

Band Subset Selection for Anomaly Detection in Hyperspectral Imagery

Lin Wang, Chein-I Chang, *Life Fellow, IEEE*, Li-Chien Lee, Yulei Wang, Bai Xue, *Student Member, IEEE*, Meiping Song, Chuanyan Yu, and Sen Li

Abstract—This paper presents a new approach, called band subset selection (BSS)-based hyperspectral anomaly detection (AD), which selects multiple bands simultaneously as a band subset rather than selecting multiple bands one at a time as the tradition band selection (BS) does, referred to as sequential multiple BS (SQMBS). Its idea is to first use virtual dimensionality (VD) to determine the number of multiple bands, n_{BS} needed to be selected as a band subset and then develop two iterative process, sequential BSS (SQ-BSS) algorithm and successive BSS (SC-BSS) algorithm to find an optimal band subset numerically among all possible n_{BS} combinations out of the full band set. In order to terminate the search process the averaged least-squares error (ALSE) and 3-D receiver operating characteristic (3D ROC) curves are used as stopping criteria to evaluate performance relative to AD using the full band set. Experimental results demonstrate that BSS generally performs better background suppression while maintaining target detection capability compared to target detection using full band information.

Index Terms—3-D receiver operating characteristic (ROC) analysis, averaged least-squares error (ALSE), band selection (BS), band subset selection (BSS), dimensionality reduction (DR), sequential BSS (SQ-BSS), sequential multiple BS (SQMBS), simultaneous MBS (SMMBS), single BS (SBS), successive BSS (SC-BSS).

Manuscript received January 1, 2017; accepted March 6, 2017. Date of publication June 12, 2017; date of current version August 25, 2017. The work of L. Wang was supported by the 111 Project (B17035). The work of C.-I. Chang was supported by the Fundamental Research Funds for Central Universities under Grant 3132016331. The work of M. Song was supported by the National Nature Science Foundation of China under Grant 61601077, Fundamental Research Funds for the Central Universities (3132017124) and State Key Laboratory of Integrated Services Networks. The work of Y. Wang was supported in part by the Fundamental Research Funds for Central Universities under Grant 3132017080 and State Key Laboratory of Integrated Services Networks. (*Corresponding author: Chein-I Chang.*)

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

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

L.-C. Lee and B. Xue are with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD 21250 USA (e-mail: baixuel@umbc.edu).

Y. Wang, M. Song, C. Yu, and S. Li are with the Information and Technology College, Dalian Maritime University, Dalian, China (e-mail: wangyulei.heu@gmail.com; smping@163.com; 36202986@qq.com; listen@dlmu.edu.cn).

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.2017.2681278

0196-2892 © 2017 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

I. INTRODUCTION

HYPERSPECTRAL imaging sensors use hundreds of contiguous spectral channels to reveal subtle material substances in the data. As a consequence, hyperspectral imagery generally has enormous data volume and contains vast amount of spectral information which is also expected to be highly correlated among bands. Two common practices are generally taken [1]. One is data dimensionality reduction (DR) [1] which compacts data in a lower dimensional space via various transforms. The other is band selection (BS) [1]–[33] which selects appropriate bands from the original set of spectral bands that can well represent original data, while discarding all unselected bands. Compared to DR which transforms data, BS has an advantage of preserving original information from the data. This paper is mainly focused on BS and presents a new simultaneous multiple BS (SMMBS) approach to BS, to be called band subset selection (BSS), which selects multiple bands simultaneously as a band set instead of other MBS, referred to as sequential MBS (SQMBS), which selects one single band at a time sequentially via the conventional single band selection (SBS). Therefore, SBS can be considered as a special case of BSS where the band subset used by BSS is simply a singleton set.

There are several crucial differences between SQMBS and BSS. First and foremost is how multiple bands are selected. BSS is designed to select multiple bands altogether at once simultaneously, while SQMBS selects multiple bands one at a time sequentially. As a result, a second major difference is that SQMBS is generally performed by SBS which requires band prioritization (BP) to rank individual single bands as well as band decorrelation (BD) to remove highly correlated bands, whereas BSS considers selection of multiple bands as a whole in which case it already takes care of BP and BD. A third important difference is that SBS used to perform SQMBS usually takes advantage of data statistics and properties as BP criteria such as variance, signal-to-noise ratio (SNR), entropy, and so on to prioritize bands [2], [3]. Therefore, SBS has nothing to do with applications. That is, once bands are selected, the selected bands will be used for all different applications. By contrast, BSS is primarily determined by various applications and thus, different applications select different sets of bands. A fourth difference is that since SQMBS selects bands sequentially, it requires BD to decorrelate with previously selected bands to avoid selecting redundant bands. However, this leads to a

challenging issue in how to select an appropriate threshold to determine BD for which BSS does not have such issue. A fifth difference is that many SQMBS applications are supervised such as classification using training samples to generate features for BP [4], [7], [14]–[20], whereas BSS is completely unsupervised. Finally and most importantly, BSS and SQMBS are completely different problems; thus, they require different approaches. Specifically, there is an issue arising in BSS which does not exist in SBS. That is how to effectively select simultaneous multiple bands as an optimal band subset. Interestingly, this issue is similar to endmember extraction. If we interpret each data sample vector as an individual single band, then the issue of extracting a set of endmembers from the entire data space is same as the issue of finding an optimal set of multiple bands from the full band set. With this interpretation, the well-known N-finder algorithm (N-FINDR) developed by Winter [34] is an excellent candidate to be used for BSS. Since N-FINDR needs to conduct an exhaustive search for all possible endmember sets, two sequential versions of N-FINDR, sequential N-FINDR (SQ N-FINDR), and successive N-FINDR (SC N-FINDR) were developed for this purpose in [1] and [35]–[37]. By taking advantage of these two algorithms, we can also develop their counterparts for BSS, to be called sequential BSS (SQ-BSS) algorithm and successive BSS (SC-BSS) algorithm, respectively, for finding an optimal BSS to avoid an exhaustive search for all possible band subsets.

It is known that BS is generally performed in either an unsupervised or a supervised manner. When it comes to unsupervised BS, the bands to be selected are usually determined by data characteristics or statistics such as variance, SNR, entropy, and information divergence (ID) [2], [3]. As a result, the selected bands are basically independent of applications. When it comes to applications, BS methods are generally supervised and require training samples to generate BS features such as classification features in [14]–[16], [23], [24], and [26]–[29], target detection [19], [20], endmember extraction [22], and spectral unmixing [31]. Interestingly, BS for anomaly detection (AD) has not received much interest [5], [24]. This may be due to the fact that AD is unsupervised, and there are no training samples that can be used to find anomaly features to select bands. Most importantly, according to a recent study [38], background suppression is a very important and crucial factor in effectiveness of AD. This is because without prior knowledge as ground truth AD is generally evaluated by visual inspection where background suppression plays a key element in assessing performance of AD. Specifically, a better background suppression can bring up weak anomalies which could be overwhelmed and compromised by other stronger anomalies, while also reducing falsely alarmed targets. This paper takes up this issue to explore two major anomaly detectors since many currently being used anomaly detectors are variants of one of these two anomaly detectors. One is developed by Reed and Yu [39], referred to as K-anomaly detector (K-AD) with \mathbf{K} specified by a sample covariance matrix \mathbf{K} . The other is the R-anomaly detector (R-AD) with \mathbf{R} specified by a sample correlation matrix [40]. Although K-AD and R-AD using SBS have been

also studied in the literature, this paper extends the work in [41] which is believed to be the first work using BSS to select multiple bands simultaneously for K-AD. In particular, in order to evaluate the issue of background suppression, a 3-D receiver operating characteristic (3-D ROC) analysis developed in [1] and [42] is further used for background suppression performance analysis for K-AD as well as R-AD, both of which are not found [41].

II. BAND SUBSET SELECTION

Since different spectral bands provide different levels of the information of interest, the primary goal of BS is to select an appropriate band subset from the original band set to represent the original data in some sense of optimality. Therefore, the information preserved by BS has significant impact on data analysis because the information of unselected bands will be completely discarded after BS. So, a key success in BS is how to design effective criteria for BS to meet various applications.

Solving a general BS problem generally requires an exhaustive search for all possible Ω_{BS} -combinations out of the total number of spectral bands, L in Ω where $|\Omega_{\text{BS}}|$ is the number of bands to be selected in Ω_{BS} .

More specifically, assume that $J(\cdot)$ is a generic objective function of Ω_{BS} for BS to be optimized. For a given number of selected bands, n_{BS} , a BS technique is to find an optimal band subset, Ω_{BS}^* with $|\Omega_{\text{BS}}^*| = n_{\text{BS}}$ which satisfies the following optimization problem:

$$\Omega_{\text{BS}}^* = \arg\{\max/\min_{\Omega_{\text{BS}} \subset \Omega, |\Omega_{\text{BS}}| = n_{\text{BS}}} J(\Omega_{\text{BS}})\}. \quad (1)$$

Depending upon how the objective function $J(\Omega_{\text{BS}})$ is designed, the optimization in (1) can be performed by either maximization or minimization over all possible band subsets Ω_{BS} in Ω with $|\Omega_{\text{BS}}| = n_{\text{BS}}$.

Over the past years, many BS techniques have been investigated by designing various criteria or features to define $J(\Omega_{\text{BS}})$ in (1). In what follows, we describe a rather different BS technique, which is based on the concept of finding endmembers.

A. Interpretation of BSS as Endmember Finding

Finding endmembers has received considerable interests in recent years [1] since an endmember is assumed to have purity signature present in the data that can be used to specify a particular spectral class. Let E_p be a set of p endmembers to be found in the data set S with total number of data sample vectors denoted by N , i.e., $|S| = N$. Then finding an optimal set of p endmembers requires exhausting all possible p -combinations out of N data sample vectors. For example, if N-FINDR [34] is used as a desired endmember finding algorithm (EFA), the objective function in (1) can be interpreted as finding the maximal volume of simplexes embedded in the data set S . In this case, (1) can be expressed as

$$E_p^{N\text{-FINDR}} = \arg\{\max_{E_p \subset S} \text{SV}_{\text{embedded}}(E_p)\} \quad (2)$$

where $\text{SV}_{\text{embedded}}$ is defined as the volume of a simplex formed by E_p embedded in the data set S . Comparing (1) to (2)

immediately realizes that finding an optimal set of p bands by (1) is similar to finding an optimal set of p endmembers. With this interpretation, the well-established theory of N-FINDR developed in [1] is readily applied to BS provided that $J(\Omega_{BS})$ in (1) can be defined appropriately.

B. Band Subset Selection Algorithms

In order to take advantage of the N-FINDR theory to solve the BS problem, we need to appropriately define the $J(\Omega_{BS})$ in (1). This can be done by specifying a particular application. In other words, we define η_A as an application-based performance measure to replace $J(\Omega_{BS})$ in (1) for BS where the subscript of η_A , “A” is used to specify a particular application. For example, if an application is specified by AD, the subscript A in η_A will be specified by AD or a particular algorithm used to implement AD. Then performance measure, η_A can be the area under ROC curve, denoted by area under curve (AUC) for evaluation of detection performance. As another example, if an application is specified by finding endmembers, the subscript “A” in η_A will be replaced by simplex volume (SV) as defined in (2). However, when there is no ground truth available for performance evaluation, a more general performance criterion is to measure the difference such as least-squares error (LSE), denoted by $\eta_{LSE}(\Omega_{BS}; \Omega)$ between the results produced by full bands and the results by selected bands.

Using η_A as a general performance measure for BS, we develop two numerical search algorithms to select BSS simultaneously from all possible n_{BS} -combinations out of L bands to avoid an exhaustive search. Since N-FINDR suffers from the same issue of an exhaustive search for (2) as BS does for (1), two sequential versions of N-FINDR developed in [1], called SQ N-FINDR and SC N-FINDR, were developed to avoiding conducting an exhaustive search for an optimal p -endmember set E_p . By interpreting finding p endmembers as finding an optimal p -band subset SQ N-FINDR and SC N-FINDR can be modified and rederived for finding optimal set of bands to be selected.

1) *Sequential Band Subset Selection Algorithm*: The first algorithm is derived from SQ N-FINDR and called SQ-BSS algorithm which can be described in Algorithm 1.

It should be noted that $\eta_A(\mathbf{B}_1^{(l)}, \dots, \mathbf{B}_{j-1}^{(l)}, \mathbf{B}_j, \mathbf{B}_{j+1}^{(l)}, \dots, \mathbf{B}_p^{(l)}; \Omega)$ is determined by various applications specified by the subscript of η_A .

2) *Successive Band Subset Selection Algorithm*: The second algorithm to be developed from SC N-FINDR is called SC-BSS and its detailed implementation is given in Algorithm 2.

III. BSS-BASED ANOMALY DETECTION

Applications are generally used to justify the utility of BS. Specifically, classification has been widely used for this purpose [4], [7], [9], [10], [15]–[17], [23]–[30], [33]. However, AD seems to have received little attention in BS except some experiments done in [24]. Accordingly, this paper has mainly focused on AD [38], [43] and conducted extensive experiments in performance evaluation for BSS.

Algorithm 1 SQ-BSS Algorithm

1. Initialization:
 - a. Let p be the number of selected bands determined by VD.
 - b. Let $\{\mathbf{B}_1^{(0)}, \mathbf{B}_2^{(0)}, \dots, \mathbf{B}_p^{(0)}\}$ be a set of initial bands randomly selected from the entire band set Ω . Set $l = 1$.
2. Outer Loop: (using index l as a counter to keep track band \mathbf{B}_l)

Check $l = L$. If it is, the algorithm terminated. Otherwise, let $l \leftarrow l + 1$ and continue.
3. Input the l^{th} band, \mathbf{B}_l . (Note that the \mathbf{B}_l here is now the $l + 1^{\text{st}}$ band, \mathbf{B}_{l+1}).
4. Inner Loop: (using m as a counter to keep track the j^{th} band \mathbf{B}_j)

For $1 \leq j \leq p$, we re-calculate $\eta_A(\mathbf{B}_1^{(l)}, \dots, \mathbf{B}_{j-1}^{(l)}, \mathbf{B}_j, \mathbf{B}_{j+1}^{(l)}, \dots, \mathbf{B}_p^{(l)}; \Omega)$ for the band \mathbf{B}_j . If any of these p recalculated the performance measure, $\eta_A(\mathbf{B}_1, \mathbf{B}_2^{(l)}, \dots, \mathbf{B}_p^{(l)}; \Omega)$, $\eta_A(\mathbf{B}_1^{(l)}, \mathbf{B}_l, \mathbf{B}_3^{(l)}, \dots, \mathbf{B}_p^{(l)}; \Omega), \dots, \eta_A(\mathbf{B}_1^{(l)}, \dots, \mathbf{B}_{p-1}^{(l)}, \mathbf{B}_l; \Omega)$, is greater than $\eta_A(\mathbf{B}_1^{(l)}, \mathbf{B}_2^{(l)}, \dots, \mathbf{B}_p^{(l)}; \Omega)$, go to step 5. Otherwise, go to step 2.
5. Replacement rule:

Find an index j^* by

$$j^* = \arg \left\{ \min_{1 \leq j \leq p} \eta_A(\mathbf{B}_1^{(l)}, \dots, \mathbf{B}_{j-1}^{(l)}, \underbrace{\mathbf{B}_l}_j, \mathbf{B}_{j+1}^{(l)}, \dots, \mathbf{B}_p^{(l)}; \Omega) \right\} \quad (3)$$

which specified the band be replaced by the l^{th} band \mathbf{B}_l . Assume that such an band is now denoted by $\mathbf{B}_j^{(l+1)}$. A new set of bands is then produced by letting $\mathbf{B}_{j^*}^{(l+1)} = \mathbf{B}_l$ and $\mathbf{B}_j^{(l+1)} = \mathbf{B}_j^{(l)}$ for $j \neq j^*$ and go to step 3.

In order to effectively detect anomalies, an algorithm developed by Reed and Yu [39] referred to as Reed-Xiaoli detector (RXD) has been widely used. Since its development, many RXD-like anomaly detectors have been proposed [36]–[38]. Of particular interest are anomaly detectors which modify RXD by replacing the global sample covariance matrix, \mathbf{K} , with the global sample correlation matrix \mathbf{R} . In this case, the resulting RXD is called R-AD, while RXD using \mathbf{K} is denoted by K-AD for distinction.

Assume that $\{\mathbf{r}_i\}_{i=1}^N$ where N is the total number of entire data sample vectors in the data and $\mathbf{r}_i = (r_{i1}, r_{i2}, \dots, r_{iL})^T$ is the i th data sample vector where L is the total number of spectral bands. The K-AD, denoted by $\delta^{\text{K-AD}}(\mathbf{r})$, is specified by

$$\delta^{\text{K-AD}}(\mathbf{r}) = (\mathbf{r} - \boldsymbol{\mu})^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}) \quad (5)$$

where $\boldsymbol{\mu}$ is the global sample mean given by $\boldsymbol{\mu} = (1/N) \sum_{i=1}^N \mathbf{r}_i$ and \mathbf{K} is the global sample data covari-

Algorithm 2 SC-BSS Algorithm

1. Initialization:
Let p be the number of endmembers required to generate and $\{\mathbf{B}_1^{(0)}, \mathbf{B}_2^{(0)}, \dots, \mathbf{B}_p^{(0)}\}$ be a set of initial bands randomly selected from Ω where $\Omega = \{\mathbf{B}_l\}_{l=1}^L$ is the set of all spectral bands.
2. Outer Loop
For $1 \leq j \leq p$ find $\mathbf{B}_j^{(*)}$ to replace $\mathbf{B}_j^{(0)}$.
3. Inner Loop for a given j in the outer loop:
For $1 \leq l \leq L$ calculate $\eta_A = (\mathbf{B}_1^{(*)}, \dots, \mathbf{B}_{j-1}^{(*)}, \mathbf{B}_l, \mathbf{B}_{j+1}^{(*)}, \dots, \mathbf{B}_p^{(*)}; \Omega)$ for all $\{\mathbf{B}_l\}_{l=1}^L$, while fixing other bands $\mathbf{B}_i^{(*)}$ with $i < j$ and $\mathbf{B}_i^{(0)}$ with $i > j$. Find

$$\mathbf{B}_j^{(*)} = \arg \left\{ \min_{\mathbf{B}_l} \eta_A(\mathbf{B}_1^{(*)}, \dots, \mathbf{B}_{j-1}^{(*)}, \mathbf{B}_l, \mathbf{B}_{j+1}^{(*)}, \dots, \mathbf{B}_p^{(*)}; \Omega) \right\}. \quad (4)$$
4. Stopping rule:
If $j \leq p$, then $j \leftarrow j + 1$ and go step 2. Otherwise, the final set of $\{\mathbf{B}_1^{(*)}, \mathbf{B}_2^{(*)}, \dots, \mathbf{B}_p^{(*)}\}$ is the desired p bands.

ance matrix given by $\mathbf{K} = (1/N) \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu})(\mathbf{r}_i - \boldsymbol{\mu})^T$. Interestingly, the form of $\delta^{\mathbf{K}-\text{RXD}}(\mathbf{r})$ in (6) is actually the well-known Mahalanobis distance. Another is R-AD, denoted by $\delta^{\text{R-AD}}(\mathbf{r})$, is specified by

$$\delta^{\text{R-AD}}(\mathbf{r}) = \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} \quad (6)$$

with $\mathbf{R} = (1/N) \sum_{i=1}^N \mathbf{r}_i \mathbf{r}_i^T$.

In regard to AD two stopping criteria for SQ/SC-BSS algorithm can be defined for η_A . One is that we assume ground truth is available for performance evaluation in which case $\eta_A = (\text{AUC})$. In this case, we can find

$$\Omega_{\text{BS}}^{\text{K-AD/R-AD}} = \arg\{\max_{\Omega_{\text{BS}} \subset \Omega} \text{AUC}^{\text{K-AD/R-AD}}(\Omega_{\text{BS}})\} \quad (7)$$

with $|\Omega_{\text{BS}}| = n_{\text{BS}} = p$ where $\text{AUC}(\Omega_{\text{BS}})$ is calculated by the area under the ROC curve produced by K-AD/R-AD using only bands selected from Ω_{BS} .

The other is that we assume no ground truth is available for performance evaluation. In this case, $\eta_A = \text{LSE}$. In this case, we need to calculate LSE between the results produced by full bands and the results by selected bands, Ω_{BS} , that is

$$\Omega_{\text{BS}}^{\text{K-AD/R-AD}} = \arg\{\max_{\Omega_{\text{BS}} \subset \Omega} \text{LSE}^{\text{K-AD/R-AD}}(\Omega_{\text{BS}})\}. \quad (8)$$

The two stopping rules specified by (7) and (8) will be used in the experiments performed in this paper.

IV. DISCUSSION ON MULTIPLE BAND SELECTION

Generally, MBS can be carried out in two ways, one band at a time sequentially as SQMBS does and multiple bands simultaneously as SMMBS does by selecting multiple bands to be a band subset as a whole. In theory, most SBS-based approaches can be extended to SQMBS such as BP criterion-based BS methods [3], constrained band selection (CBS) [5], and most recently, sequential feature search/sequential feature

forward search-based algorithms [11] by augmenting selected bands gradually [13], [14]. However, as for SMMBS there is a more complicated issue, which is requirement of an exhaustive search for finding an optimal set of bands. For selecting an optimal set of p bands among the total number of L bands, it requires running through all p -band combinations $\binom{L}{p} = (L!/(p!(L-p)!))$. Practically, this is impossible to do so if L is large such as hyperspectral imagery. In order to mitigate this problem, several approaches have been studied. One approach is to use band clustering to group all L bands into a predetermined number of clusters where the cluster centers or representatives are selected as desired bands [24], [30]. In particular, the concept in [25] is similar to Fisher's ratio using mutual information as a BP criterion for clustering. As an alternative, some approaches based on band groups or band combinations are proposed in [15], [16], [23], and [24]. For example, [23] is a band group-wise method which uses compressive sensing along with the multitasks sparsity pursuit (MTSP)-based criterion to select band combinations based on linear sparse representation where the used search strategies are evolution-based algorithms. Unfortunately, such methods did not run through all possible band combinations because each band combination is considered as antibody and the set of antibody populations was fixed at a predetermined parameter $N = 10$. In other words, for a p -band combination, the approach in [23] only runs through a predetermined number of antibody populations specified by N . However, in order to conduct an exhaustive search, this N must be sufficiently enough to represent $\binom{L}{p} = (L!/(p!(L-p)!))$ p -combinations. So, technically speaking, this algorithm is practically impossible to be implemented in this way. This same problem also arises in [15] and [16] where an antibody specified in [23] is replaced by a firefly in [16] and a swam in [15] with antibody populations N replaced by the total number of swarm particles, $M = 25$, used by particle swarm optimization (PSO) in [15] and the total number of fireflies, $m = 10$, used by the firefly algorithm (FA) in [16], both of which have exactly the same issue as N had in [23]. This implies that all the approaches in [15], [16], and [23] only run through band combinations with fixed $M = 25$ or $m = 10$ or $N = 10$, all of which are empirically predetermined. There were no provided guidelines of how to determine these values. Most recently, Yuan *et al.* [24] proposed a graph-based SMMBS method, called multigraph determinantal point process (MDPP) which makes use of multiple graphs to discover a structure and diverse band subset from a graph where each band is considered as a node and the edge is specified by similarity between bands. Accordingly, a path represents a possible band subset. Then a search algorithm, called mixture DPP was further developed to find a diverse subset that can be a potential optimal band combination. Compared to the aforementioned works our proposed BSS belongs to a completely new category which is particularly designed to directly search for an optimal p -band subset out of $\binom{L}{p} = (L!/(p!(L-p)!))$ p -band combinations. It does

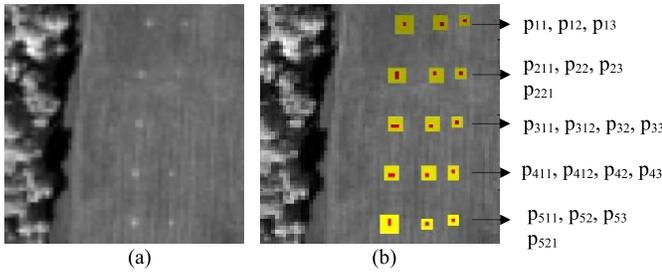


Fig. 1. (a) HYDICE panel scene which contains 15 panels. (b) Ground truth map of spatial locations of the 15 panels.

not use any linear representation form or minimum estimated abundance covariance in [14]–[16] or sparse representation in [23] or m , number of spectral subspaces in multigraph in [24]. Most importantly, it does not require parameters, m , M , or N required by [15], [16], [23], and [24]. As a matter of fact, BSS runs through all possible p -band subsets among all L spectral bands via SQ and SC search processes without fixing a particular number specified by M , m , or N . It is our brief that MDPP in [24], MTSP in [23], PSO in [15], and FA in [16] cannot work this way practically because they only run for specific values of M , m , and N with $1 \leq M, m, N \leq \binom{L}{p} = (L!/(p!(L-p)!))$. So, to our best knowledge, the proposed BSS is probably the only one algorithm designed for running through all p -band combination among all L bands by numerical search algorithms.

V. REAL IMAGE EXPERIMENTS

A real image scene collected by the Hyperspectral Digital Imagery Collection Experiments (HYDICE) shown in Fig. 1 was used for experiments. It has a total of 169 bands along with spatial resolution 1.56 m and spectral resolution 10 nm. Its detailed description can be found in [1] with a size of 64×64 pixel vectors and 15 panels in the scene. The ground truth map is provided in Fig. 1(b). The reason that this scene was selected for experiments is because it provides complete ground truth of small 15 man-made panels which can be considered as anomalies. In this case, we can use ROC analysis to evaluate detection performance in terms of detection probability P_D and false alarm probability P_F .

First of all, we need to determine the number of bands, p , required to be selected, which can be estimated by virtual dimensionality (VD) [44], [45]. If we assume that each signature can be accommodated by a single band, the number of signatures can be then used to estimate the number of bands n_{BS} . For our experiments, VD for this scene was chosen to be 9 according to [1], [44], and [45]. Four experiments were conducted according to two criteria, AUC and LSE. Also, K-AD and R-AD are used for AD. Since AD is completely unsupervised, it is blind target detection. So, in order to conduct a fair comparison, all supervised BS methods are excluded from the study and only uniform band selection (UBS) as well as SBS using various BP criteria, variance, SNR, entropy, and ID considered in [3] was compared for performance analysis.

A. $\eta_A = AUC$

In this section, we assume that the ground truth is provided such that AUC can be used for η_A as a stopping rule to terminate SQ/SC-BSS algorithms. Table I tabulates nine bands selected by UBS, SBS using various BP criteria, variance, SNR, entropy, ID, and SQ/SC-BSS algorithms where “/” is used to separate two selected bands. In the last two columns of Table I, we also calculated AUC produced by K-AD and R-AD using corresponding nine bands. As we can see SQ/SC-BSS algorithms selected much better band subsets than those selected by SBS. Most interestingly, K-AD and R-AD using bands selected by SQ/SC-BSS algorithms produced better AUC values than using full bands. These experimental results demonstrated that selecting effective nine bands could perform better than blindly using full bands in terms of AUC values.

According to [38], one crucial measure to assess effectiveness of AD is background suppression which cannot be simply analyzed by AUC values in Table I. Figs. 2 and 3 show the respective detection maps of K-AD and R-AD using full bands and nine bands selected in Table I where it is very obvious that the detection maps produced by K-AD and R-AD using nine bands selected by SQ/SC-BSS algorithms had better background suppression compared to other BS algorithms. It is also interesting to note that the detection map using full bands actually produced worst background suppression despite that it produced better AUC values than those produced by SBS using various BP criteria. In order to better explain these phenomena, 3-D ROC analysis provides evidence of all the answers. Figs. 4(a)–(d) and 5(a)–(d) plot 3-D ROC curve of (P_D, P_F, τ) and three 2-D ROC curves, 2-D ROC curves of (P_D, P_F) , 2-D ROC curves of (P_D, τ) , and 2-D ROC curves of (P_F, τ) produced by K-AD and R-AD, respectively. By looking at Figs. 4(d) and 5(d), P_F produced by using full bands produced highest P_F and did not converge to 0, which indicated poor background suppression. In addition to AUC obtained by 2-D ROC curve of (P_D, P_F) tabulated in Table I, we set the threshold value τ less than 0.05 for 2-D ROC curves generated by K-AD and R-AD using nine bands selected by SQ/SC-BSS algorithms. In this case, P_F would be close to 0 according to Figs. 4(d) and 5(d), and P_D would be approaching to 1 according to Figs. 4(c) and 5(c). This implies that if a threshold τ is selected appropriately around 0.05, we could have $P_F \rightarrow 0$ and $P_D \rightarrow 1$.

B. $\eta_A = LSE$

Following the same experiments conducted in Section V-A which used AUC as a stopping rule, similar to experiments were also performed for SQ/SC-BSS algorithms using LSE for η_A as a stopping rule where we assumed that there was no ground truth available to calculate AUC values. In this case, we used the results produced by full bands as a gold standard for comparison. Table II tabulates nine bands selected by UBS, SBS using various BP criteria, variance, SNR, entropy, ID as well as SQ/SC-BSS algorithms. In the last two columns of Table II, we also calculated LSE produced by K-AD and R-AD using corresponding nine bands. Once again, we can

TABLE I
K-AD AND R-AD USING AUC AS A CRITERION WITH NINE BANDS SELECTED BY FULL BANDS,
UBS, VARIOUS SBS METHODS, AND SQ/SC-BSS ALGORITHMS

Method	Bands selected by K-AD	Bands selected by R-AD	AUC (K-AD)	AUC (R-AD)
full bands	1:169	1:169	0.9898	0.9900
UBS	9/27/45/63/81/99/117/135/153	9/27/45/63/81/99/117/135/153	0.9679	0.9741
variance	52/36/47/50/13/54/55/96/152	52/36/47/50/13/54/55/96/152	0.9860	0.9857
SNR	102/104/119/139/117/90/140/134/101	102/104/119/139/117/90/140/134/101	0.9872	0.9887
entropy	65/53/59/78/77/49/56/47/58	65/53/59/78/77/49/56/47/58	0.9790	0.9801
ID	154/157/156/153/150/158/145/164/163	154/157/156/153/150/158/145/164/163	0.9915	0.9933
SC-BSS	57/65/111/115/129/134/136/158/165	111/115/129/132/134/135/136/158/165	0.9974	0.9980
SQ-BSS	57/65/66/106/107/111/115/120/168	36/37/124/129/132/133/153/158/168	0.9978	0.9980

TABLE II
K-AD AND R-AD USING LSE AS A CRITERION WITH NINE BANDS SELECTED BY FULL BANDS,
UBS, VARIOUS SBS METHODS, AND SQ/SC-BSS ALGORITHMS

Method	Bands selected by K-AD	Bands selected by R-AD	LSE by K-AD	LSE by R-AD
Full band	1:169	1:169		
Uniform	9/27/45/63/81/99/117/135/153	9/27/45/63/81/99/117/135/153	1.2279e+08	1.2277e+08
variance	52/36/47/50/13/54/55/96/152	52/36/47/50/13/54/55/96/152	1.2356e+08	1.2389e+08
SNR	102/104/119/139/117/90/140/134/101	102/104/119/139/117/90/140/134/101	1.2048e+08	1.2286e+08
entropy	65/53/59/78/77/49/56/47/58	65/53/59/78/77/49/56/47/58	1.2318e+08	1.2228e+08
ID	154/157/156/153/150/158/145/164/163	154/157/156/153/150/158/145/164/163	1.2334e+08	1.2419e+08
SC-BSS	1/47/51/72/106/111/115/140/149	76/77/79/80/106/111/134/141/153	1.1861e+08	1.1472e+08
SQ-BSS	24/38/52/99/107/112/129/133/153	24/44/51/108/109/123/128/133/154	1.1878e+08	1.1933e+08

also see SQ/SC-BSS algorithms selected much better band subsets than those selected by SBS in the sense of producing smaller LSE.

Similarly, Figs. 6 and 7 show the respective detection maps of K-AD and R-AD using full bands and nine bands selected in Table II where the detection maps produced by K-AD and R-AD using nine bands selected by SQ/SC-BSS algorithms had better background suppression compared to other BS algorithms. Furthermore, as shown in Figs. 2(g) and (h) and 3(g) and (h) for the case of AUC, both SQ/SC-BSS algorithms had nearly the same background suppression. But this was not true for the case of LSE. SQ-BSS algorithm produced better band subsets than SC-BSS algorithm did in terms of background suppression. It is also interesting to note that the detection map using full bands actually produced worst background suppression despite that it was used as a gold standard for comparison. Finally, if we used the ground truth to calculate AUC using the nine bands selected by LSE as a stopping rule, Figs. 8(a)–(d) and 9(a)–(d) plot 3-D ROC curve of (P_D, P_F, τ) and three 2-D ROC curves, 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) produced by K-AD and R-AD, respectively. Like Figs. 4 and 5, using full bands did produce highest P_F , which indicated poor background suppression. If we set the threshold value τ less than 0.05 for 2-D ROC curves generated by K-AD and R-AD using nine bands selected by BSS SQ/SC algorithms. In this case, P_F would be close to 0 according to Figs. 8(d) and 9(d). On the other hand, if the threshold τ is chosen to be less than 0.05, P_D would be approaching to 1 according to Figs. 8(c) and 9(c). This implies that if a threshold τ is selected appropriately around 0.05, we could have $P_F \rightarrow 0$ and $P_D \rightarrow 1$. Finally, Tables III and IV tabulate

the AUC values for 2-D ROC curves of (P_D, P_F) obtained by K-AD and R-AD using the nine bands in Table II. As we can see from the table, the best result was the one produced by SQ-BSS algorithm.

As concluding remarks, several observations from the experimental results are noteworthy.

- 1) The conducted experiments demonstrated that if the ground truth was available, AUC would be a better stopping rule than LSE. If there was no ground truth, there was no way to compute AUC in which case LSE must be used.
- 2) According to the above experimental results, BSS could find much better bands as a band subset than SBS did for individual bands since the former selects multiple bands simultaneously as a band subset compared to SBS which makes use of BP to rank all bands and then selects multiple bands one single band at time.
- 3) Using simultaneously selected multiple bands by SQ/SC-BSS algorithms produced the best results among all test SBS algorithms including UBS.
- 4) It seems a general understanding that using spectral information provided by full bands is supposed to produce the best results. Our experiments showed otherwise via 3-D ROC analysis. This is mainly due to the fact that background suppression was never been considered as an evaluation criterion as discussed in [38]. SQ/SC-BSS algorithms always have better background suppression compared to full bands producing the worst background suppression. This implies that using full bands may not be necessary to produce the best results.
- 5) It is important to note that AD cannot be evaluated solely by 2-D ROC analysis as commonly done in the past. Background suppression is a crucial element to assess

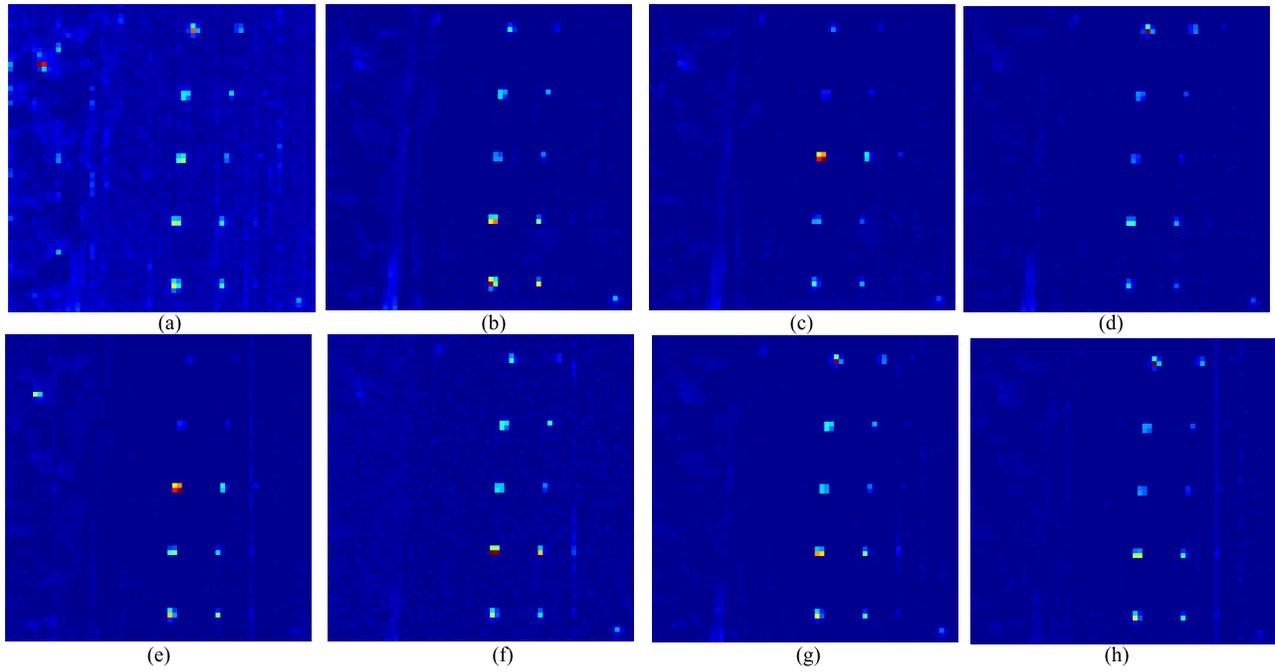


Fig. 2. Detection maps of K-AD using AUC as a criterion and nine bands selected by UBS, various SBS methods, and SQ/SC-BSS algorithms. (a) Full bands. (b) UBS. (c) Variance. (d) SNR. (e) Entropy. (f) ID. (g) SC-BSS-K-AD. (h) SQ-BSS-K-AD.

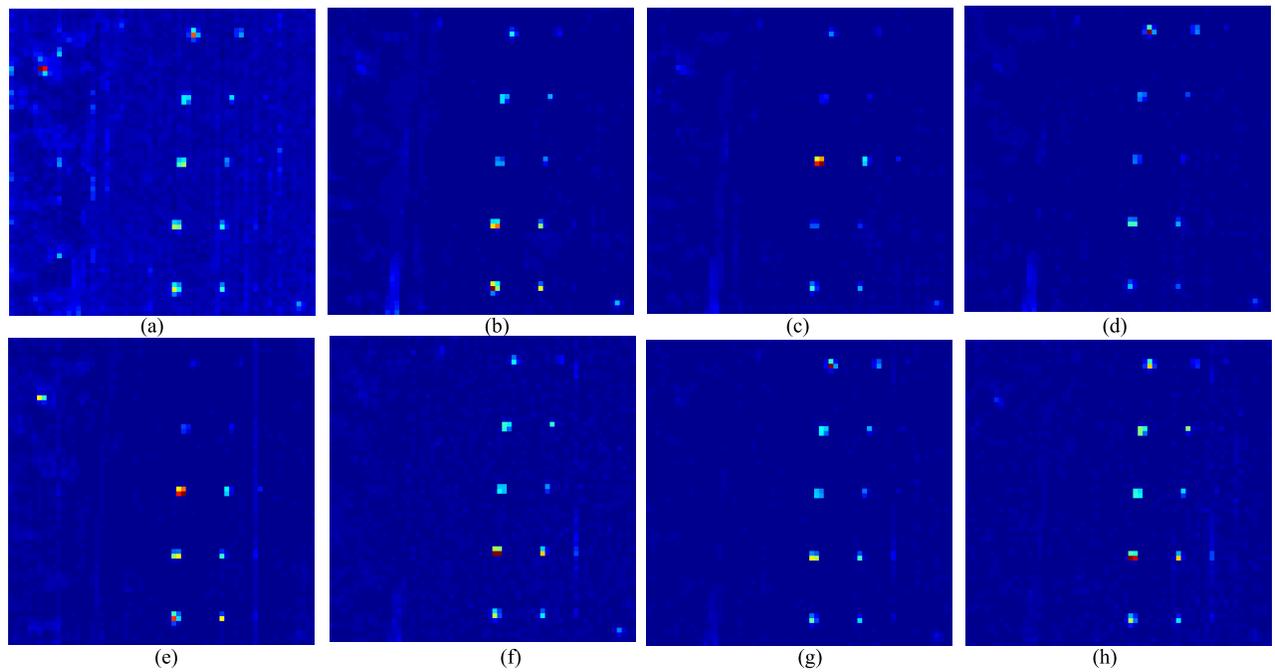


Fig. 3. Detection maps of R-AD using AUC as a criterion and nine bands selected by UBS, various SBS methods, and SQ/SC-BSS algorithms. (a) Full bands. (b) UBS. (c) Variance. (d) SNR. (e) Entropy. (f) ID. (g) SC-BSS-R-AD. (h) SQ-BSS-R-AD.

TABLE III
AUC CALCULATED FROM FIG. 8

Method	Full bands	UBS	variance	SNR	entropy	ID	SC-BSS-K-AD	SQ-BSS-K-AD
AUC	0.9898	0.9679	0.9860	0.9872	0.9790	0.9915	0.9935	0.9943

effectiveness of AD as clearly shown in Figs. 2–9 by visual inspection.

6) As also shown by experiments, 3-D ROC analysis is

a better evaluation tool than the traditional 2-D ROC analysis in the sense that it can use 2-D ROC curves of (P_F, τ) and (P_D, τ) to measure background sup-

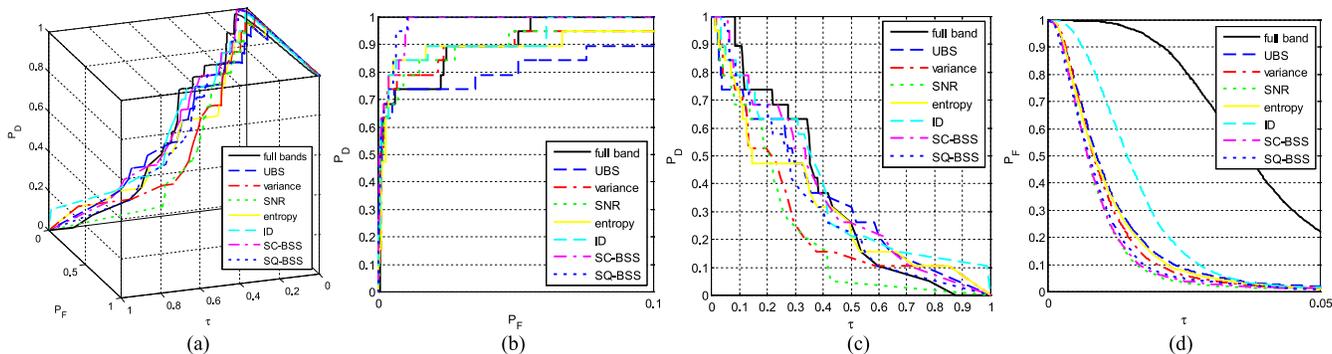


Fig. 4. Three-dimensional ROC and three 2-D ROC curves of K-AD. (a) Three-dimensional ROC curves. (b) Two-dimensional ROC curves of P_D versus P_F . (c) Two-dimensional ROC curves of P_D versus τ . (d) Two-dimensional ROC curves of P_F versus τ .

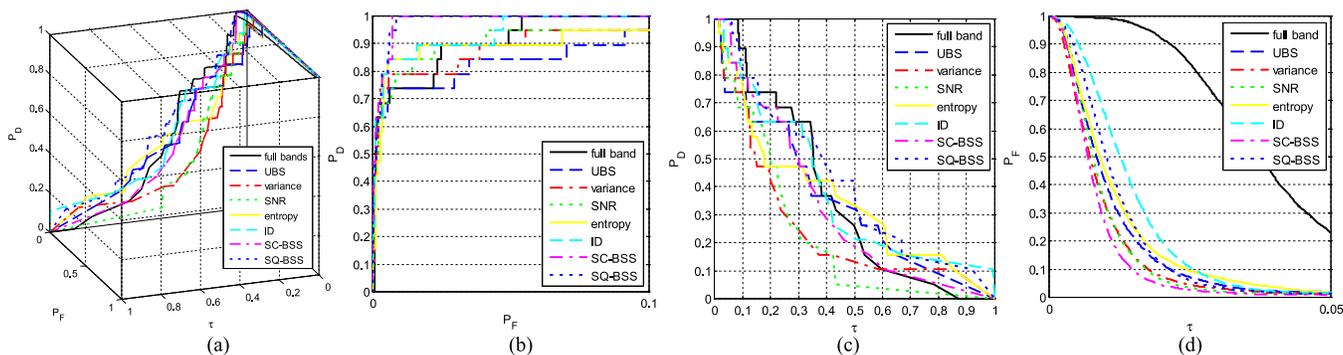


Fig. 5. Three-dimensional ROC and three 2-D ROC curves of R-AD. (a) Three-dimensional ROC curves. (b) Two-dimensional ROC curves of P_D versus P_F . (c) Two-dimensional ROC curves of P_D versus τ . (d) Two-dimensional ROC curves of P_F versus τ .

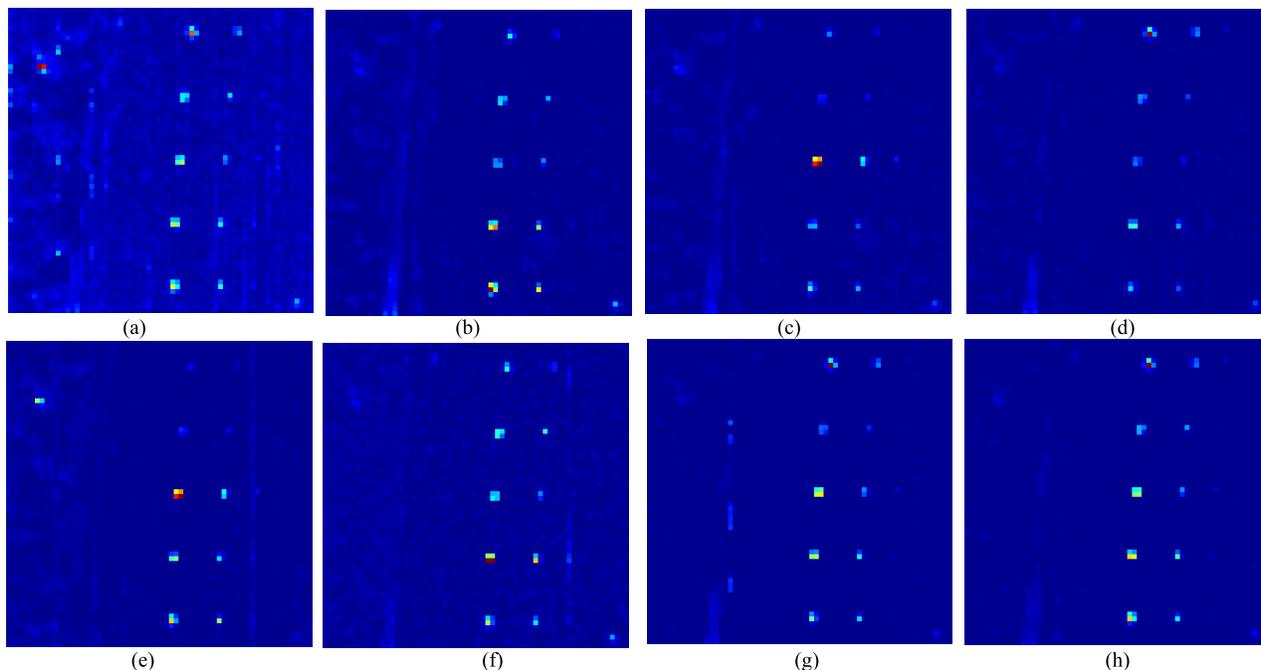


Fig. 6. Detection maps of K-AD using LSE as a criterion and nine bands selected by UBS, various SBS methods, and SQ/SC-BSS algorithms. (a) Full bands. (b) UBS. (c) Variance. (d) SNR. (e) Entropy. (f) ID. (g) SC-BSS-K-AD. (h) SQ-BSS-K-AD.

pression and detection power with the threshold τ as a parameter compared to 2-D ROC curve of (P_D, P_F) which only evaluates P_D relative to P_F , both of which

are actually functions of the threshold parameter τ . As a consequence, 2-D ROC curve completely discards the issue of background suppression because the false alarm

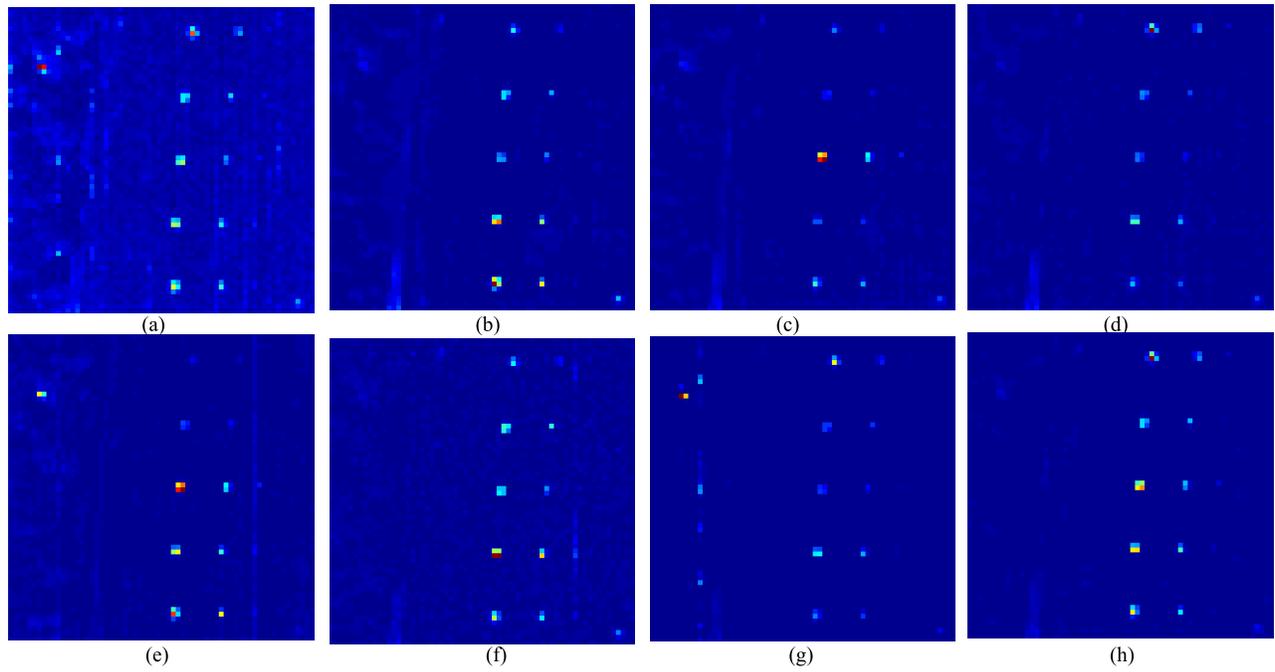


Fig. 7. Detection maps of R-AD using LSE as a criterion and nine bands selected by UBS, various SBS methods, and SQ/SC-BSS algorithms. (a) Full bands. (b) UBS. (c) Variance. (d) SNR. (e) Entropy. (f) ID. (g) SC-BSS-R-AD. (h) SQ-BSS-R-AD.

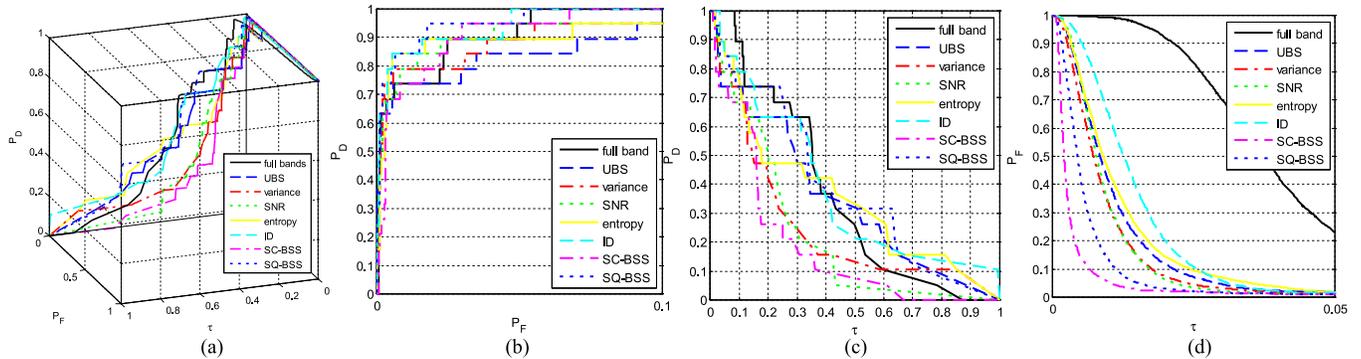


Fig. 8. Three-dimensional ROC and three 2-D ROC curves of K-AD. (a) Three-dimensional ROC curves. (b) Two-dimensional ROC curves of P_D versus P_F . (c) Two-dimensional ROC curves of P_D versus τ . (d) Two-dimensional ROC curves of P_F versus τ .

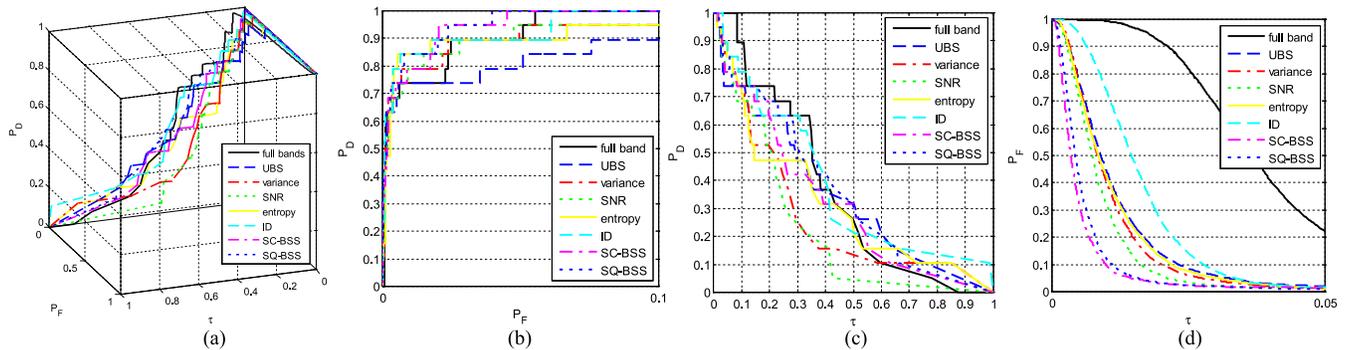


Fig. 9. Three-dimensional ROC and three 2-D ROC curves of R-AD. (a) Three-dimensional ROC curves. (b) Two-dimensional ROC curves of P_D versus P_F . (c) Two-dimensional ROC curves of P_D versus τ . (d) Two-dimensional ROC curves of P_F versus τ .

TABLE IV
AUC CALCULATED FROM FIG. 9

Method	Full bands	UBS	variance	SNR	entropy	ID	SC-BSS-R-AD	SQ-BSS-R-AD
AUC	0.9900	0.9741	0.9857	0.9887	0.9801	0.9933	0.9897	0.9946

probability P_F is used as an independent parameter to measure the detection probability, P_D and itself cannot be used to measure background suppression. On the other hand, 3-D ROC analysis is developed to treat τ as an independent variable of the 3-D ROC curve of (P_D, P_F, τ) from which three types of 2-D ROC curves of (P_D, P_F) , (P_D, τ) and (P_F, τ) can be generated for performance evaluation, specifically, (P_F, τ) can be used to assess the background suppression.

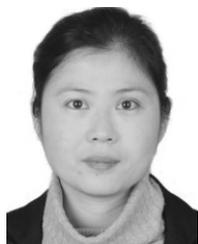
VI. CONCLUSION

This paper develops a new approach for BSS-based AD which selects multiple bands as a band subset simultaneously. There are several contributions made in this paper. First of all, it extends the conventional SBS to select multiple bands altogether as a band subset. Its idea is very close to that used for endmember extraction where the search algorithm for finding an optimal endmember set can be considered to be equivalent to algorithms that are used to find an optimal band subset. Second, two sequential algorithms, SQ N-FINDR, and SC N-FINDR developed for N-FINDR are further used to derive for BSS as their counterparts, SQ-BSS and SC-BSS algorithms. Third, since the criterion used for BSS is determined by a particular application of interest, AD is chosen for this purpose due to the fact that very little work has been done for AD using BS. Two reasons are attributed to this cause. One is that anomaly is a blind target detection with no required target knowledge. In this case, ROC analysis is not applicable and it is very challenging to evaluate its performance without ground truth. Another is that how can we evaluate effectiveness of AD without prior knowledge? As a matter of fact, as shown in [38], this issue can be addressed by background suppression. Fourth, in order to effectively evaluate the effect of background suppression on AD, we extend traditional 2-D ROC analysis to 3-D ROC analysis for AD where two more 2-D ROC curves of (P_D, τ) and (P_F, τ) can be further generated to analyze target detection probability P_D and false alarm probability P_F individually and separately via a threshold parameter τ . Finally and most importantly, the experimental results showed that using full bands produced the worst background suppression even though P_D is very high. Such phenomenon can be only addressed by P_F which has been overlooked and never reported in the literature. Furthermore, the experiments also demonstrated that BSS found better band subsets than SBS did for AD and also performed better than using full bands in terms of background suppression with lower P_F . It is our belief that BSS is indeed a promising BS technique in many other applications yet to explore.

REFERENCES

- [1] C.-I. Chang, *Hyperspectral Data Processing: Signal Processing Algorithm Design and Analysis*. Hoboken, NJ, USA: Wiley, 2013.
- [2] C. Conese and F. Maselli, "Selection of optimum bands from TM scenes through mutual information analysis," *ISPRS J. Photogram. Remote Sens.*, vol. 48, no. 3, pp. 2–11, 1993.
- [3] 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.
- [4] R. Huang and M. He, "Band selection based on feature weighting for classification of hyperspectral data," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 2, pp. 156–159, Apr. 2005.
- [5] 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.
- [6] N. Keshava, "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1552–1565, Jul. 2004.
- [7] P. W. Mausel, W. J. Kramber, and J. K. Lee, "Optimum band selection for supervised classification of multispectral data," *Photogram. Eng. Remote Sens.*, vol. 56, no. 1, pp. 55–60, 1990.
- [8] P. Bajcsy and P. Groves, "Methodology for hyperspectral band selection," *Photogram. Eng. Remote Sens.*, vol. 10, no. 7, pp. 793–802, Jul. 2004.
- [9] S. Jia, Z. Ji, Y. Qian, and L. Shen, "Unsupervised band selection for hyperspectral imagery classification without manual band removal," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 5, no. 2, pp. 531–543, Apr. 2012.
- [10] S. D. Stearns, B. E. Wilson, and J. R. Peterson, "Dimensionality reduction by optimal band selection for pixel classification of hyperspectral imagery," *Proc. SPIE*, vol. 2028, p. 118, Oct. 1993.
- [11] P. Pudil, J. Novovičova, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognit. Lett.*, vol. 15, no. 11, pp. 1119–1125, 1994.
- [12] S. de Backer, P. Kempeneers, W. Debruyn, and P. Scheunders, "Band selection for hyperspectral remote sensing band selection for hyperspectral remote sensing," *Pattern Recognit.*, vol. 2, no. 3, pp. 319–323, 2005.
- [13] 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.
- [14] 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.
- [15] H. Su, Q. Du, G. Chen, and P. Du, "Optimized hyperspectral band selection using particle swarm optimization," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 7, no. 6, pp. 2659–2670, Jun. 2014.
- [16] H. Su, B. Yang, and Q. Du, "Hyperspectral band selection using improved firefly algorithm," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 68–72, Jan. 2016.
- [17] W. Wei, Q. Du, and N. H. Younan, "Fast supervised hyperspectral band selection using graphics processing unit," *J. Appl. Remote Sens.*, vol. 6, no. 1, pp. 061504-1–061504-12, 2012.
- [18] 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.
- [19] X. Geng, K. Sun, and L. Ji, "Band selection for target detection in hyperspectral imagery using sparse CEM," *Remote Sens. Lett.*, vol. 5, no. 12, pp. 1022–1031, 2014.
- [20] 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.
- [21] K. Sun, X. Geng, and L. Ji, "Exemplar component analysis: A fast band selection method for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 998–1002, May 2015.
- [22] 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.
- [23] 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.
- [24] 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.
- [25] A. Martínez-Usó, F. Pla, J. M. Sotoca, and 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.
- [26] 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.
- [27] J. Feng, L. Jiao, F. Liu, T. Sun, and X. Zhang, "Mutual-information-based semi-supervised hyperspectral band selection with high discrimination, high information, and low redundancy," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2956–2969, May 2015.

- [28] 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.
- [29] H. Yang, H. Su, 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.
- [30] H. Su and Q. Du, "Hyperspectral band clustering and band selection for urban land cover classification," *Geocarto Int.*, vol. 27, no. 5, pp. 395–411, 2012.
- [31] J. E. Ball, L. M. Bruce, and N. H. Younan, "Hyperspectral pixel unmixing via spectral band selection and DC-insensitive singular value decomposition," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 3, pp. 382–386, Jul. 2007.
- [32] K. Koonsanit, C. Jaruskulchai, and A. Eiumnroh, "Band selection for dimension reduction in hyper spectral image using integrated information gain and principal components analysis technique," *Int. J. Mach. Learn. Comput.*, vol. 2, no. 3, pp. 248–251, Jun. 2012.
- [33] W. Xia, B. Wang, and L. Zhang, "Band selection for hyperspectral imagery: A new approach based on complex networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1229–1233, Sep. 2013.
- [34] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proc. SPIE*, vol. 3753, p. 266, Oct. 1999.
- [35] C. C. Wu, S. Chu, and C.-I. Chang, "Sequential N-FINDR algorithms," *Proc. SPIE*, vol. 7086, p. 70860C, Aug. 2008.
- [36] W. Xiong, C.-I. Chang, C. C. Wu, K. Kalpakis, and H. M. Chen, "Fast algorithms to implement N-FINDR for hyperspectral endmember extraction," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 4, no. 3, pp. 545–564, Sep. 2011.
- [37] C.-I. Chang, "Maximum simplex volume-based endmember extraction algorithms," U.S. Patent 8417748 B2, Apr. 9, 2013.
- [38] C.-I. Chang, *Real Time Progressive Hyperspectral Image Processing: Endmember Finding and Anomaly Detection*. New York, NY, USA: Springer, 2016, chs. 5–16.
- [39] I. S. Reed and X. Yu, "Adaptive Multiple-Band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.
- [40] C.-I. Chang and S.-S. Chiang, "Anomaly detection and classification for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 6, pp. 1314–1325, Jun. 2002.
- [41] L. Wang and C.-I. Chang, "Multiple band selection for anomaly detection in hyperspectral imagery," in *Proc. IEEE Geosci. Remote Sens. Symp.*, Beijing, China, Jul. 2016, pp. 7022–7025.
- [42] C.-I. Chang, "Multiple-parameter receiver operating characteristic analysis for signal detection and classification," *IEEE Sensors J.*, vol. 10, no. 3, pp. 423–442, Mar. 2010.
- [43] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectra images," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 25, no. 7, pp. 5–27, Jul. 2010.
- [44] C.-I. Chang, *Hyperspectral Imaging: Spectral Techniques for Detection and Classification*. New York, NY, USA: Kluwer, 2003.
- [45] 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.
- [46] H. C. Li, C.-I. Chang, Y. Li, and L. Wang, "Constrained multiple band selection for hyperspectral imagery," in *Proc. IEEE Geosci. Remote Sens. Symp.*, Beijing, China, Jul. 2016, pp. 6149–6152.



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

From 2015 to 2016, she was a Visiting Research Associate Scholar at the Remote Sensing Signal and Image Processing Laboratory, University of Maryland, Baltimore County, 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.



Chin-I Chang (S'81–M'87–SM'92–F'10–LF'17) received the B.S. degree in mathematics from Soochow University, in 1973, Taipei, Taiwan, the M.S. degree in mathematics from the Institute of Mathematics, National Tsing Hua University, in 1975, Hsinchu, Taiwan, the M.A. degree in mathematics from The State University of New York, in 1977, Stony Brook, NY, USA, the M.S. and M.S.E.E. degrees from the University of Illinois at Urbana–Champaign, in 1980 and 1982, Urbana, IL, USA, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, MD, USA.

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

Dr. Chang is a Fellow of the Society for Photo-Optical Instrumentation Engineers (SPIE). He was a recipient of 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. He was a Plenary Speaker of SPIE Optics+Applications, Remote Sensing Symposium in 2009. He was also a Keynote Speaker at the User Conference of Hyperspectral Imaging 2010, Industrial Technology Research Institute, Hsinchu; the 2009 Annual Meeting of the Radiological Society of the Republic of China, Taichung; the 2008 International Symposium on Spectral Sensing Research; and the Conference on Computer Vision, Graphics, and Image Processing 2003, Kimen and 2013, Nantou, Taiwan. He was the Guest Editor of a special issue of the *Journal of High Speed Networks on Telemedicine and Applications* in 2000 and the 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, and *Multiplexed, Hyperspectral, and Polarimetric Imaging Technology* for the *Journal of Sensors*. He was an Associate Editor in the area of hyperspectral signal processing of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, from 2001 to 2007. He is currently an Associate Editor of *Artificial Intelligence Research*. He currently serves 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*.



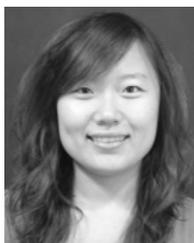
Li-Chien Lee received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 2014, and the M.S. degree in electrical engineering from the University of Maryland, Baltimore County, Baltimore, MD, USA, in 2016, where he is currently pursuing the Ph.D. degree with the Remote Sensing, Signal, and Image Processing Laboratory.

His research interests include hyperspectral imaging, remote sensing, statistic signal processing, machine learning, and large-scale data analysis.



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, Baltimore County, Baltimore, 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.



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

She was awarded by the China Scholarship Council in 2011, as a joint Ph.D. Student to study at the Remote Sensing Signal and Image Processing Laboratory, University of Maryland, Baltimore County, 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 signs signal processing.



Chuanyan Yu 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, where she was a Post-Doctoral Fellow with the Information Science and Technology College, from 2013 to 2016. 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 a Lecturer with the Information Science and Technology College, Dalian Maritime University. Her research interests include image segmentation, hyperspectral image classification, and pattern recognition.



Bai Xue (S'15) received the B.E. degree in automation from the Huazhong University of Science and Technology, Wuhan, China, in 2015. He is currently pursuing the Ph.D. degree in electrical engineering with the University of Maryland, Baltimore County, Baltimore, MD, USA.

From 2014 to 2015, he was a Program Exchange Undergraduate Student at the Department of Electrical and Computer Engineering, University of Detroit Mercy, Detroit, MI, USA. He is currently a Visiting Research Assistant with the Center for Hyperspectral

Imaging in Remote Sensing, Dalian Maritime University, Dalian, China. His research interests include multispectral/hyperspectral imaging, pattern recognition, medical imaging, and medical data analysis.



Sen Li received the B.S. degree in microelectronics from Liao Ning University, Shenyang, China, in 1996, and the M.S. degree in information and communication engineering and the Ph.D. degree in information and signal processing from Dalian Maritime University, Dalian, China, in 1999 and 2011, respectively.

In 2013, she held a post-doctoral position at the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada. She is currently an Associate Professor with the Department of Information Science and Technology, Dalian Maritime University. Her research interests include array signal processing, communication signal processing, non-Gaussian signal processing, and hyperspectral image processing.