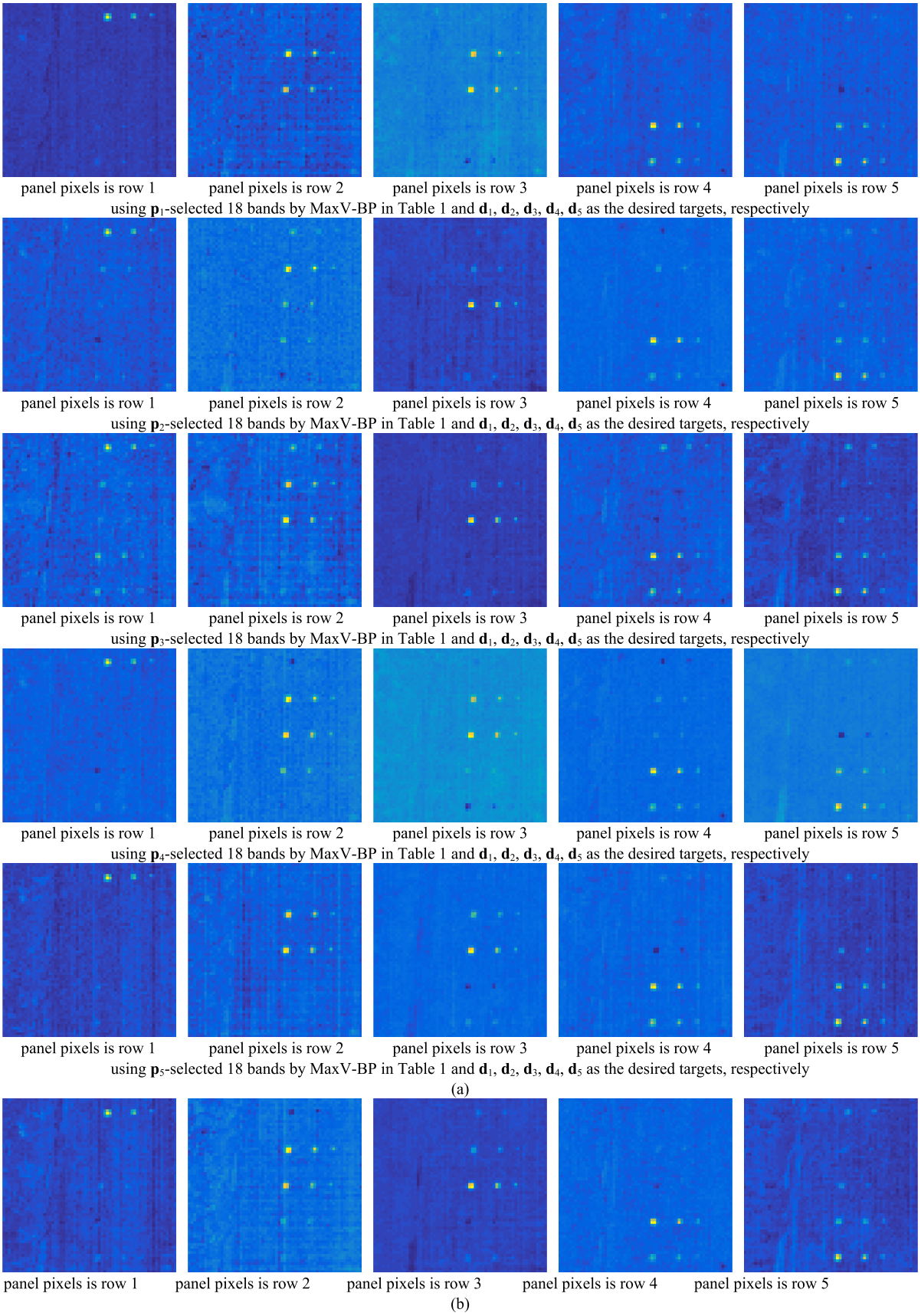


Fig. 3. CEM detection of 19 R panel pixels using 18 bands selected by (a) MaxV-BP and (b) BF-MaxV-BP in Table I.

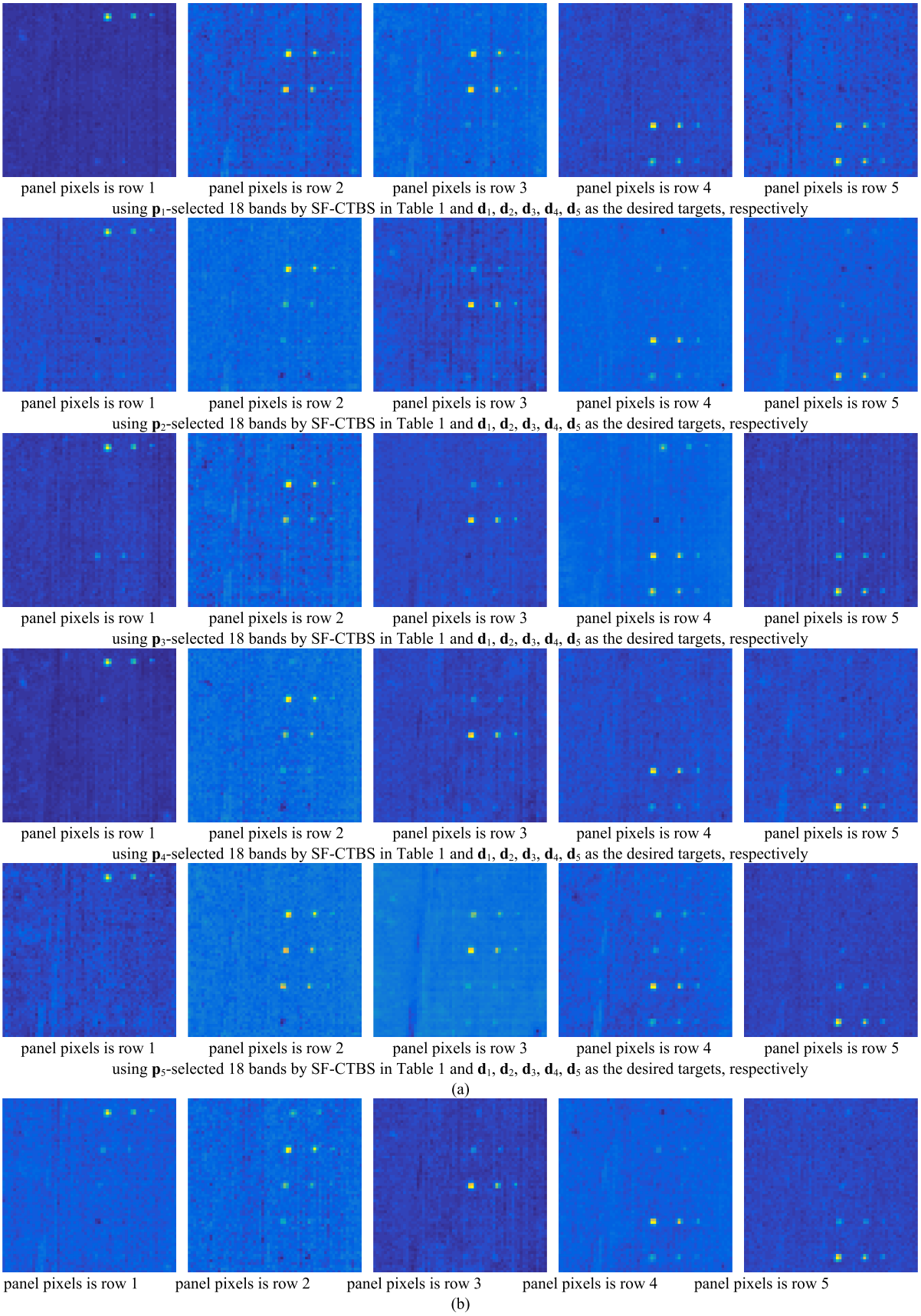


Fig. 4. CEM detection of 19 R panel pixels using 18 bands selected by (a) SF-CTBS and (b) BF-SF-CTBS in Table I.

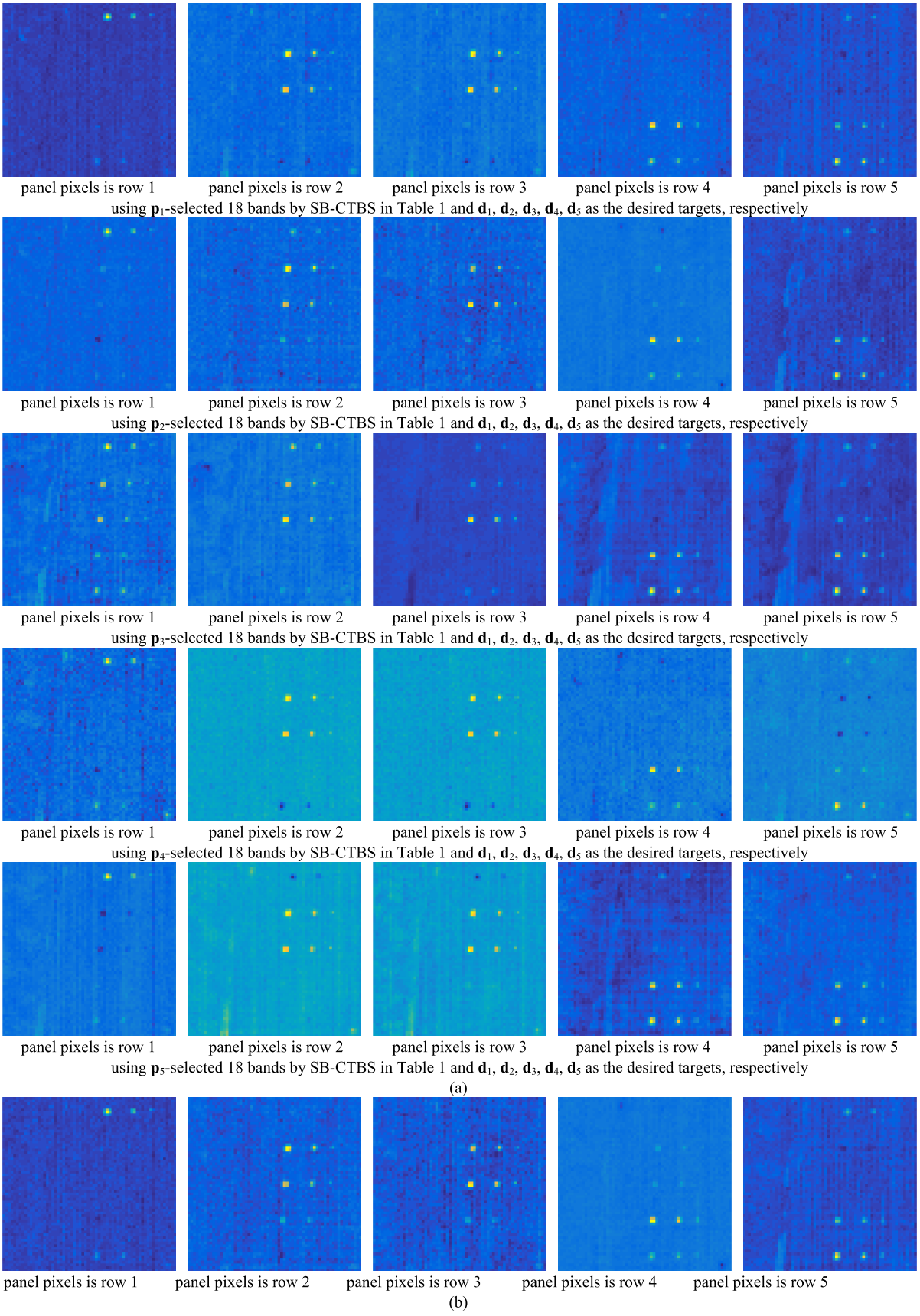


Fig. 5. CEM detection of 19 R panel pixels using 18 bands selected by (a) SB-CTBS and (b) BF-SB-CTBS in Table I.

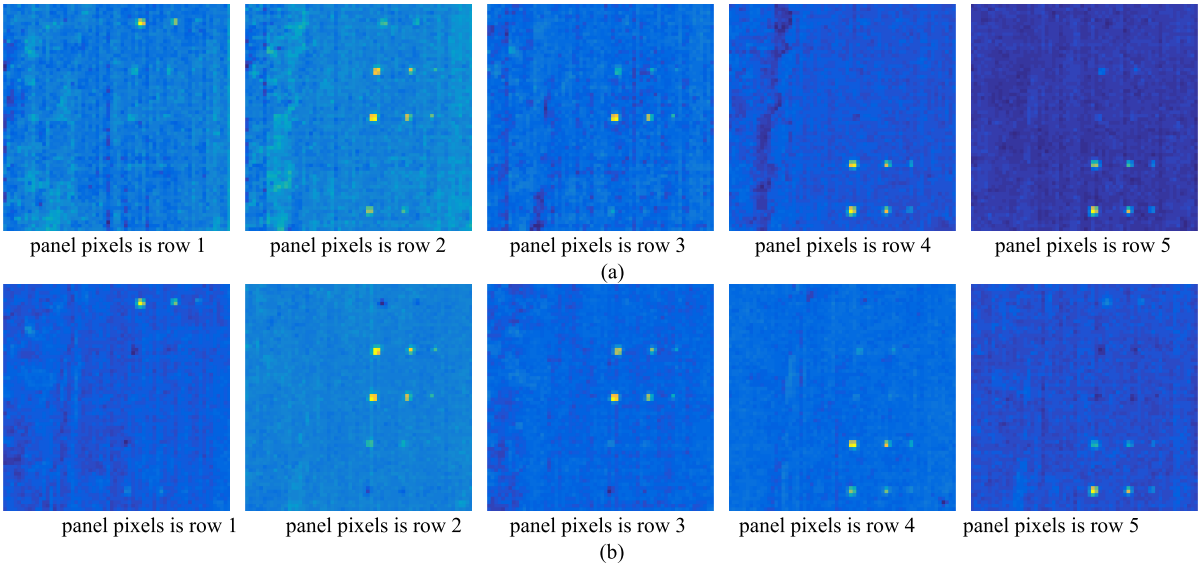


Fig. 6. CEM detection of 19 R panel pixels using 18 bands selected by (a) CBS and (b) UBS in Table I.

for VD estimation [1], [23], [29], [30], specifically, a review of VD in [31] which provides useful information and guidelines.

Once the number of bands is determined for BS by such as VD or $n_{\mathrm{BS}} = \max_{1\le j\le p}\{n_j\}$, we can select n_{BS} bands from $\{\Omega^j_{n_j}\}^p_{j=1}$ according to (42). When SF-CTBS is implemented in the above BFS method, it is called BF-SF-CTBS. Similarly, it is called BF-SB-CTBS if SB-CTBS is implemented in conjunction with the BFS method.

Several comments are noteworthy.

1) As noted in CEM, the target signature $\mathbf{d}$ is a spectral signature and is not necessary one of data sample vectors. It can be obtained by either a spectral library/database or training sample vectors.
2) n_{BS} can be determined by VD, in which case we assume that the value of VD, n_{VD}, is sufficiently large such that $n_{\mathrm{BS}} = n_{\mathrm{VD}} = \max_{1\le j\le p}\{n_j\}$.
3) For the worse case, we can assume that $n_{\mathrm{BS}} = L$, i.e., total number of bands. In this case, we can use one of the four above CTBS methods to produce entire bands prioritized by their corresponding prioritization equations (19), (22), (28), and (33). Since these L bands are prioritized, different target signatures will select different numbers from these L prioritized bands ranked in different orders. Then we can further use the progressive band processing of CEM developed in [32] to implement CEM for each $\mathbf{d}_j$ to produce L-band progressive profiles of CEM-detection maps for all target signatures $\{\mathbf{d}_j\}^p_{j=1}$ up to $n_{\mathrm{BS}} = L$. From these L-band progressive profiles of CEM-detection maps, each $\mathbf{d}_j$ can determine its own n_j by finding a band that has insignificant changes in the CEM detection map. The priority order assigned to this particular band will be n_j.

VIII. Real Hyperspectral Image Experiments

In this section an airborne hyperspectral digital imagery collection experiment (HYDICE) scene shown in Fig. 1 was used to demonstrate the utility of CTBS. The data were collected in August 1995 from a flight altitude of 10 000 ft. This scene has been studied extensively by many reports such as [1], [23], [27], and [30].

There are 15 square panels in Fig. 1(b) with three different sizes, 3 m × 3 m, 2 m × 2 m, and 1 m × 1 m, respectively. Due to the ground sampling distance approximately 1.56 m the each of panels in the first column except the first row contains two panel pixels highlighted by red, p_{211}, p_{221} in row 2, p_{311}, p_{312} in row 3, p_{411}, p_{412} in row 4, p_{511}, p_{521} in row 5 as shown in Fig. 1. All the 11 remaining panels in Fig. 1 contain one single panel pixel for each panel also highlighted by red, p_{11}, p_{12}, p_{13} in row 1, p_{22}, p_{23} in row 2, p_{32}, p_{33} in row 3, p_{42}, p_{43} in row 4, p_{52}, p_{53} in row 5. Therefore, there are a total of 19 R panel pixels. Fig. 1(b) shows their precise spatial locations with the pixels in yellow (Y pixels) indicating panel pixels mixed with the BKG. Fig. 1(c) shows the five panel spectral signatures, $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$, $\mathbf{p}_5$ obtained from Fig. 1(b), where the ith panel signature, denoted by $\mathbf{p}_i$ was generated by averaging the red panel center pixels in row i. These panel signatures will be used as desired target signatures for CTBS to represent target knowledge of the panels in each row.

A. Experiments

The VD of this scene was estimated by the HFC method in [23] and [33] as 9. However, it was demonstrated in [34], [35] that 18 is a more appropriate value than 9. So, in the following experiments, 18 was used as the number of bands to be selected. Table I lists 18 bands selected by uniform band selection (UBS), CBS [5], MinV-BP, MaxV-BP, SF-CTBS, SB-CTBS, and their corresponding BFS methods where UBS was selected because it is the simplest BS method and also generally performs reasonably well. There are also two reasons to choose CBS for comparison. The most important one is that its designed rationale is also based on CEM. A second one is that CBS was compared and shown to perform better than

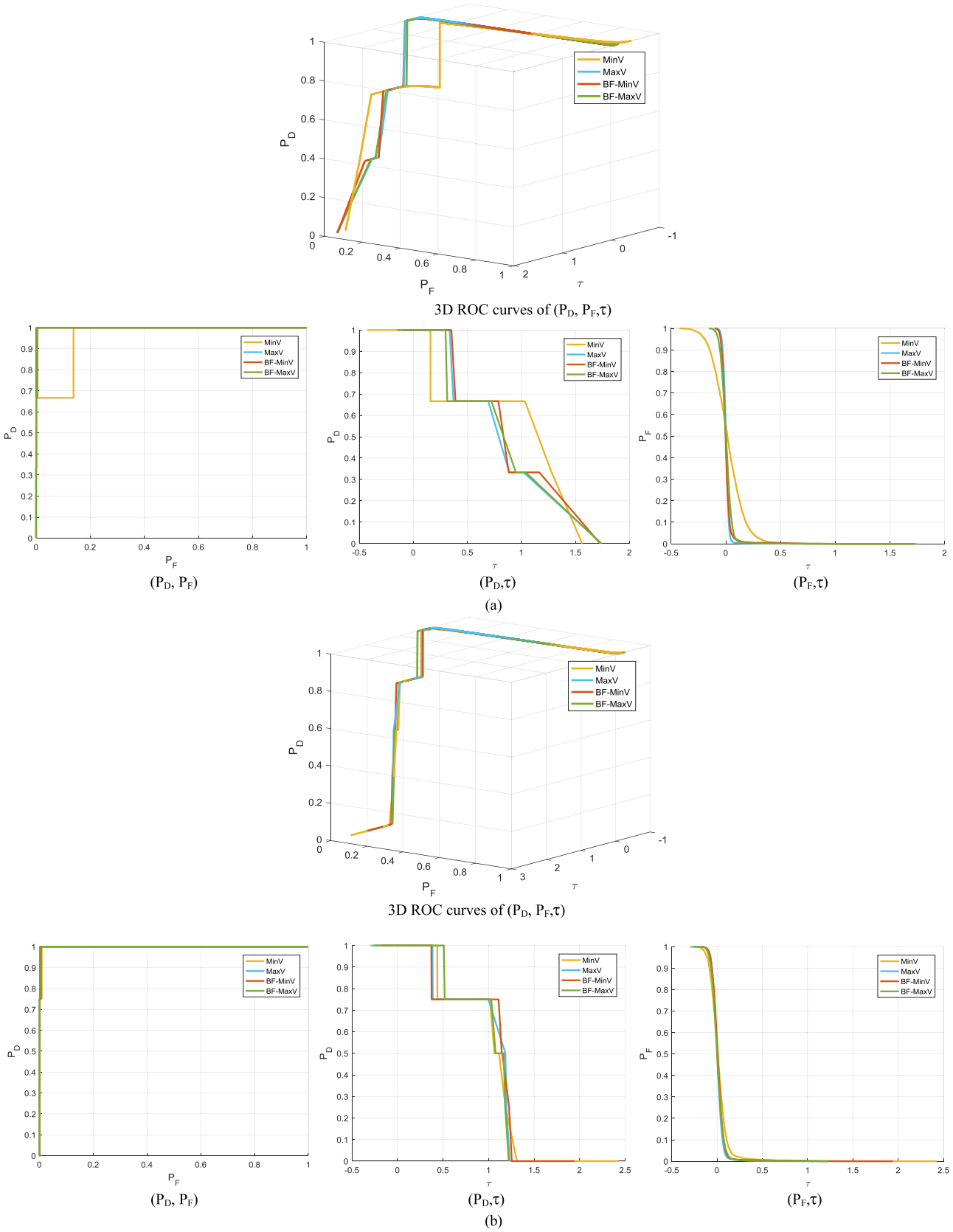
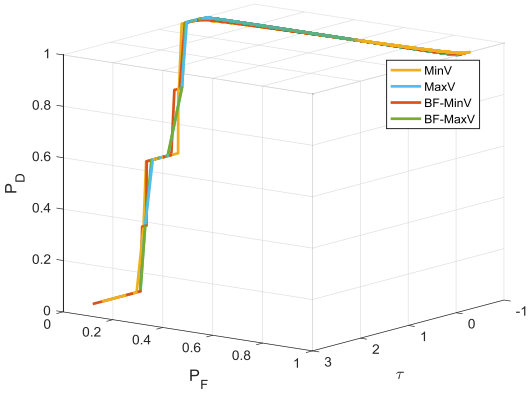
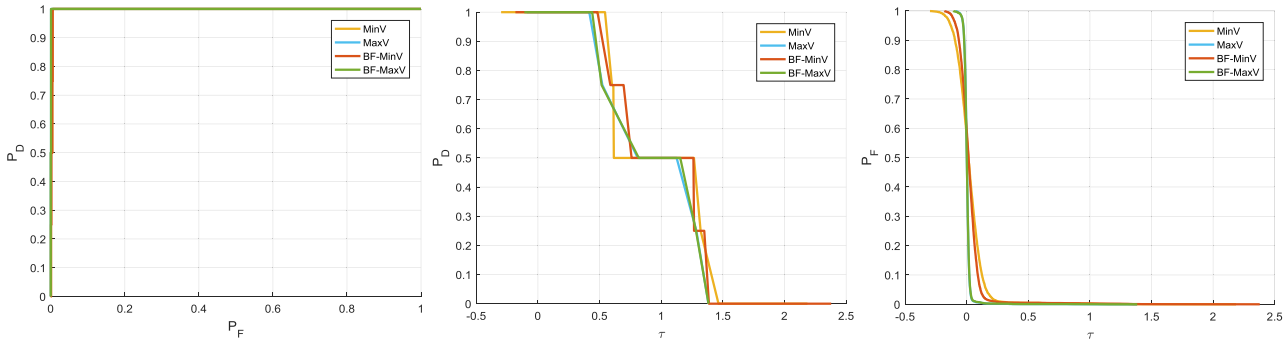


Fig. 7. CEM detection of 19 R panel pixels using 18 bands selected by MinV-BP, MaxV-BP, BF-MinV-BP, and BF-MaxV-BP in Table I. (a) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 1. (b) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 2. (c) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 3. (d) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 4. (e) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 5.

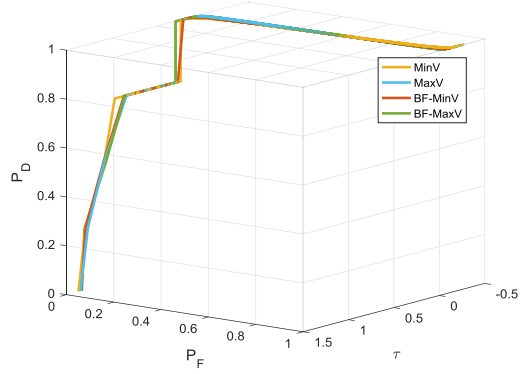


3D ROC curves of (P_D, P_F,τ)

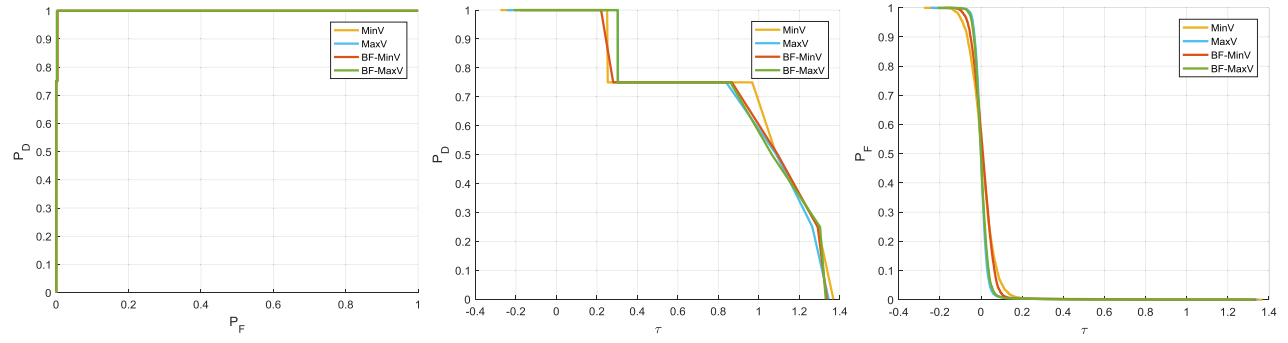


(P_D, P_F) (P_D,τ) (P_F,τ)

(c)



3D ROC curves of (P_D, P_F,τ)



(P_D, P_F) (P_D,τ) (P_F,τ)

(d)

Fig. 7. *(Continued.)* CEM detection of 19 R panel pixels using 18 bands selected by MinV-BP, MaxV-BP, BF-MinV-BP, and BF-MaxV-BP in Table I. (a) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 1. (b) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 2. (c) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 3. (d) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 4. (e) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 5.

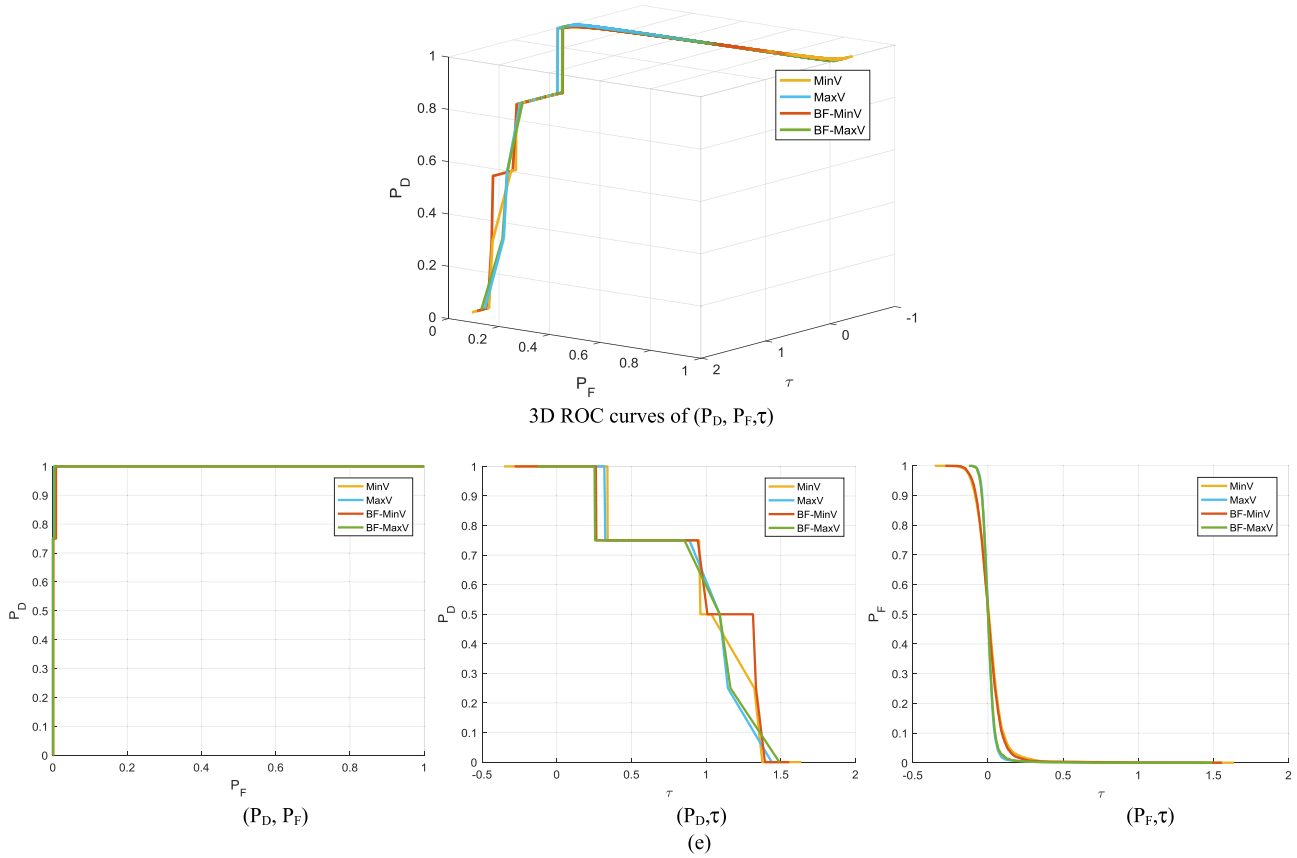


Fig. 7. *(Continued.)* CEM detection of 19 R panel pixels using 18 bands selected by MinV-BP, MaxV-BP, BF-MinV-BP, and BF-MaxV-BP in Table I. (a) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 1. (b) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 2. (c) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 3. (d) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 4. (e) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 5.

many other BS methods in [5] in which case there are no need to include other methods here for comparison as long as we can show that CTBS indeed performs better than CBS, which is the case according to Tables II–IV.

Figs. 2(a), 3(a), 4(a), and 5(a) show detection results of CEM using 18 bands selected in Table I by four CTBS-based methods, MinV-BP, MaxV-BP, SF-CTBS, and SB-CTBS while each of $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$, $\mathbf{p}_5$ was used to select bands to detect panel pixels in row 1–row 5. Figs. 2(b), 3(b), 4(b), and 5(b) also show detection results of CEM using 18 bands selected in Table I by four BFS-CTBS methods, BF-MinV-BP, BF-MaxV-BP, BF-SF-CTBS, and BF-SB-CTBS which used fused bands selected by MinV-BP, MaxV-BP, SF-CTBS, and SB-CTBS to detect panel pixels in row 1–row 5. For a further comparison, CBS and UBS were also implemented using 18 bands and their results were shown in Fig. 6(a) and (b).

Since the panels in rows 2 and 3 were made by the same material and painted by slightly different colors. As expected, detecting panels in row 2 will be very likely also to detect panels in row 3 and vice versa. This was indeed the case in our experiments and so was the detection of panels in rows 4 and 5 since the panels in rows 4 and 5 were made by the same material and painted by slightly different colors. According to visual inspection of Figs. 2–6, it is difficult to see which one would produce best results since there was not much appreciable difference to make a call. Nevertheless, it seemed that the CEM detection results using bands selected by particular signatures were generally not as good as the results produced by their fusion counterparts.

B. 3-D ROC Analysis

Since Figs. 2–6 only provide qualitative analysis, to further evaluate detection performance of Figs. 2–6, the 3-D receiver operating characteristic (ROC) analysis developed in [1], [36], and [37] was used for detection analysis where a 3-D ROC curve is a plot specified by three parameters, detection probability (P_D), false alarm probability (P_F), the threshold τ that is used to calculate P_D and P_F. This 3-D ROC curve can be used to generate three types of 2-D ROC curves, i.e., 2-D ROC curve of (P_D,P_F), 2-D ROC curve of (P_D, τ), and 2-D ROC curve of (P_F, τ). Figs. 7 and 8 show 3-D ROC curves along with their corresponding three 2-D ROC curves of detections of 19 R panel pixels in five rows in Fig. 1(b) produced by CEM using 18 bands selected by four CTBS-BP based methods, MinV-BP, MaxV-BP, BF-MinV-BP, and BF-MaxV-BP in Tables I and IV and CTBS-BS based methods, SF-CTBS, SB-CTBS, BF-SF-CTBS, BF-SB-CTBS, CBS, and UBS in Table I, respectively.

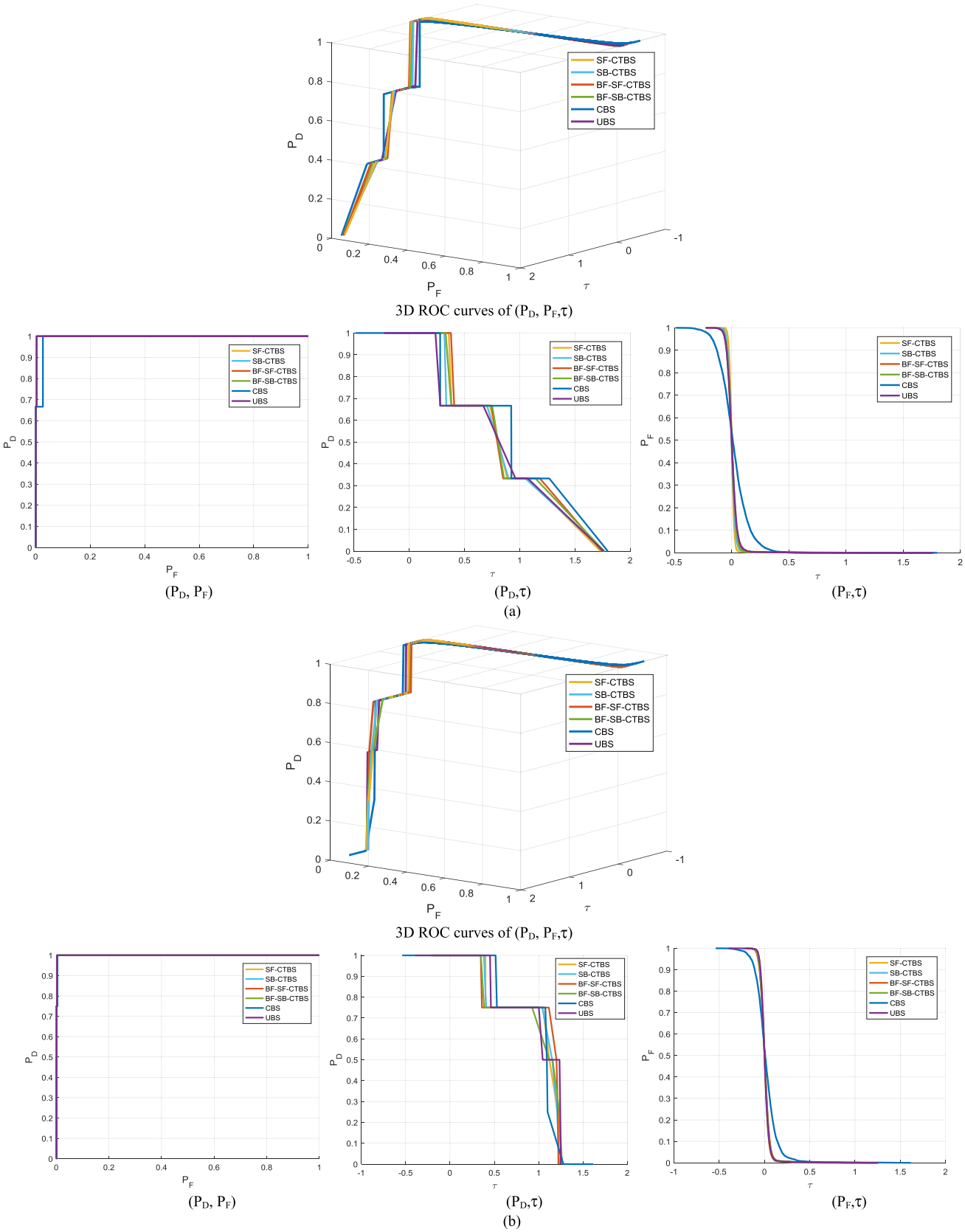


Fig. 8. CEM detection of 19 R panel pixels using 18 bands selected by SF-CTBS, SB-CTBS, BF-SF-CTBS, BF-SB-CTBS, CBS, and UBS in Table I. (a) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 1. (b) 3-D ROC and corresponding 2-D ROC curves for detection of panel pixels in row 2. (c) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 3. (d) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 4. (e) 3-D ROC cures and corresponding 2-D ROC curves for detection of panel pixels in row 5.

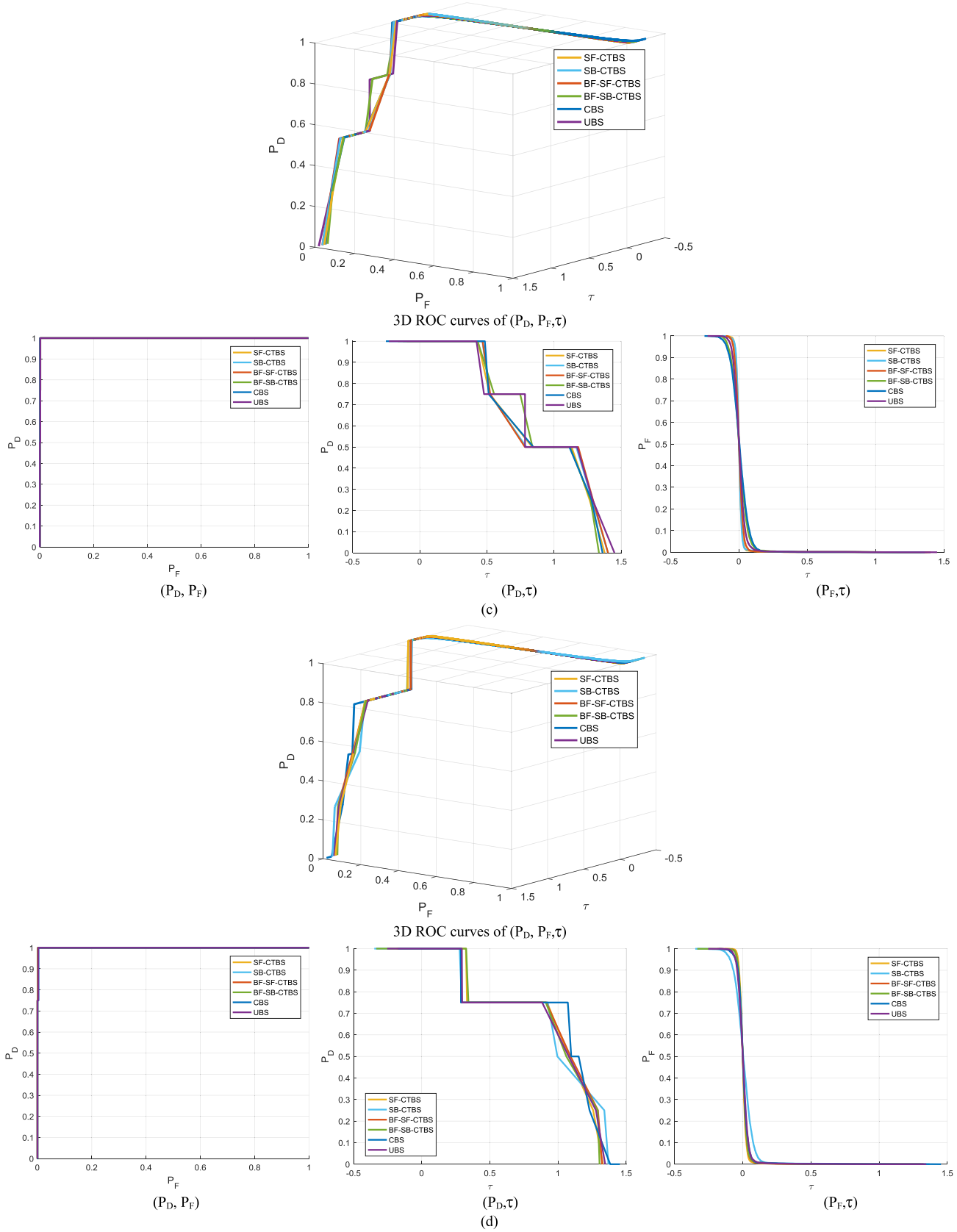


Fig. 8. *(Continued.)* CEM detection of 19 R panel pixels using 18 bands selected by SF-CTBS, SB-CTBS, BF-SF-CTBS, BF-SB-CTBS, CBS, and UBS in Table I. (a) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 1. (b) 3-D ROC and corresponding 2-D ROC curves for detection of panel pixels in row 2. (c) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 3. (d) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 4. (e) 3-D ROC cures and corresponding 2-D ROC curves for detection of panel pixels in row 5.

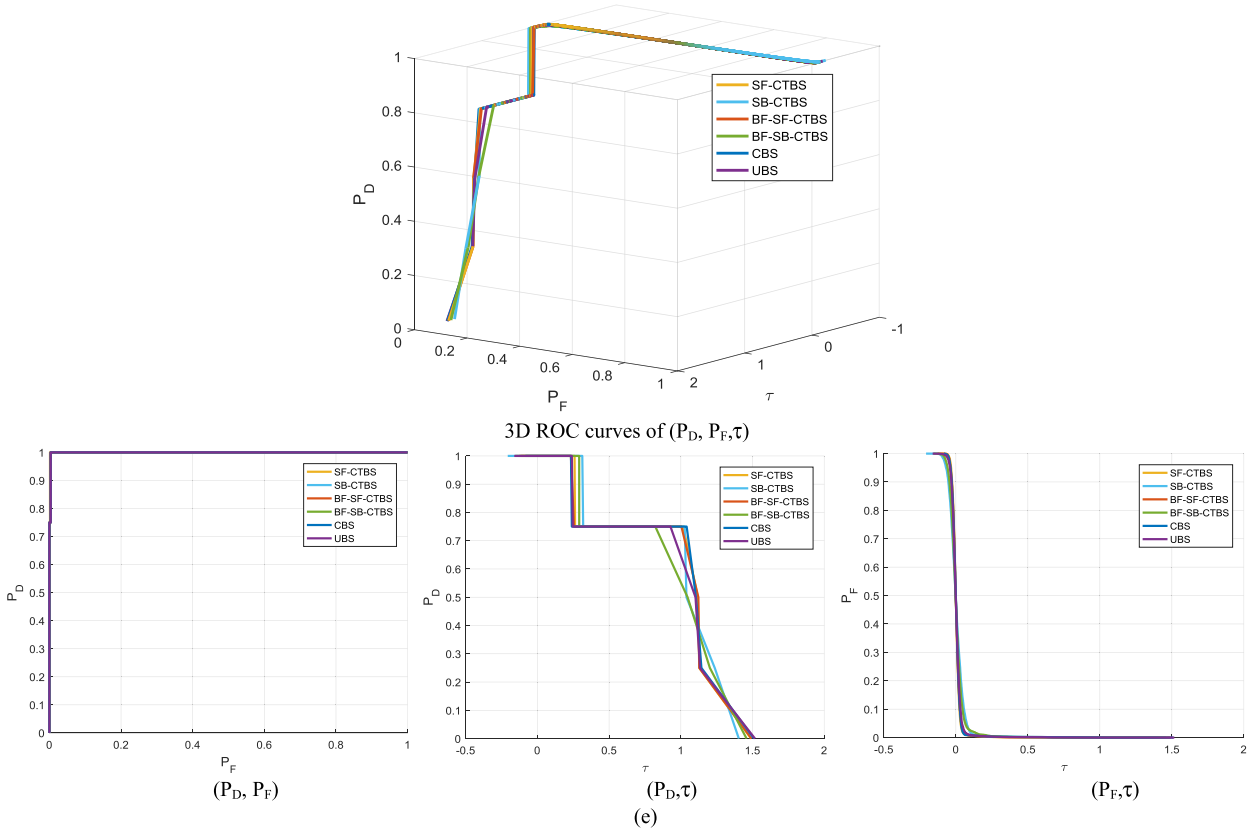


Fig. 8. *(Continued.)* CEM detection of 19 R panel pixels using 18 bands selected by SF-CTBS, SB-CTBS, BF-SF-CTBS, BF-SB-CTBS, CBS, and UBS in Table I. (a) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 1. (b) 3-D ROC and corresponding 2-D ROC curves for detection of panel pixels in row 2. (c) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 3. (d) 3-D ROC curves and corresponding 2-D ROC curves for detection of panel pixels in row 4. (e) 3-D ROC cures and corresponding 2-D ROC curves for detection of panel pixels in row 5.

As shown in Figs. 7 and 8, it is very difficult to see performance difference among different CTBS-BP and CTBS-BS methods from the traditional 2-D ROC curves of (P_D, P_F). However, the other 2-D ROC curves of (P_D, τ) and (P_F, τ) did provide very useful information as the threshold τ varies. Specifically, the 2-D ROC curves of (P_F, τ) can be used to evaluate CEM in the BKG suppression.

It is very obvious that the visual inspection of Figs. 2–6 did not tell the whole story since BKG suppression is a key issue in image analysis. It is noted that BKG suppression plays a significant role in progressive target detection in [27] and real-time target detection in [38], and [39]. Similarly, the same issue of BKG suppression is encountered in progressive anomaly detection [40] and real-time anomaly detection [41], [42]. This is because anomaly detection also uses the inverse of the sample correlation matrix, i.e., $\mathbf{R}^{-1}$, as CEM does to suppress BKG. Such issue of BKG suppression has been discussed in great detail in [27, Ch. 5]. Accordingly, BKG suppression should be included to account for the performance of target detection. Apparently, plots of P_D versus P_F used by the traditional ROC analysis cannot address this issue. Fortunately, a 3-D ROC analysis allows users to generate 2-D ROC curve of P_D versus τ, 2-D ROC curve of P_F versus τ, and 2-D ROC curve of P_D versus P_F, all of which can be jointly used to evaluate detection performance. Specifically, the 2-D ROC curve of P_F versus τ is particularly designed to assess the impact of BKG suppression on target detection. Fig. 7(a)–(d) shows 3-D ROC curves in Fig. 7(a), 2-D ROC curves of P_D versus τ in Fig. 7(b), 2-D ROC curves of P_F versus τ in Fig. 7(c), and 2-D ROC curves of P_D versus P_F in Fig. 7(d) produced by the results in Figs. 2 and 3. Similarly, Fig. 8(a)–(d) also shows 3-D ROC curves in Fig. 8(a), 2-D ROC curves of P_D versus τ in Fig. 8(b), 2-D ROC curves of P_F versus τ in Fig. 8(c), and 2-D ROC curves of P_D versus P_F in Fig. 8(d) produced by the results in Figs. 4–6.

In order to further conduct a quantitative study, Tables II–IV calculate the area under curve (AUC) of the three types of 2-D ROC curves of (P_D, P_F), (P_D, τ), and (P_F, τ), respectively, where best results are boldfaced. In other words, the better the detection, the higher the AUC value of (P_D, P_F), the higher the AUC value of (P_D, τ), and the smaller the AUC value of (P_F, τ). Due to limited space, the figures showing the CEM detection results using nine bands from Table I are not included here, but their AUC results are included in Tables II–IV for comparison. Since the bands in Table I were selected sequentially, the nine bands used by CTBS were simply to select the first nine bands in Table I to produce results. In addition, the last column in Tables II–IV also calculated the averaged AUC of panels in the five rows using selected 9 and 18 bands for comparison.

TABLE I

18 BANDS PRIORITIZED BY BF-MINV-BP, BF-MAXV-BP, BF-MINV-BP, AND BF-MAXV-BP, AND 18 BANDS SELECTED BY UBS, SF-CTBS, SB-CTBS, BF-SF-CTBS, AND BF-SB-CTBS

MinV-BP	$\mathbf{p}_1$:	169	122	123	168	167	166	124	162	164	165	161	160	158	132	163	157	159	133
	$\mathbf{p}_2$:	122	169	123	132	133	131	134	168	135	126	124	130	125	127	136	98	167	137
	$\mathbf{p}_3$:	122	123	132	133	124	131	134	135	52	169	167	130	168	53	51	136	127	125
	$\mathbf{p}_4$:	123	122	124	125	127	128	126	129	130	131	132	133	98	134	99	135	100	168
	$\mathbf{p}_5$:	122	123	124	12	126	168	127	169	128	129	130	167	166	131	165	132	164	133
BF-MinV-BP		122	123	124	132	133	168	169	125	127	167	131	130	126	134	135	128	166	129
MaxV-BP	$\mathbf{p}_1$:	109	6	110	128	17	142	122	116	57	24	28	130	106	132	140	51	111	112
	$\mathbf{p}_2$:	133	44	50	129	143	118	45	46	154	47	71	144	29	101	38	153	115	66
	$\mathbf{p}_3$:	123	44	45	51	50	52	19	99	131	98	38	46	5	169	113	60	88	94
	$\mathbf{p}_4$:	158	152	154	61	155	63	98	138	52	105	118	25	27	107	153	82	18	84
	$\mathbf{p}_5$:	61	57	136	52	67	124	27	28	73	114	113	128	38	25	63	86	65	101
BF-MaxV-BP		52	38	61	44	57	45	50	154	51	128	63	118	27	98	28	46	113	25
SF-CTBS	$\mathbf{p}_1$:	169	127	122	105	112	149	130	116	128	123	148	17	22	2	106	109	142	132
	$\mathbf{p}_2$:	122	128	133	154	143	137	142	148	153	11	27	19	6	138	46	49	42	129
	$\mathbf{p}_3$:	122	126	53	41	118	111	15	52	45	51	38	107	44	5	49	26	37	98
	$\mathbf{p}_4$:	123	166	50	128	158	101	152	141	149	46	153	25	134	9	38	112	104	140
	$\mathbf{p}_5$:	122	49	25	17	39	154	136	28	153	18	27	38	52	132	155	26	151	14
BF-SF-CTBS		122	49	128	153	38	123	25	17	154	112	149	142	52	148	46	27	132	26
SB-CTBS	$\mathbf{p}_1$:	109	110	111	112	122	113	132	137	138	140	123	139	108	114	136	142	141	120
	$\mathbf{p}_2$:	133	129	143	144	118	44	46	47	45	43	115	78	154	153	142	33	35	32
	$\mathbf{p}_3$:	123	45	44	46	47	43	51	52	50	49	53	54	55	56	57	48	70	39
	$\mathbf{p}_4$:	158	152	154	155	153	63	156	151	150	149	159	148	147	157	161	163	166	165
	$\mathbf{p}_5$:	61	62	73	67	124	58	94	114	27	28	25	26	24	113	136	23	21	29
BF-SB-CTBS		123	45	44	154	46	47	153	43	113	114	136	142	61	109	133	158	62	110
CBS		62	77	63	61	13	91	30	69	76	56	38	45	16	20	39	34	24	47
UBS		1	10	19	28	37	46	55	64	73	82	91	100	109	118	127	136	145	154

TABLE II

AUC OF (P_D, P_F) FOR CEM DETECTION RESULTS USING 9 AND 18 BANDS PRIORITIZED BY MINV-BP, BF-MINV-BP, MAXV-BP, AND BF-MAXV-BP, AND SELECTED BY SF-CTBS, BF-SF-CTBS, SB-CTBS, AND BF-SB-CTBS ALONG WITH 9 AND 18 BANDS BY CBS AND UBS

		Detection of panel pixels in rows											
BS algorithms	Desired targets for BS	panel pixels in row 1		panel pixels in row 2		panel pixels in row 3		panel pixels in row 4		panel pixels in row 5		Average AUCs of 5 rows	
		9 bands	18 bands	9 bands	18 bands	9 bands	18 bands	9 bands	18 bands	9 bands	18 bands	9 bands	18 bands
	$\mathbf{p}_1$	0.997	0.961	0.996	0.987	0.996	0.981	0.997	0.997	0.998	0.998	0.997	0.985
	$\mathbf{p}_2$	0.922	0.990	0.903	0.992	0.865	0.983	0.971	0.997	0.995	0.998	0.931	0.992
MinV-BP	$\mathbf{p}_3$	0.994	0.992	0.988	0.992	0.911	0.772	0.992	0.998	0.996	**0.999**	0.976	0.951
	$\mathbf{p}_4$	0.987	0.989	0.985	0.966	0.982	0.750	0.995	0.957	0.996	0.996	0.989	0.932
	$\mathbf{p}_5$	0.994	0.979	0.994	0.970	0.993	0.944	0.991	0.995	0.997	0.998	0.994	0.977
BF-MinV-BP	Fusion	0.982	0.976	0.985	0.986	0.955	0.981	0.990	0.996	0.997	0.997	0.982	0.987
	$\mathbf{p}_1$	0.997	0.905	0.786	0.591	0.571	0.983	0.995	0.998	0.997	0.999	0.869	0.895
	$\mathbf{p}_2$	0.938	0.929	0.637	0.778	0.991	0.962	0.996	0.997	0.998	0.999	0.912	0.933
MaxV-BP	$\mathbf{p}_3$	0.997	**0.995**	0.995	0.987	0.994	0.960	0.997	0.996	0.997	0.998	0.996	0.987
	$\mathbf{p}_4$	0.765	0.818	0.760	0.946	0.998	0.997	0.997	0.997	**0.999**	0.999	0.904	0.951
	$\mathbf{p}_5$	0.864	0.845	0.848	0.528	0.959	0.991	0.997	0.998	0.998	0.999	0.933	0.872
BF-MaxV-BP	Fusion	0.992	0.971	0.992	0.920	0.986	0.984	**0.999**	0.997	0.999	0.999	0.994	0.974
	$\mathbf{p}_1$	0.995	0.985	0.778	0.573	0.826	0.941	0.997	0.997	0.998	0.999	0.919	0.899
	$\mathbf{p}_2$	0.991	0.989	0.875	0.857	0.850	0.986	0.997	0.997	0.998	0.999	0.942	0.966
SF-CTBS	$\mathbf{p}_3$	0.997	0.925	0.936	0.778	0.991	0.981	0.995	0.997	0.998	0.999	0.983	0.936
	$\mathbf{p}_4$	0.757	0.786	0.711	0.879	0.631	0.975	0.998	0.997	0.999	0.999	0.819	0.927
	$\mathbf{p}_5$	0.969	0.983	0.916	0.895	0.818	0.977	0.996	0.990	0.999	0.999	0.940	0.969
BF-SF-CTBS	Fusion	0.906	0.943	0.986	0.914	0.980	0.981	0.997	0.996	0.999	0.999	0.974	0.967
	$\mathbf{p}_1$	**0.999**	0.913	0.998	0.984	0.998	**0.999**	0.998	0.998	0.999	0.999	**0.998**	0.979
	$\mathbf{p}_2$	0.999	0.962	0.544	0.569	0.617	0.623	0.996	0.997	0.997	0.998	0.831	0.830
SB-CTBS	$\mathbf{p}_3$	0.999	0.991	0.995	0.989	0.994	0.993	0.995	0.998	0.997	0.998	0.996	**0.994**
	$\mathbf{p}_4$	0.999	0.797	**0.999**	**0.996**	**0.999**	0.996	0.932	0.994	0.998	0.999	0.985	0.956
	$\mathbf{p}_5$	0.999	0.925	0.547	0.896	0.684	0.659	0.998	0.998	0.999	0.999	0.845	0.895
BF-SB-CTBS	Fusion	0.999	0.962	0.992	0.872	0.978	0.920	0.997	0.997	0.997	0.999	0.993	0.950
CBS	CBS	0.936	0.819	0.642	0.983	0.843	0.817	0.998	**0.999**	0.998	0.999	0.883	0.923
UBS	UBS	0.989	0.984	0.963	0.984	0.870	0.984	0.998	0.997	0.997	0.999	0.963	0.990

C. Results and Discussion

Many intriguing findings can be observed from Tables II–IV.

1) First of all, by examining Tables II–IV, it is very obvious that the AUC values in Table II based on 2-D ROC curves of (P_D, P_F) were always nearly one which shows

TABLE III

AUC OF (P_D, τ) FOR CEM DETECTION RESULTS USING 9 AND 18 BANDS PRIORITIZED BY MINV-BP, BF-MINV-BP, MAXV-BP, AND BF-MAXV-BP, AND SELECTED BY SF-CTBS, BF-SF-CTBS, SB-CTBS, AND BF-SB-CTBS ALONG WITH 9 AND 18 BANDS BY CBS AND UBS

BS algorithms	Desired targets for BS	Detection of panel pixels in rows: panel pixels in row 1		panel pixels in row 2		panel pixels in row 3		panel pixels in row 4		panel pixels in row 5		Average AUCs of 5 rows	
		9 bands	18 bands	9 bands	18 bands	9 bands	18 bands	9 bands	18 bands	9 bands	18 bands	9 bands	18 bands
	$\mathbf{p}_1$	**0.747**	0.475	**0.778**	0.461	0.625	**0.450**	0.470	0.420	0.650	0.642	**0.654**	**0.490**
	$\mathbf{p}_2$	0.442	0.185	0.491	0.357	0.522	0.258	0.493	0.511	0.699	0.713	0.529	0.405
MinV-BP	$\mathbf{p}_3$	0.535	0.239	0.572	0.281	0.375	0.157	**0.704**	0.562	0.695	0.700	0.576	0.388
	$\mathbf{p}_4$	0.281	0.135	0.296	0.173	0.292	0.104	0.678	0.554	0.608	**0.757**	0.431	0.345
	$\mathbf{p}_5$	0.461	0.139	0.597	0.184	0.641	0.183	0.489	0.413	0.732	0.658	0.584	0.315
BF-MinV-BP	Fusion	0.252	0.147	0.290	0.166	0.330	0.149	0.498	0.443	0.729	0.688	0.420	0.318
	$\mathbf{p}_1$	0.367	0.110	0.306	0.101	0.383	0.115	0.563	0.458	0.673	0.677	0.458	0.292
	$\mathbf{p}_2$	0.271	0.178	0.104	0.108	0.179	0.140	0.403	0.344	0.688	0.666	0.329	0.287
MaxV-BP	$\mathbf{p}_3$	0.718	0.355	0.345	0.289	0.211	0.118	0.615	0.467	0.670	0.666	0.512	0.379
	$\mathbf{p}_4$	0.340	0.192	0.241	0.123	0.259	0.120	0.499	0.415	0.726	0.709	0.413	0.312
	$\mathbf{p}_5$	0.506	0.128	0.388	0.128	0.205	0.251	0.679	0.522	0.670	0.672	0.489	0.341
BF-MaxV-BP	Fusion	0.511	0.159	0.195	0.239	0.174	0.147	0.681	0.360	0.664	0.648	0.445	0.311
SF-CTBS	$\mathbf{p}_1$	0.153	0.092	0.100	0.106	0.114	0.068	0.433	0.419	0.636	0.677	0.287	0.272
	$\mathbf{p}_2$	0.265	0.153	0.092	0.065	0.077	0.181	0.431	0.340	0.699	0.655	0.313	0.279
	$\mathbf{p}_3$	0.406	0.088	0.258	0.161	0.195	0.141	0.656	0.620	0.667	0.643	0.436	0.331
	$\mathbf{p}_4$	0.190	0.076	0.194	0.081	0.128	0.134	0.473	0.265	0.654	0.662	0.328	0.244
	$\mathbf{p}_5$	0.385	0.147	0.163	0.102	0.427	0.294	0.460	0.333	0.660	0.657	0.419	0.307
BF-SF-CTBS	Fusion	0.187	0.160	0.427	0.064	**0.717**	0.111	0.429	0.313	0.654	0.656	0.483	0.261
	$\mathbf{p}_1$	0.299	0.168	0.150	0.068	0.124	0.070	0.554	0.501	0.684	0.671	0.362	0.296
	$\mathbf{p}_2$	0.234	0.221	0.095	0.109	0.176	0.149	0.657	0.431	0.688	0.651	0.370	0.312
SB-CTBS	$\mathbf{p}_3$	0.722	**0.501**	0.410	0.318	0.221	0.168	0.674	0.664	0.693	0.673	0.544	0.465
	$\mathbf{p}_4$	0.612	0.382	0.152	0.154	0.143	0.138	0.701	0.448	**0.741**	0.706	0.470	0.366
	$\mathbf{p}_5$	0.613	0.312	0.249	0.318	0.352	0.319	0.695	0.690	0.715	0.719	0.525	0.472
BF-SB-CTBS	Fusion	0.689	0.224	0.453	0.055	0.532	0.060	0.478	0.439	0.689	0.666	0.568	0.289
CBS	CBS	0.635	0.146	0.510	**0.511**	0.637	0.219	0.579	**0.695**	0.590	0.642	0.590	0.443
UBS	UBS	0.664	0.202	0.450	0.136	0.341	0.063	0.697	0.480	0.664	0.647	0.563	0.306

perfect detection. However, the AUC results calculated from the 2-D ROC curves of (P_D, τ) show otherwise with detection rates well below 0.5 and best values around 0.7. This is true if we visually inspect the results in Figs. 2–6. The results shown in Tables II–IV provide evidence that the traditional measure using 2-D ROC curve of (P_D, P_F) to evaluate detection performance is not reliable and could be misleading since it only tells half of a story in detection performance analysis.

2) As shown in Table III, the detection results were not good. This is primarily due to the fact that the spectral signatures of panel pixels in rows 2 and 3, $\mathbf{p}_2$ and $\mathbf{p}_3$ are very close. So, using either panel signature would detect panel pixels in both rows as illustrated in Figs. 2–6. Similarly, the spectral signatures of panel pixels in rows 4 and 5, $\mathbf{p}_4$ and $\mathbf{p}_5$ are also very close. So, detecting panel pixels in one row will also detects panel pixel in the other row.
3) To evaluate BKG suppression, a best way is suggested to visually inspect the results in Figs. 2–6 in conjunction with comparison of the AUC values in Table IV since the AUC value calculated from 2-D ROC curve of (P_F, τ) reflects the degree of BKG suppression. That is, the smaller the AUC value of 2-D ROC curve of (P_F, τ), the better is the BKG suppression. For example, the best performance of CEM using $\mathbf{p}_1$ as the desired target signature to detect the panel pixels in row 1 was SF-CTBS which yielded the smallest value of AUC value of 2-D ROC curve of (P_F, τ), 0.041 among all other AUC values. If we examine the top row of Fig. 4(a) produced by SF-CTBS, its $\mathbf{p}_1$-detection map was very clean compared to $\mathbf{p}_1$-detection map produced by other BS methods. Due to the lack of prior knowledge on BKG, it is crucial to include the detection maps in Figs. 2–6 for visual inspection. Unfortunately, the BKG suppression issue has been overlooked in the traditional 2-D ROC analysis.
4) As expected, BF–BS methods may not be the best in detecting individual panel pixels but generally work reasonably well over all the panel pixels.
5) The band fusion results using nine bands were generally better than their corresponding counterparts using 18 bands. However, the conclusions are reversed for using individual bands without fusion. That is, for individual target, signature results using 18 bands were generally better than those using nine bands for all test BS methods. This makes sense. Since the band fusion methods integrate bands from different BS methods, the number of bands is expected to be smaller.
6) Although the TI and TE scenarios designed in [1], [43], and [44] can also be used for experiments with complete ground truth of simulated panel pixels,

TABLE IV

AUC OF (P_F, τ) FOR CEM DETECTION RESULTS USING 9 AND 18 BANDS PRIORITIZED BY MINV-BP, BF-MINV-BP, MAXV-BP, AND BF-MAXV-BP, AND SELECTED BY SF-CTBS, BF-SF-CTBS, SB-CTBS, AND BF-SB-CTBS ALONG WITH 9 AND 18 BANDS BY CBS AND UBS

BS algorithms	Desired targets for BS	Detection of panel pixels in rows: panel pixels in row 1		panel pixels in row 2		panel pixels in row 3		panel pixels in row 4		panel pixels in row 5		Average AUCs of 5 rows	
		9 bands	18 bands	9 bands	18 bands	9 bands	18 bands	9 bands	18 bands	9 bands	18 bands	9 bands	18 bands
	$\mathbf{p}_1$	0.283	0.224	0.240	0.181	0.192	0.175	0.175	0.132	0.206	0.166	0.219	0.176
	$\mathbf{p}_2$	0.159	0.065	0.161	0.116	0.209	0.088	0.175	0.113	0.214	0.118	0.183	0.100
MinV-BP	$\mathbf{p}_3$	0.167	0.073	0.166	0.126	0.204	0.128	0.121	0.110	0.096	0.102	0.151	0.108
	$\mathbf{p}_4$	0.077	0.052	0.121	0.108	0.111	0.087	0.226	0.173	0.158	0.229	0.139	0.130
	$\mathbf{p}_5$	0.144	0.053	0.196	0.098	0.193	0.090	0.172	0.122	0.203	0.184	0.182	0.109
BF-MinV-BP	Fusion	0.087	0.055	0.103	**0.089**	0.119	0.079	0.191	0.125	0.217	0.158	0.143	0.101
	$\mathbf{p}_1$	0.141	0.058	0.272	0.112	0.392	0.246	0.109	0.115	0.090	0.114	0.201	0.129
	$\mathbf{p}_2$	0.168	0.105	0.155	0.187	0.090	0.075	0.142	0.143	0.102	0.126	0.132	0.127
MaxV-BP	$\mathbf{p}_3$	0.213	0.122	**0.084**	0.134	0.094	**0.063**	0.159	0.113	0.085	0.075	0.127	0.101
	$\mathbf{p}_4$	0.257	0.123	0.294	0.198	0.422	0.295	0.298	0.156	0.213	0.220	0.297	0.198
	$\mathbf{p}_5$	0.186	0.078	0.266	0.130	0.401	0.147	0.149	0.143	0.111	0.076	0.222	0.115
BF-MaxV-BP	Fusion	0.132	0.083	0.088	0.175	**0.089**	0.075	0.118	0.138	0.092	0.081	0.104	0.111
	$\mathbf{p}_1$	**0.048**	**0.041**	0.158	0.110	0.196	0.144	0.082	**0.081**	0.131	0.100	0.123	0.095
	$\mathbf{p}_2$	0.138	0.087	0.186	0.158	0.152	0.092	0.177	0.137	0.152	0.114	0.161	0.118
SF-CTBS	$\mathbf{p}_3$	0.100	0.060	0.182	0.136	0.129	0.085	**0.073**	0.144	0.069	0.063	0.111	0.097
	$\mathbf{p}_4$	0.115	0.044	0.156	0.171	0.168	0.076	0.153	0.090	0.078	0.087	0.134	0.094
	$\mathbf{p}_5$	0.282	0.082	0.254	0.192	0.395	0.208	0.127	0.115	0.085	0.071	0.229	0.134
BF-SF-CTBS	Fusion	0.153	0.117	0.235	0.145	0.219	0.066	0.185	0.124	0.065	0.079	0.171	0.106
	$\mathbf{p}_1$	0.095	0.065	0.388	0.166	0.363	0.195	0.096	0.130	0.127	0.098	0.214	0.131
	$\mathbf{p}_2$	0.080	0.136	0.092	0.149	0.163	0.136	0.107	0.199	0.092	0.081	0.107	0.140
SB-CTBS	$\mathbf{p}_3$	0.157	0.162	0.131	0.198	0.093	0.083	0.100	0.099	0.108	0.080	0.118	0.124
	$\mathbf{p}_4$	0.265	0.157	0.382	0.364	0.383	0.348	0.220	0.206	0.323	0.252	0.315	**0.265**
	$\mathbf{p}_5$	0.193	0.193	0.273	0.363	0.253	0.337	0.116	0.103	0.142	0.130	0.196	0.225
BF-SB-CTBS	Fusion	0.146	0.077	0.196	0.127	0.203	0.113	0.167	0.205	0.104	0.098	0.163	0.124
CBS	CBS	0.466	0.220	0.495	0.258	0.487	0.160	0.168	0.109	0.171	**0.054**	**0.357**	0.160
UBS	UBS	0.140	0.116	0.306	0.244	0.241	0.140	0.130	0.160	0.116	0.097	0.219	0.176

the results are nearly perfect and not very interesting for analysis. Accordingly, their results are not included here. On the other hand, HYDICE data set provides precise knowledge of panel pixels in Fig. 1(b). Such ground truth can be used for verification and validation. To the authors' best knowledge, this scene may be one of the best data sets that can be used to evaluate the effectiveness of target detection.

7) Finally, a remark on the data sets used in this paper is noteworthy. Although there are many real data sets that can be used for experiments, a major issue in using these data sets is no availability of complete ground truth for fact validation. For example, three well-known data sets, Purdue's Indian Pines, Salinas, and University of Pavia available on the website http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes have been widely used for hyperspectral image classification but not target detection. As pointed out above, the BKG suppression is a serious issue in target detection. It is our belief that many targets of interest in these three scenes are indeed in the BKG [31] which cannot be provided by ground truth. So, using these image scenes for target detection is not appropriate.

IX. CONCLUSION

This paper develops a new approach to BS for multiple-target detection, called CTBS. It first specifies a target signature of interest, $\mathbf{d}$ and then constrains $\mathbf{d}$ to minimize the variance caused by other signal sources via an FIR filter. The resulting variance is further used as a measure to select bands. Two types of CTBS can be derived. One is to use the CEM-produced variance to design two CTBS-based BP criteria, called MinV-BP and MaxV-BP. The other is to use the CEM-produced variance to derive two band-searching algorithms to select bands, called SF-CTBS and SB-CTBS. Since CTBS is developed for multiple target detection, a different target may result in a different set of selected bands. In order to fuse different sets of bands selected by multiple targets, a novel idea is also further developed to fuse bands selected by different targets so that a fixed set of selected bands can be found via the proposed BFS method to work for all targets for multiple-target detection. Despite that CTBS also takes advantage of the idea of CEM, it is indeed quite different from the CEM-based BS, CBS [5] in several perspectives which can be considered as main contributions of this paper. One is that CTBS is designed to constrain targets of interest using partial band subsets compared to CBS which only constrains bands of interest. As a result, CTBS takes care of both targets of interest and selected bands, whereas CBS only deals with bands. Another is that the sample correlation matrix $\mathbf{R}$ used by CTBS varies with selected bands, Ω_{BS} as opposed to the $\mathbf{R}$ used by CBS which is fixed. A third one is that the computational cost required by CTBS is determined by number of selected bands, n_{BS}, while the computing time required by CBS is

determined by the size of the hyperspectral image cube to be processed. As a result, the computing time for CBS is significantly increased as an image size grows. A fourth one is that CTBS provides crucial information about selected bands that are significant in target detection, which CBS cannot. A fifth one is that the band number, n_{BS} and the set of selected bands, Ω_{BS} selected by CTBS vary with different desired target signatures determined by specific applications, whereas CBS produces the same set of selected bands regardless of applications. Accordingly, a new fusion BS, BFS, is also developed for CTBS, which is not needed for CBS. Finally, the most and foremost is the derivation of Theorem 1 for CTBS which does not exist in CBS but is vital to CTBS.

To conclude this paper, two remarks are particularly noteworthy.

1) One is to extend CTBS for multiple-target detection to multiple-class classification. In fact, an idea similar to constraining multiple targets can also be investigated for constraining multiple classes for classification [45]. However, such an extension is by no means trivial. This is because CEM only works for target detection and is no longer applicable to multiclass classification. In this case, LCMV developed in [22], [23], and [25] must be used for this purpose. The work in [45] can be considered as a companion paper to this paper to form a nice pair of constrained-target detection and constrained-multiple-class classification approaches to BS. However, it should be noted that there are several key differences between [45] and this paper. First of all, CTBS uses CEM as a detector to select different band subsets for each of targets of interest in target detection, while reference [45] uses LCMV as a classifier to select bands for all classes in classification. As a result, CTBS requires BFS to fuse band subsets produced by different targets, while reference [45] does not since it produces its band subset for all classes. This is because CEM can detect one target at a time so as to achieve multiple-target detection. In addition, Theorem 1 derived in this paper shows that the work in [26] is its special case but is different from [45, Th. 1]. Finally, the experiments are completely different. The data sets used in [45] cannot be used for target detection.
2) In a future development of hyperspectral imaging in space BS along with progressive/real-time hyperspectral image processing will play a major role in hyperspectral data communication and transmission. The CTBS developed in this paper can be implemented in conjunction with recently developed hyperspectral imaging techniques in [27] and [30] for these applications.

APPENDIX

Proof of Theorem 1: For simplicity of notation, we assume that $\Omega_k = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k\}$ without loss of generality. Also let the total number of data sample vectors be $\{\mathbf{r}_i^{(k)}\}_{i=1}^N$ using k bands with $\mathbf{r}_i^{(k)} = (r_{i1}^{(k)}, r_{i2}^{(k)}, \ldots, r_{ik}^{(k)})^T$. Let the data matrix be denoted by

$$\mathbf{X}_k = [\mathbf{r}_1^{(k)}\mathbf{r}_2^{(k)}\ldots\mathbf{r}_N^{(k)}] = \begin{bmatrix} r_{11} & r_{21} & \cdots & r_{N1} \\ r_{12} & r_{22} & \cdots & r_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1k} & r_{2k} & \cdots & r_{Nk} \end{bmatrix}$$

using k bands. Then, the data matrix using $(k+1)$ bands is given by

$$\mathbf{X}_{k+1} = \begin{bmatrix} r_{11} & r_{21} & \cdots & r_{(N-1)1} & r_{N1} \\ r_{12} & r_{22} & \cdots & r_{(N-1)2} & r_{N2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r_{1k} & r_{2k} & \cdots & r_{(N-1)k} & r_{Nk} \\ r_{1(k+1)} & r_{2(k+1)} & \cdots & r_{(N-1)(k+1)} & r_{N(k+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_k \\ \mathbf{r}^T(k+1) \end{bmatrix} \tag{A1}$$

where $\mathbf{r}(k+1) = (r_{1(k+1)}, r_{2(k+1)}, \ldots, r_{N(k+1)})^T$. Let $\mathbf{R}_{\Omega_k} = (1/N)\sum_{i=1}^N (\mathbf{r}_i^{(k)}(\mathbf{r}_i^{(k)})^T) = (1/N)\mathbf{X}_k\mathbf{X}_k^T$. We have

$$\begin{aligned} &\mathbf{R}_{\Omega_k\cup\{\mathbf{b}_{k+1}\}} \\ &= \frac{1}{N}\begin{bmatrix} \mathbf{X}_k \\ \mathbf{r}^T(k+1) \end{bmatrix}\begin{bmatrix} \mathbf{X}_k^T & \mathbf{r}(k+1) \end{bmatrix} \\ &= \frac{1}{N}\begin{bmatrix} \mathbf{X}_k\mathbf{X}_k^T & \mathbf{X}_k\mathbf{r}(k+1) \\ \mathbf{r}^T(k+1)\mathbf{X}_k^T & \mathbf{r}^T(k+1)\mathbf{r}(k+1) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{R}_{\Omega_k} & (1/N)\mathbf{X}_k\mathbf{r}(k+1) \\ (1/N)\mathbf{r}^T(k+1)\mathbf{X}_k^T & (1/N)\mathbf{r}^T(k+1)\mathbf{r}(k+1) \end{bmatrix}. \end{aligned} \tag{A2}$$

Using the following matrix identity in [1] and [39]:

$$\begin{aligned} &\begin{bmatrix} \mathbf{U}^T\mathbf{U} & \mathbf{U}^T\mathbf{d} \\ \mathbf{d}^T\mathbf{U} & \mathbf{d}^T\mathbf{d} \end{bmatrix}^{-1} \\ &\qquad = \begin{bmatrix} (\mathbf{U}^T\mathbf{U})^{-1} + \beta\mathbf{U}^{\#}\mathbf{d}\mathbf{d}^T(\mathbf{U}^{\#})^T & -\beta\mathbf{U}^{\#}\mathbf{d} \\ -\beta\mathbf{d}^T(\mathbf{U}^{\#})^T & \beta \end{bmatrix} \end{aligned} \tag{A3}$$

$$\mathbf{R}_{\Omega_k\cup\{\mathbf{b}_{k+1}\}}^{-1} = \frac{1}{N}\begin{bmatrix} N\mathbf{R}_{\Omega_k} & \mathbf{X}_k\mathbf{r}(k+1) \\ \mathbf{r}^T(k+1)\mathbf{X}_k^T & \mathbf{r}^T(k+1)\mathbf{r}(k+1) \end{bmatrix}^{-1} = \frac{1}{N}\begin{bmatrix} N^{-1}\mathbf{R}_{\Omega_k}^{-1} + \beta\left(\mathbf{X}_k^T\right)^{\#}\mathbf{r}(k+1)\mathbf{r}^T(k+1)\mathbf{X}_k^{\#} & \left(\mathbf{X}_k^T\right)^{\#}\mathbf{r}(k+1) \\ \mathbf{r}^T(k+1)\mathbf{X}_k^{\#} & \beta \end{bmatrix} \tag{A4}$$

$$\begin{aligned} N^{-1}\mathbf{d}_{\Omega_k\cup\{\mathbf{b}_{k+1}\}}^T\mathbf{R}_{\Omega_k\cup\{\mathbf{b}_{k+1}\}}^{-1}\mathbf{d}_{\Omega_k\cup\{\mathbf{b}_{k+1}\}} &= \begin{bmatrix} \mathbf{d}_k^T & d_{k+1} \end{bmatrix}\begin{bmatrix} N^{-1}\mathbf{R}_{\Omega_k}^{-1} + \beta\left(\mathbf{X}_k^T\right)^{\#}\mathbf{r}(k+1)\mathbf{r}^T(k+1)\mathbf{X}_k^{\#} & -\beta\left(\mathbf{X}_k^T\right)^{\#}\mathbf{r}(k+1) \\ -\beta\mathbf{r}^T(k+1)\mathbf{X}_k^{\#} & \beta \end{bmatrix}\begin{bmatrix} \mathbf{d}_k \\ d_{k+1} \end{bmatrix} \\ &= N^{-1}\mathbf{d}_k^T\mathbf{R}_{\Omega_k}^{-1}\mathbf{d}_k + \beta\mathbf{d}_k^T\left(\mathbf{X}_k^T\right)^{\#}\mathbf{r}(k+1)\mathbf{r}^T(k+1)\mathbf{X}_k^{\#}\mathbf{d}_k - 2\beta d_{k+1}\mathbf{r}^T(k+1)\mathbf{X}_k^{\#}\mathbf{d}_k + d_{k+1}^2\beta \\ &= N^{-1}\mathbf{d}_k^T\mathbf{R}_{\Omega_k}^{-1}\mathbf{d}_k + \beta\left[\mathbf{d}_k^T\left(\mathbf{X}_k^T\right)^{\#}\mathbf{r}(k+1) - d_{k+1}\right]^2 > N^{-1}\mathbf{d}_k^T\mathbf{R}_{\Omega_k}^{-1}\mathbf{d}_k \end{aligned} \tag{A5}$$

with $\mathbf{U} = \mathbf{X}_k^T$ and $\mathbf{d} = \mathbf{r}(k+1)$, we can derive, (A4), as shown at the bottom of the previous page, where $(\mathbf{X}_k^T)^{\#} = (\mathbf{X}_k\mathbf{X}_k^T)^{-1}\mathbf{X}_k^T$ and $\beta = (\mathbf{r}^T(k+1)(\mathbf{I}-\mathbf{X}_k^T(\mathbf{X}_k^T)^{\#})\mathbf{r}(k+1))^{-1} = (\mathbf{r}^T(k+1)P_{\mathbf{X}_k^T}^{\perp}\mathbf{r}(k+1))^{-1}$ due to the fact that $P_{\mathbf{X}_k^T}^{\perp} = \mathbf{I} - \mathbf{X}_k^T(\mathbf{X}_k^T)^{\#}$ is idempotent, i.e., $(P_{\mathbf{X}_k^T}^{\perp})^2 = P_{\mathbf{X}_k^T}^{\perp}$.

Using (A4) and letting $\mathbf{d}_{\Omega_k\cup\{\mathbf{b}_{k+1}\}}^T = \begin{bmatrix}\mathbf{d}_k \\ d_{k+1}\end{bmatrix}$ and $\mathbf{d}_k = (d_1, d_2, \ldots, d_k)^T$, we can derive, (A5), as shown at the bottom of the previous page. Since $[\mathbf{d}_k^T(\mathbf{X}_k^T)^{\#}\mathbf{r}(k+1) - d_{k+1}]^2 > 0$ and $\beta > 0$, the second term, $\beta[\mathbf{d}_k^T(\mathbf{X}_k^T)^{\#}\mathbf{r}(k+1) - d_{k+1}]^2 > 0$. Therefore, (14) is proven. **QED**

REFERENCES

[1] C.-I Chang, *Hyperspectral Data Processing: Algorithm Design and Analysis*. Hoboken, NJ, USA: Wiley, 2013.

[2] C.-I Chang, Q. Du, T.-L. Sun, and M. L. G. Althouse, "A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 6, pp. 2631–2641, Nov. 1999.

[3] 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.

[4] A. Martínez-Usó, F. Pla, J. M. Sotoca, P. García-Sevilla, "Clustering-based hyperspectral band selection using information measures," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 4158–4171, Dec. 2007.

[5] H. Su, H. Yang, Q. Du, and Y. Sheng, "Semisupervised band clustering for dimensionality reduction of hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 6, pp. 1135–1139, Nov. 2011.

[6] J. Feng, L. C. Jiao, X. Zhang, and T. Sun, "Hyperspectral band selection based on trivariate mutual information and clonal selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4092–4105, Jul. 2014.

[7] 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.

[8] H. Yang, Q. Du, H. Su, and Y. Sheng, "An efficient method for supervised hyperspectral band selection," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 1, pp. 138–142, Jan. 2011.

[9] H. Su, Q. Du, G. Chen, and P. Du, "Optimized hyperspectral band selection using particle swarm optimization," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2659–2670, Jun. 2014.

[10] H. Su, B. Yong, and Q. Du, "Hyperspectral band selection using improved firefly algorithm," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 68–72, Jan. 2016.

[11] Y. Yuan, G. Zhu, and Q. Wang, "Hyperspectral band selection by multitask sparsity pursuit," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 631–644, Feb. 2015.

[12] Y. Yuan, X. Zheng, and X. Lu, "Discovering diverse subset for unsupervised hyperspectral band selection," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 51–64, Jan. 2017.

[13] C. Wang, M. Gong, M. Zhang, and Y. Chan, "Unsupervised hyperspectral image band selection via column subset selection," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 7, pp. 1411–1415, Jul. 2015.

[14] A. Zare and P. Gader, "Hyperspectral band selection and endmember detection using sparsity promoting priors," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 256–260, Apr. 2008.

[15] K. Sun, X. Geng, L. Ji, and Y. Lu, "A new band selection method for hyperspectral image based on data quality," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2697–2703, Jun. 2014.

[16] X. Geng, K. Sun, and L. Ji, "Band selection for target detection in hyperspectral imagery using sparse CEM," *Remote Sens. Lett.*, vol. 5, no. 3, pp. 1022–1031, 2014.

[17] K. Sun, X. Geng, and L. Ji, "A new sparsity-based band selection method for target detection of hyperspectral image," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 329–333, Feb. 2015.

[18] Y. Xu, Q. Du, and N. H. Younan, "Particle swarm optimization-based band selection for hyperspectral target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 554–558, Apr. 2017.

[19] M. A. Hossain, X. Jia, and J. A. Benediktsson, "One-class oriented feature selection and classification of heterogeneous remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Ses.*, vol. 9, no. 4, pp. 1606–1612, Apr. 2016.

[20] J. C. Harsanyi, "Detection and classification of subpixel spectral signatures in hyperspectral image sequences," Dept. Elect. Eng., Univ. Maryland, College Park, MD, USA, Aug. 1993.

[21] W. H. Farrand and J. C. Harsanyi, "Mapping the distribution of mine tailings in the Coeur d'Alene river valley, Idaho, through the use of a constrained energy minimization technique," *Remote Sens. Environ.*, vol. 59, no. 1, pp. 64–76, Jan. 1997.

[22] 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.

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

[24] H. Ren and C.-I Chang, "Target-constrained interference-minimized approach to subpixel target detection for hyperspectral images," *Opt. Eng.*, vol. 39, no. 12, pp. 3138–3145, Dec. 2000.

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

[26] X. Geng, L. Ji, K. Sun, and Y. Zhao, "CEM: More bands, better performance," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 11, pp. 1876–1880, Nov. 2014.

[27] C.-I Chang, *Real-Time Progressive Hyperspectral Image Processing: Endmember Finding and Anomaly Detection*. New York, NY, USA: Springer, 2016.

[28] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognit. Lett.*, vol. 15, no. 11, pp. 1119–1125, 1994.

[29] C.-I Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.

[30] C.-I Chang, *Real-Time Recursive Hyperspectral Sample and Band Processing: Algorithm Architecture and Implementation*. New York, NY, USA: Springer, 2017.

[31] C.-I Chang, "A review of virtual dimensionality for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1285–1305, Apr. 2018.

[32] C.-I Chang, R. C. Schultz, M. C. Hobbs, S.-Y. Chen, Y. Wang, and C. Liu, "Progressive band processing of constrained energy minimization for subpixel detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1626–1637, Mar. 2015.

[33] J. C. Harsanyi, W. Farrand, and C.-I Chang, "Detection of subpixel spectral signatures in hyperspectral image sequences," in *Proc. Annu. Meeting, Proc. Amer. Soc. Photogram. Remote Sens.*, Reno, NV, USA, 1994, pp. 236–247.

[34] C.-I Chang, X. Jiao, Y. Du, C.-C. Wu, Y. Du, and M.-L. Chang, "A review of unsupervised spectral target analysis for hyperspectral imagery," *EURASIP J. Adv. Signal Process.*, vol. 2010, Apr. 2010, Art. no. 503752.

[35] C.-I Chang, X. Jiao, E. Y. Du, C. C. Wu, and H.-M. Chen, "Component analysis-based unsupervised linear spectral mixture analysis for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4123–4137, Nov. 2011.

[36] C.-I Chang, H. Ren, Q. Du, S.-S. Chiang, and A. Ifarragaerri, "An ROC analysis for subpixel detection," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Sydney, NSW, Australia, Jul. 2001, pp. 2355–2357.

[37] C.-I Chang, "Multiparameter receiver operating characteristic analysis for signal detection and classification," *IEEE Sensors J.*, vol. 10, no. 3, pp. 423–442, Mar. 2010.

[38] C.-I Chang, H.-C. Li, M. Song, C. Liu, and L. Zhang, "Real-time constrained energy minimization for subpixel detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2545–2559, Jun. 2015.

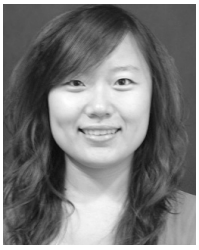
[39] C.-I Chang, *Real-Time Recursive Hyperspectral Sample and Band Processing: Algorithm Architecture and Implementation*. New York, NY, USA: Springer, 2017.

[40] C.-I Chang, Y. Li, M. C. Hobbs, R. C. Schultz, and W.-M. Liu, "Progressive band processing of anomaly detection in hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 7, pp. 3558–3571, Jul. 2015.

[41] S.-Y. Chen, Y. Wang, C. C. Wu, C. Liu, and C.-I Chang, "Real-time causal processing of anomaly detection for hyperspectral imagery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 1511–1534, Apr. 2014.

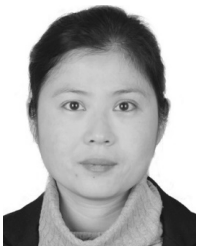
[42] L. Zhang *et al.*, "Fast real-time causal linewise progressive hyperspectral anomaly detection via Cholesky decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 10, pp. 4614–4629, Oct. 2017.

[43] Y.-C. C. Chang, H. Ren, C.-I Chang, and R. S. Rand, "How to design synthetic images to validate and evaluate hyperspectral imaging algorithms," *Proc. SPIE*, vol. 6966, Apr. 2008, Art. no. 69661P.
[44] C. C. Wu, C. S. Lo, and C.-I Chang, "Improved process for use of a simplex growing algorithm for endmember extraction," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 523–527, Jul. 2009.
[45] C. Yu, M. Song, Y. Wang, and C.-I Chang, "Class signature-constrained background-suppressed approach to band selection for classification of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 14–31, Jan. 2019.



Yulei Wang received the B.S. and Ph.D. degrees in signal and information processing, Harbin Engineering University, Harbin, China, in 2009 and 2015, respectively.

She is awarded by China Scholarship Council in 2011 as a joint Ph.D. Student to study in Remote Sensing Signal and Image Processing Laboratory, University of Maryland, Baltimore, MD, USA. She is currently a Lecturer with the College of Information Science and Technology, Dalian Maritime University, Dalian, China. Her research interests include hyperspectral image processing and vital sign signal processing.



Lin Wang received the B.S. degree in optoelectronic technology and the M.S. and Ph.D. degrees in physical electronics from Xidian University, Xi'an, China, in 2000, 2013, and 2018, respectively.

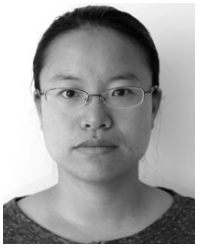
From 2015 to 2016, she was a Visiting Research Associate Scholar with the Remote Sensing Signal and Image Processing Laboratory, University of Maryland, Baltimore, MD, USA. She is currently an Associate Professor with the School of Physics and Optoelectronic Engineering, Xidian University. Her research interests include hyperspectral image processing, automatic target recognition, and real-time image processing.

Chunyan Yu (M'17) received the B.Sc. and Ph.D. degrees in environment engineering from Dalian Maritime University, Dalian, China, in 2004 and 2012, respectively.

In 2004, she joined the College of Computer Science and Technology, Dalian Maritime University. From 2013 to 2016, she was a Post-Doctoral Fellow with the Information Science and Technology College, Dalian Maritime University. From 2014 to 2015, she was a Visiting Scholar with the College of Physicians and Surgeons, Columbia University, New York, NY, USA. She is currently an Associate Professor with the Information Science and Technology College, Dalian Maritime University. Her research interests include image segmentation, hyperspectral image classification, and pattern recognition.

Enyu Zhao received the Ph.D. degree from the College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China, in 2017, and the joint Ph.D. degree in cartography and geographic information system from the University of Chinese Academy of Sciences, Beijing, China, and the Engineering Science, Computer Science and Imaging Laboratory, University of Strasbourg, Strasbourg, France.

He is currently a Lecturer with the College of Information Science and Technology, Dalian Maritime University, Dalian, China. His research interests include remote sensing and hyperspectral image processing.



Meiping Song received the Ph.D. degree from the College of Computer Science and Technology, Harbin Engineering University, Harbin, China, in 2006.

From 2013 to 2014, she was a Visiting Associate Research Scholar at the University of Maryland, College Park, MD, USA. She is currently an Associate Professor with the College of Information Science and Technology, Dalian Maritime University, Dalian, China. Her research interests include remote sensing and hyperspectral image processing.

Chia-Hsien Wen received the B.S. degree in computer science from Tamkang College, Taipei, Taiwan, in 1976, and the M.S. degree in applied mathematics and the Ph.D. degree in computer science from National Tsing Hua University, Hsinchu, Taiwan, in 1978 and 1994, respectively.

He was the Director of the Computing Center at Taichung Veterans General Hospital, Taichung, Taiwan. He is currently a Professor with the Department of Computer Science and Information Management and the Executive Vice President with Providence University, Taichung. His research interests include medical image processing, medical informatics, database management, and machine learning.

Chein-I Chang (S'81–M'87–SM'92–F'10–LF'17) received the B.S. degree in mathematics from Soochow University, Taipei, Taiwan, the M.S. degree in mathematics from the Institute of Mathematics, National Tsing Hua University, Hsinchu, Taiwan, the M.A. degree in mathematics from the State University of New York at Stony Brook, Stony Brook, NY, USA, the M.S. and M.S.E.E. degrees from the University of Illinois at Urbana–Champaign, Urbana, IL, USA, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, MD, USA.

He has been with the University of Maryland, Baltimore County (UMBC), Baltimore, MD, USA, since 1987. He was a Visiting Research Specialist with the Institute of Information Engineering, National Cheng Kung University, Tainan, Taiwan, from 1994 to 1995, and also a Distinguished Visiting Fellow/Fellow Professor, both of which were sponsored by National Science Council in Taiwan from 2009 to 2010. He was a Distinguished Lecturer Chair with National Chung Hsing University, Taichung, Taiwan, sponsored by the Ministry of Education in Taiwan from 2005 to 2006, where he was a Distinguished Chair Professor from 2014 to 2017. He has been an adjunct Chair Professor with Providence University, Taichung, Taiwan, since 2012. He has been a Hua Shan Scholar Chair Professor with Xidian University, Xian, China, since 2016. He has been a Feng-Tai Chair Professor with the National Yunlin University of Science and Technology, Douliu, Taiwan, since 2017. He is currently holding Chang Jiang Scholar Chair Professorship and the Director of the Center for Hyperspectral Imaging in Remote Sensing with Dalian Maritime University, Dalian, China. He is currently a Professor with the Department of Computer Science and Electrical Engineering, UMBC. He has authored four books, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification* published by Kluwer Academic Publishers in 2003, *Hyperspectral Data Processing: Algorithm Design and Analysis* by John Wiley & Sons in 2013, *Real Time Progressive Hyperspectral Image Processing: Endmember Finding and Anomaly Detection* by Springer in 2016, and *Recursive Hyperspectral Sample and Band Processing: Algorithm Architecture and Implementation* by Springer in 2017. In addition, he also edited two books, *Recent Advances in Hyperspectral Signal and Image Processing* in 2006 and *Hyperspectral Data Exploitation: Theory and Applications*, John Wiley & Sons, in 2007, and co-edited with A. Plaza a book on *High Performance Computing in Remote Sensing,* CRC Press, in 2007. He holds seven patents on hyperspectral image processing. His research interests include multispectral/hyperspectral image processing, automatic target recognition, and medical imaging.

Dr. Chang is a fellow of SPIE. He received the National Research Council Senior Research Associateship Award from 2002 to 2003 sponsored by the U.S. Army Soldier and Biological Chemical Command, Edgewood Chemical and Biological Center, Aberdeen Proving Ground, MD, USA. He was a Plenary Speaker of the Society for Photo-optical Instrumentation Engineers Optics+Applications, Remote Sensing Symposium in 2009. He was also a Keynote Speaker at the User Conference of Hyperspectral Imaging 2010, in 2010, Industrial Technology Research Institute, Hsinchu; 2009 Annual Meeting of the Radiological Society of the Republic of China in 2009, Taichung; 2008 International Symposium on Spectral Sensing Research in 2008; and the Conference on Computer Vision, Graphics, and Image Processing 2003, Kimen and 2013, Nan-Tou, Taiwan. He was the Guest Editor of a special issue of the *Journal of High Speed Networks* on *Telemedicine and Applications* in 2000 and a Co-Guest Editor of another special issue of the same journal on *Broadband Multimedia Sensor Networks in Healthcare Applications* in 2007. He is also the Co-Guest Editor of special issues on *High Performance Computing of Hyperspectral Imaging* for the *International Journal of High Performance Computing Applications* in 2007, *Signal Processing and System Design in Health Care Applications* for the *EURASIP Journal on Advances in Signal Processing* in 2009, *Multipsectral, Hyperspectral, and Polarimetric Imaging Technology* for the *Journal of Sensors* in 2016, and *Hyperspectral Imaging and Applications* for the *Remote Sensing* in 2018.

He was an Associate Editor in the area of hyperspectral signal processing for the IEEE TRANSACTION ON GEOSCIENCE AND REMOTE SENSING from 2001 to 2007. He is currently an Associate Editor of the *Artificial Intelligence Research* and also on the editorial boards of the *Journal of High Speed Networks*, the *Recent Patents on Mechanical Engineering*, the *International Journal of Computational Sciences and Engineering*, the *Journal of Robotics and Artificial Intelligence*, and the *Open Remote Sensing Journal*.