Decomposition Model With Background Dictionary Learning for Hyperspectral Target Detection

Tongkai Cheng and Bin Wang, *Senior Member, IEEE*

***Abstract*—Representation-based target detectors for hyperspectral imagery have attracted considerable attention in recent years. However, their detection performance is still unsatisfactory due to the independent manner of the recovery process on each test pixel. Moreover, the background dictionary generated through the dual windows is susceptible to target contamination. Aiming to address these issues, in this article, we propose a decomposition model (DM) with background dictionary learning (BDL) for hyperspectral target detection. The observed data are decomposed into three parts: background, target, and noise. The background and target dictionaries are utilized to represent the background and target components, respectively. In order to achieve a satisfactory recovery of the background and target components, the proposed DM exploits the spatial smoothness of background pixels and the scarcity of the targets of interest in the whole scene via the total variation and the sparsity, respectively. Then, the separated target image is directly used for the detection purpose. Furthermore, a novel BDL method based on the locality-constrained linear coding is presented, and a complete and compact background dictionary can be learned with a low computational cost. Meanwhile, the *a priori* target dictionary is also incorporated into the learning process in order to suppress the contamination of the target signal on the learned background spectra. Extensive experiments on both simulated and real hyperspectral datasets demonstrate the superiority of the proposed detector in comparison with several conventional and state-of-the-art target detectors.**

***Index Terms*—Background dictionary learning (BDL), hyperspectral imagery, locality-constrained linear coding (LLC), target detection, total variation (TV).**

I. Introduction

HYPERSPECTRAL imaging sensors capture the data from the ground surface with hundreds of narrow bands, and each pixel is characterized by a nearly contiguous spectral curve [1]. Since all materials reflect electromagnetic energy at specific wavelengths in distinctive patterns [2], the hyperspectral imagery with a fine spectral resolution is suitable for discriminating substances with subtle spectral differences [3] and has proven valuable in several applications, such as target detection and classification.

Manuscript received November 12, 2020; revised December 25, 2020; accepted January 2, 2021. Date of publication January 8, 2021; date of current version January 27, 2021. This work was supported by the National Natural Science Foundation of China under Grant 61971141 and Grant 61731021. *(Corresponding author: Bin Wang.)*

The authors are with the Key Laboratory for Information Science of Electromagnetic Waves (MoE), Fudan University, Shanghai 200433, China, and also with the Research Center of Smart Networks and Systems, School of Information Science and Technology, Fudan University, Shanghai 200433, China (e-mail: 16110720007@fudan.edu.cn; wangbin@fudan.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3049843

The objective of target detection is to determine whether a specific material is present in the scene on the condition of a known target spectral signature. When the prior target information is unavailable, the anomaly detection is applied with the goal of searching for abnormal objects with different spectral features from the natural background [3]. Generally speaking, with the information of the targets of interest being provided, target detection (supervised) can achieve more accurate result than anomaly detection (unsupervised). Target detection has been widely used in various fields [4]–[6], for instance, detecting aircraft for the purpose of search and rescue, land mines for defense and public safety, and man-made objects in reconnaissance. In this article, we focus on the target detection problem.

Theoretically, target detection can be considered as a binary hypothesis testing problem where each test pixel is labeled as a target (target present) or background (target absent) [2]. Over the past decades, a number of classical target detection approaches have emerged and they roughly fall into two categories, i.e., statistical or subspace-based, according to their ways to describe the spectral variability [2]. In the statistical techniques, the target and background spectra are assumed to follow specific distributions, e.g., multivariate normal distribution. The spectral matched filter (SMF) [7], [8] assumes that the target and background both follow the Gaussian distributions with the same covariance matrix but differ in their mean vectors. The constrained energy minimization (CEM) [9] designs a finite-impulse filter that highlights the response of targets and suppresses that of the background signal. The formulation of CEM is equivalent to that of SMF when the mean of the observed data is removed beforehand in SMF [10]. The adaptive coherence/cosine estimator [11] considers a scaled background covariance matrix in the case of a subpixel target. However, the performance of these statistical detectors is usually unsatisfactory since specific assumptions on the distributions of the background and target spectra are frequently violated in real hyperspectral data. Subspace-based approaches consider that the background pixels or targets lie in a low-dimensional subspace. The representative algorithms include the orthogonal subspace projection [12] and matched subspace detector (MSD) [13]. Unfortunately, the columns in the background and target subspaces are required to be independent in order to generate the projection operators, which limits their generalization capability.

In recent years, representation-based target detectors without any explicit assumptions on the statistical distribution of the observed data have received extensive attention. The sparsity-based target detector (STD) [14] models each test pixel as a

linear combination of few atoms from both the target and background dictionaries, and the recovered sparse representation is used for detection. Considering the correlations of pixels in a small neighborhood, a joint sparse representation-based detector [15] is proposed. Another detector named the spatially adaptive STD [16] is established in a similar way where different weights are assigned for neighboring pixels in heterogeneous areas. Moreover, in order to incorporate the class information of the training samples in the recovery process, several detectors are developed based on the binary hypothesis model. The sparse representation-based binary hypothesis (SRBBH) detector [17] utilizes the background-only training samples to represent the test pixel when the target is absent and both target and background ones to approximate the test pixel when the target is present. In the target detector by combining the sparse and collaborative representation (CSCR) [18], each test pixel is sparsely represented by target-only training samples or collaboratively represented by background-only training samples. However, there are two main problems with the aforementioned representation-based detectors. First, the recovery process is performed in an independent manner for each test pixel. Undesirable conditions may occur without exploiting the relationships of background pixels and targets in the whole scene due to the complexity of hyperspectral data caused by the inherent variability of spectral signatures. Second, the background training samples are generated locally through dual windows centered at the test pixel. However, it is not easy to set the appropriate window sizes and the obtained background dictionary is susceptible to target contamination, especially when the targets are of irregular shape and distribute closely to each other.

More recently, some efforts have been devoted to addressing the above-mentioned problems to some extent. In the anomaly detection task, several commonly used properties about the background pixels (low rank or spatially smooth) and anomalous targets (sparse) [19], [20] have also been applied in the target detection problem [21], in which a detector based on the sparse and low-rank matrix decomposition (SLRMD) of the observed data is established. However, the recovery of background is still unsatisfactory without the use of an appropriate background dictionary [22]. For the second issue, some approaches consider designing an overcomplete background dictionary through the *k*-means clustering [23], [24]. In order to achieve a satisfactory representation of the background part, the background dictionary is expected to cover all the ground materials in the scene, i.e., the dictionary is complete. However, some background classes cannot be represented well with an undercomplete background dictionary if the cluster number is underestimated. Another approach, namely, dictionary learning [25]–[32], alternates between the sparse coding of the samples based on the current dictionary and the update of dictionary atoms to better fit the observed data. The learned dictionary that comprises the whole scene is complete and compact. Nevertheless, these learning algorithms require to solve an ℓ_1-norm minimization problem, which has a high computational cost. Besides, the target spectrum may be also learned in the background dictionary.

In this article, we propose a decomposition model (DM) with background dictionary learning (BDL) for hyperspectral target detection. By using DM, the observed data are decomposed into three parts: background, target, and noise. The background and target dictionaries are utilized to represent the background and target components, respectively. Since the background is dominant and spatially homogeneous, while the targets of interest are sparsely populated in the scene, two constraints, i.e., total variation (TV) and the sparsity are imposed on the representation coefficient matrices to constrain the background and target components, respectively. With the relationships regarding the background pixels and targets in the whole scene being taken into account, a satisfactory decomposition is expected to be achieved. After that, the separated target component is directly used to detect the targets. Moreover, we also present a novel BDL method based on locality-constrained linear coding (LLC) [33] to learn a complete and compact background dictionary. The locality is more essential than sparsity since locality must lead to sparsity but the reverse does not always hold [34]. Compared with sparse coding, LLC just requires to solve an ℓ_2-norm minimization problem, which is time-saving. Moreover, in order to reduce the contamination of the target signal on the learned background dictionary, we also incorporate the *a priori* target dictionary into the learning process, and each pixel is encoded over the union dictionary that is composed of the background and target dictionaries. Only the coded vector corresponding to the background dictionary is used to update the background dictionary while the target dictionary is fixed. In this way, the purity of the learned background dictionary can be guaranteed.

The main contributions of this article can be briefly summarized as follows.

1) Based on the basic model by which the observed data are decomposed into the background, target, and noise components, the proposed DM exploits the relationships about the background pixels and targets of interest in the scene via the TV and sparsity regularizers, respectively, in order to achieve a satisfactory decomposition, and then the recovered target component is directly employed to perform the target detection.
2) The BDL method based on LLC is proposed. Since the locality can lead to the sparsity and the LLC just requires to solve the ℓ_2-norm minimization problem, which can be implemented very fast in practice. This efficiency augments the practicability of the proposed method in real applications. Moreover, the *a priori* target dictionary is also incorporated into the learning process to reduce the interference of the target signal and ensure that the learned background dictionary does not contain the target spectrum.

The rest of this article is arranged as follows. The proposed method for hyperspectral target detection is elaborated in Section II, including the BDL method and the DM. In Section III, both the simulated and real data experiments are conducted, and the corresponding experimental results are reported and analyzed. Finally, the conclusion is included in Section IV.

II. Proposed Method

In this section, we first present a BDL method based on LLC to learn a complete, compact, and pure background dictionary. Then, given the learned background dictionary and *a priori*

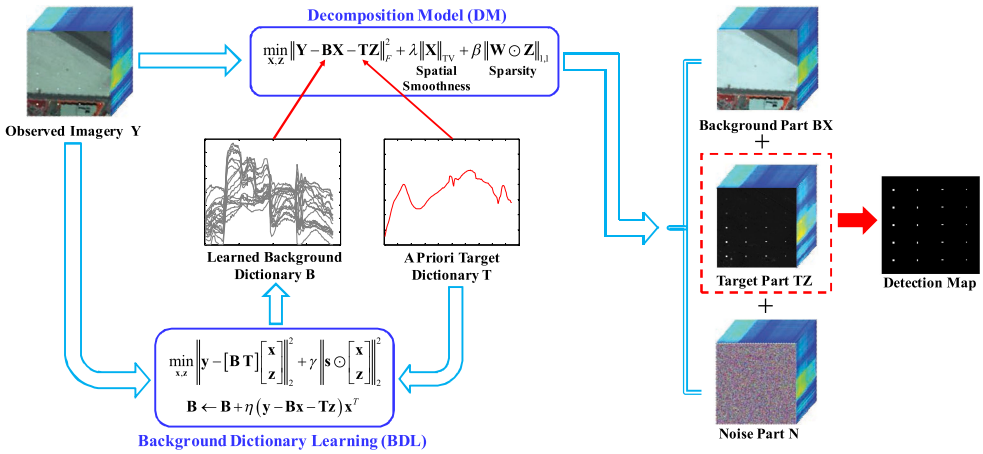


Fig. 1. Schematic flowchart of the proposed DM-BDL for hyperspectral target detection.

target dictionary, the DM is proposed to decompose the hyperspectral imagery into background, target, and noise images. Finally, the detection is accomplished based on the separated target image. Fig. 1 shows the schematic flowchart of the proposed DM-BDL for hyperspectral target detection. In what follows, each part of the proposed method will be introduced in detail.

A. BDL Based on LLC

Suppose that a resized hyperspectral data with n pixels and d spectral bands can be expressed as $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^n \in \mathbb{R}^{d\times n}$, in which each column stands for a pixel vector. Given the observed data $\mathbf{Y}$ and background dictionary $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m] \in \mathbb{R}^{d\times m}$ with m atoms, the sparse coding [26] aims to use as few atoms as possible to approximate the data well. The optimization problem of sparse coding can be written as

$$\min_{\mathbf{B},\{\mathbf{x}_i\}} \sum_{i=1}^{n} \|\mathbf{y}_i - \mathbf{B}\mathbf{x}_i\|_2^2 + \gamma\|\mathbf{x}_i\|_1 \tag{1}$$

where $\mathbf{x}_i \in \mathbb{R}^{m\times 1}$ is the coded vector, the ℓ_1-norm is the sparsity-inducing penalty, and γ is a scalar parameter trading off between the residual and the sparsity. The optimal dictionary $\mathbf{B}$ for sparse coding can be learned iteratively by solving the coded vectors $\{\mathbf{x}_i\}_{i=1}^n$ given the current dictionary and updating the dictionary $\mathbf{B}$ according to the following learning rule:

$$\mathbf{B} \leftarrow \mathbf{B} + \eta\left\langle (\mathbf{y}_i - \mathbf{B}\mathbf{x}_i)\,\mathbf{x}_i^T \right\rangle \tag{2}$$

where η is the step length and $\langle\cdot\rangle$ stands for the average over all the pixels. After the learning step in each iteration, the dictionary atoms are renormalized to have a unit norm to avoid the trivial solution. The learned spectra can gradually match the spectra of endmembers in the scene with a slowly diminishing step length η [26]. However, as mentioned in [28], this method is computationally expensive since it requires to solve the ℓ_1-norm minimization problem. Besides, the learned background dictionary may contain the target spectrum.

Aiming to solve these issues straightly, in the following, we present a novel BDL method based on LLC. To be specific, the LLC solves the ℓ_2-norm minimization problem [33] as

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{B}\mathbf{x}\|_2^2 + \gamma\,\|\mathbf{s}\odot\mathbf{x}\|_2^2 \tag{3}$$

where $\odot$ denotes the Hadamard product, i.e., elementwise product, and $\mathbf{s} \in \mathbb{R}^{m\times 1}$ is the locality adaptor [33] that measures the similarity between the test pixel and each dictionary atom. Here, the Euclidean distance is utilized to calculate the locality adaptor

$$\mathbf{s} = [\|\mathbf{y} - \mathbf{b}_1\|_2, \ldots, \|\mathbf{y} - \mathbf{b}_m\|_2]^T. \tag{4}$$

Intuitively, if s_i (the ith element of $\mathbf{s}$) is small, which means that the ith atom in $\mathbf{B}$ is similar to the pixel vector, and then the corresponding coefficient is allowed to be large. For the atoms that are different from the test pixel, the corresponding coefficients will be enforced to be small. As a result, the locality can lead to sparsity since the coded vector $\mathbf{x}$ is sparse in the sense that there are only few elements with significant values. The solution to (3) has an analytical form as

$$\mathbf{x} = \left(\mathbf{B}^T\mathbf{B} + \gamma\,\mathrm{diag}\left(\mathbf{s}^2\right)\right)^{-1}\mathbf{B}^T\mathbf{y} \tag{5}$$

where $\mathrm{diag}(\mathbf{s}^2)$ returns a square matrix with the elements of $\mathbf{s}^2$ on the diagonal. An important advantage of LLC over sparse coding is that it can be executed very fast in practice. Additionally, in the sparse coding process, quite different atoms might be selected to represent similar pixel vectors for favoring sparsity [33]. However, the locality adaptor in LLC explicitly enforces similar coded vectors for the similar test pixels, and thus, preserving the correlations between the coded vectors.

Furthermore, in order to avoid the target spectrum being learned in the background dictionary, we incorporate the *a priori* target dictionary $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_q] \in \mathbb{R}^{d\times q}$ that is composed of q available target spectra into the learning process. The union dictionary containing the background and target dictionaries is then utilized to represent the test pixel

$$\min_{\mathbf{x},\mathbf{z}} \left\|\mathbf{y} - [\mathbf{B}\mathbf{T}]\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix}\right\|_2^2 + \gamma\left\|\mathbf{s}\odot\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix}\right\|_2^2 \tag{6}$$

Algorithm 1: BDL Based on LLC.

Input: observed data $\mathbf{Y} \in \mathbb{R}^{d\times n}$, *a priori* target dictionary $\mathbf{T} \in \mathbb{R}^{d\times q}$, learning parameter γ, the number of background atoms m.

Initialize: background dictionary $\mathbf{B} \in \mathbb{R}^{d\times m}$ is set to have random positive values.

For $i = 1$ to n, do

1) Calculate the locality adaptor $\mathbf{s} = [\|\mathbf{y}_i - \mathbf{b}_1\|_2, \ldots, \|\mathbf{y}_i - \mathbf{b}_m\|_2, \|\mathbf{y}_i - \mathbf{t}_1\|_2, \ldots, \|\mathbf{y}_i - \mathbf{t}_q\|_2]^T$.
2) Solve the coded vector $[\mathbf{x}^T\ \mathbf{z}^T]^T = ([\mathbf{B}\ \mathbf{T}]^T[\mathbf{B}\ \mathbf{T}] + \gamma diag(\mathbf{s}^2))^{-1}[\mathbf{B}\ \mathbf{T}]^T\mathbf{y}_i$.
3) Update the background dictionary $\mathbf{B} \leftarrow \mathbf{B} + \eta(\mathbf{y}_i - \mathbf{T}\mathbf{z} - \mathbf{B}\mathbf{x})\mathbf{x}^T$, $\mathbf{B} = \max(\mathbf{B}, \mathbf{0})$.
4) Renormalize each column of $\mathbf{B}$ to have a unit norm.
5) Slowly diminish the step length η.

end for

Output: learned background dictionary $\mathbf{B}$.

where $\mathbf{x} \in \mathbb{R}^{m\times 1}$ and $\mathbf{z} \in \mathbb{R}^{q\times 1}$ are the representation vectors associated with the background and target dictionaries, respectively, and $\mathbf{s} \in \mathbb{R}^{(m+q)\times 1}$ is now the similarity between the test pixel and each atom in the union dictionary. In the learning process, only the background dictionary is updated while the target dictionary is fixed.

In practice, it is memory-consuming to hold all the coded vectors together in each iteration when the number of pixels n is large. In our implementation, we handle a single pixel $\mathbf{y}$ in each iteration and incrementally update the background dictionary $\mathbf{B}$. The learning rule for updating the background dictionary can be written as

$$\mathbf{B} \leftarrow \mathbf{B} + \eta\left(\mathbf{y} - \mathbf{T}\mathbf{z} - \mathbf{B}\mathbf{x}\right)\mathbf{x}^T. \tag{7}$$

Note that here the input can be viewed as $\mathbf{y} - \mathbf{T}\mathbf{z}$ with the target signal being removed. Therefore, the contamination of targets on the learned background dictionary can be well reduced. Consequently, a complete, compact, and pure background dictionary can be obtained.

The detailed procedure of the proposed BDL algorithm is described in Algorithm 1, in which the dictionary elements are enforced to be nonnegative to maintain the physical meaning in accordance with spectral reflectance [26]. It is worth mentioning that the iteration number is exactly the total number of pixels *n*. The discussion and concrete design for a sequence of diminishing step length η will be given in Section III.

B. Decomposition Model

According to the replacement signal model [35], for a subpixel target, in which the known target $\mathbf{t}$ replaces a fraction α of the background $\mathbf{b}$, the observed pixel vector $\mathbf{y}$ can be expressed as

$$\mathbf{y} = (1-\alpha)\,\mathbf{b} + \alpha\mathbf{t} + \mathbf{n}, 0 < \alpha \le 1 \tag{8}$$

where $\mathbf{n}$ is the additive noise. For a background pixel in which the target is absent ($\alpha = 0$), the observed spectrum can be written as $\mathbf{y} = \mathbf{b} + \mathbf{n}$. Since the background and target components of all the pixels can be linearly represented by some common background and target spectra, respectively, the replacement model in (8) can be rewritten in a representation-based form

$$\begin{aligned} H_0 &: \mathbf{y} = \mathbf{B}\mathbf{x} + \mathbf{T}\cdot\mathbf{0} + \mathbf{n}, \text{ (Target absent)} \\ H_1 &: \mathbf{y} = \mathbf{B}\mathbf{x}' + \mathbf{T}\mathbf{z}' + \mathbf{n}, \text{(Target present)} \end{aligned} \tag{9}$$

where $\mathbf{x}' = (1-\alpha)\mathbf{x}$ and $\mathbf{z}' = \alpha\mathbf{z}$. In the following, the superscript will be omitted for simplicity. Taking all the pixels into consideration and writing (9) in a unified form as

$$\mathbf{Y} = \mathbf{B}\mathbf{X} + \mathbf{T}\mathbf{Z} + \mathbf{N} \tag{10}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ and $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n]$ concatenate the representation vectors corresponding to the background and target dictionaries, respectively.

In order to achieve a satisfactory separation of the background, target, and noise components from the observed data, the characteristics of these components in hyperspectral imagery should be properly exploited. Since the background is spatially smooth in the sense that the neighboring pixels usually consist of similar materials, and, therefore, if two background pixels $\mathbf{y}_i$ and $\mathbf{y}_j$ are adjacent in spatial domains, their representation vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ associated with the background dictionary $\mathbf{B}$ are expected to be similar. The TV regularizer [36]–[39], which is effective in preserving the piecewise smoothness, provides an appropriate means to characterize the background part

$$TV\left(\mathbf{X}\right) = \sum_i \|\mathbf{x}_i - \mathbf{x}_{i_h}\|_1 + \|\mathbf{x}_i - \mathbf{x}_{i_v}\|_1. \tag{11}$$

Intuitively, the TV term sums up the differences between the representation vector of each pixel and those of its spatial neighbors (in both horizontal and vertical directions). For the convenience of subsequent calculation, two linear operators $\mathbf{H}_h$ and $\mathbf{H}_v$ are introduced to simplify the TV term in (11). Specifically, $\mathbf{H}_h$ is designed such that $\mathbf{H}_h\mathbf{X} = [\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_n]$, in which $\mathbf{d}_i = \mathbf{x}_i - \mathbf{x}_{i_h}$ is the difference between the representation vectors of the *i*th pixel and its horizontal neighbor. The operator $\mathbf{H}_v$ computing the vertical differences can be designed in a similar way. Here, the boundaries are assumed to be periodic [39]. In this way, the TV term in (11) can be written in a compact form

$$TV\left(\mathbf{X}\right) = \left\| \begin{bmatrix} \mathbf{H}_h\mathbf{X} \\ \mathbf{H}_v\mathbf{X} \end{bmatrix} \right\|_{1,1} = \|\mathbf{H}\mathbf{X}\|_{1,1} \tag{12}$$

where $\|\cdot\|_{1,1}$ is defined as the sum of ℓ_1-norm of each column in a matrix.

On the other hand, the representation vector $\mathbf{z}$ corresponding to the target dictionary $\mathbf{T}$ is null when the target is absent and has nonzero elements when the target is present according to (9). Since the targets of interest usually occupy a few pixels in the whole scene [40], $\mathbf{Z}$ is actually a sparse matrix in which there are

very few entries with significant values. Obviously, for a pixel with high target fraction α, the corresponding column in $\mathbf{Z}$ has greater values. Therefore, the columns of $\mathbf{Z}$ naturally measure the degree of a pixel belonging to the target. In order to enable this scheme to work, we further incorporate a weight matrix $\mathbf{W}$ to suppress the false alarms that reside in $\mathbf{Z}$ and highlight the response of true targets depending on the distances of the atoms in target dictionary $\mathbf{T}$ from each pixel vector

$$[\mathbf{W}]_{i,j} = \|\mathbf{y}_j - \mathbf{t}_i\|_2. \tag{13}$$

Finally, the additive noise $\mathbf{N}$ is assumed to be densely distributed and follow the identically and independently Gaussian distributions. The mean-squared error is utilized to model the noise. In consideration of the above-stated characteristics regarding the background, target, and noise components, the proposed DM can be formulated as

$$\min_{\mathbf{X},\mathbf{Z}} \|\mathbf{Y} - \mathbf{B}\mathbf{X} - \mathbf{T}\mathbf{Z}\|_F^2 + \lambda\|\mathbf{H}\mathbf{X}\|_{1,1} + \beta\|\mathbf{W} \odot \mathbf{Z}\|_{1,1} \tag{14}$$

where $\|\cdot\|_F$ is the Frobenius norm, λ and β control the spatial smoothness of $\mathbf{X}$ and the sparsity level in $\mathbf{Z}$, respectively. It is expected that a satisfactory decomposition can be achieved by DM since the intrinsic characteristics of the background and target parts have been explicitly imposed. After the decomposition, the separated target component $\mathbf{TZ}$ is directly used to accomplish the target detection. The degree of the ith pixel being claimed to be the target can be determined by the magnitude of the ith column of $\mathbf{TZ}$ as

$$\left\|[\mathbf{TZ}]_{:,i}\right\|_2 = \sqrt{\sum_{j=1}^{d}\left([\mathbf{TZ}]_{j,i}\right)^2}. \tag{15}$$

It is expected to alternately perform the BDL and DM to achieve more accurate detection results. However, in view of the computational burden, we only perform one round of BDL and DM since it is sufficient to obtain satisfactory results.

In what follows, we solve the optimization problem in (14) via the alternating direction method of multipliers [41]. In order to make the objective function separable, three auxiliary variables $\mathbf{V}_1$, $\mathbf{V}_2$, and $\mathbf{V}_3$ are introduced. The problem (14) is then converted to the following equivalent formulation:

$$\min_{\{\mathbf{V}\},\mathbf{X},\mathbf{Z}} \|\mathbf{Y} - \mathbf{B}\mathbf{X} - \mathbf{T}\mathbf{Z}\|_F^2 + \lambda\|\mathbf{V}_2\|_{1,1} + \beta\|\mathbf{W} \odot \mathbf{V}_3\|_{1,1},$$
$$\text{s.t. } \mathbf{V}_1 = \mathbf{X}, \mathbf{V}_2 = \mathbf{H}\mathbf{V}_1, \mathbf{V}_3 = \mathbf{Z}. \tag{16}$$

The augmented Lagrangian formulation for the above-mentioned optimization problem can be written as

$$\begin{aligned}\min_{\{\mathbf{V}\},\mathbf{X},\mathbf{Z}} & \|\mathbf{Y} - \mathbf{B}\mathbf{X} - \mathbf{T}\mathbf{Z}\|_F^2 + \lambda\|\mathbf{V}_2\|_{1,1} + \beta\|\mathbf{W} \odot \mathbf{V}_3\|_{1,1} \\ & + \frac{\mu}{2}\Big(\|\mathbf{V}_1 - \mathbf{X} - \mathbf{D}_1\|_F^2 + \|\mathbf{V}_2 - \mathbf{H}\mathbf{V}_1 - \mathbf{D}_2\|_F^2 \\ & + \|\mathbf{V}_3 - \mathbf{Z} - \mathbf{D}_3\|_F^2\Big)\end{aligned} \tag{17}$$

where $(\mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3)$ are the Lagrange multipliers and μ is a positive penalty parameter. The minimization of this multiple-variable problem can be achieved by alternately minimizing the objective function with respect to one variable while fixing other variables. Since the objective function in (16) is closed, proper, and convex, the convergence of the solution can be guaranteed [41].

Algorithm 2: Optimization Procedure for DM.

Input: observed data $\mathbf{Y} \in \mathbb{R}^{d\times n}$, learned background dictionary $\mathbf{B}$ and *a priori* target dictionary $\mathbf{T}$, and tradeoff parameters λ and β.

Initialize: all variables $\mathbf{X}$, $\mathbf{Z}$, $\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{V}_3$, $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{D}_3$ are set to zero matrices, $\mu_0 = 1e^{-3}$, $\mu_{\max} = 1e^{10}$, $\rho = 1.2$, $\varepsilon = 1e^{-4}$, iter $= 0$.

While not converged, do

1) Update $\mathbf{X}$ and $\mathbf{Z}$
$\mathbf{X} = (2\mathbf{B}^T\mathbf{B} + \mu\mathbf{I})^{-1}(2\mathbf{B}^T(\mathbf{Y} - \mathbf{T}\mathbf{Z}) + \mu(\mathbf{V}_1 - \mathbf{D}_1))$
$\mathbf{Z} = (2\mathbf{T}^T\mathbf{T} + \mu\mathbf{I})^{-1}(2\mathbf{A}^T(\mathbf{Y} - \mathbf{B}\mathbf{X}) + \mu(\mathbf{V}_3 - \mathbf{D}_3))$.
2) Update $\mathbf{V}_1$, $\mathbf{V}_2$, and $\mathbf{V}_3$
$\mathbf{V}_1 = (\mathbf{H}^T\mathbf{H} + \mathbf{I})^{-1}(\mathbf{H}^T(\mathbf{V}_2 - \mathbf{D}_2) + \mathbf{X} + \mathbf{D}_1)$
$\mathbf{V}_2 = \text{soft}(\mathbf{H}\mathbf{V}_1 + \mathbf{D}_2, \lambda/\mu)$
$\mathbf{V}_3 = \text{soft}(\mathbf{Z} + \mathbf{D}_3, (\beta/\mu) \cdot \mathbf{W})$.
3) Update Lagrange multipliers and penalty parameter
$\mathbf{D}_1 = \mathbf{D}_1 - (\mathbf{V}_1 - \mathbf{X})$,
$\mathbf{D}_2 = \mathbf{D}_2 - (\mathbf{V}_2 - \mathbf{H}\mathbf{V}_1)$,
$\mathbf{D}_3 = \mathbf{D}_3 - (\mathbf{V}_3 - \mathbf{Z})$, $\mu = \min(\rho\mu, \mu_{\max})$.
4) Check for convergence
$\|\mathbf{V}_1 - \mathbf{X}\|_F + \|\mathbf{V}_2 - \mathbf{H}\mathbf{V}_1\|_F + \|\mathbf{V}_3 - \mathbf{Z}\|_F < \varepsilon$.
5) iter $\leftarrow$ iter $+ 1$.

end while

Output: an optimal solution $(\mathbf{X}^*, \mathbf{Z}^*)$.

Algorithm 2 outlines the step-by-step optimization procedure for the proposed DM. Note that the update rule for $\mathbf{V}_1$ can be applied for each band in an independent manner, and the discrete Fourier transform diagonalization is adopted for efficient implementation [39]. In the update rules for $\mathbf{V}_2$ and $\mathbf{V}_3$, $soft(\cdot, \cdot)$ represents the shrinkage operator [42]. The MATLAB implementations for the proposed Algorithm 1 and Algorithm 2 have been made available online.[1]

III. Experimental Results

In this section, we conduct both the simulated and real data experiments to evaluate the performance of the proposed DM-BDL in hyperspectral target detection. Specifically, one simulated hyperspectral dataset is generated and used to give a comprehensive analysis of the proposed DM-BDL, while two real-world hyperspectral datasets are utilized to validate its effectiveness in real scenarios. The receiver operating characteristic (ROC) curve [43] is employed to evaluate the detection performance quantitatively, which describes the relationship between the probability of detection and the false alarm rate. The ROC curve closer to the upper left corner indicates the better

[1][Online]. Available: https://github.com/FDU-ctk/HSI-detection

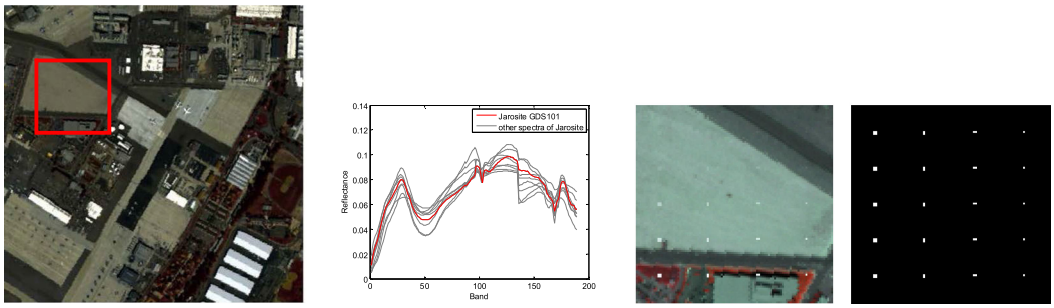


Fig. 2. Simulated hyperspectral dataset. (a) False color image of the whole scene. (b) Buried target spectrum $\mathbf{t}$ (Jarosite GDS101), and the other eight spectra of the jarosite mineral. (c) False color image of the simulated data. (d) Ground-truth map for the embedded targets.

detection result. The area under the ROC curve (AUC) [44] that provides an intuitive and quantitative evaluation of detection performance is also considered. The experimental platform is a computer with an eight-core Intel CPU Xeon CPU 3.07-GHz and 3.06-GHz processors, 48-GB memory, and 64-bit Windows 10. All the experiments are executed in the environment of MATLAB R2013b.

A. Simulated Data Experiments

In this section, a simulated dataset is generated based on the real hyperspectral imagery, which was collected by the airborne visible/infrared imaging spectrometer over San Diego airport in California, USA. It consists of 224 spectral channels spanning wavelengths from 370 to 2510 nm. The spatial resolution is 3.5 m per pixel. Before making some analysis with this imagery, the spectral bands with indices 1–6, 33–35, 97, 107–113, 153–166, and 221–224 are removed due to low SNR or water vapor absorption. As a result, a total of 189 spectral bands are left. The whole imagery has a size of 400 $\times$ 400 pixels, as shown in Fig. 2(a). A 100 $\times$ 100 region (pixels in rows 71–170 and columns 41–140) is selected from the original image (encircled by the red square) to construct the simulated data. The targets are embedded according to the replacement signal model in (8). The target spectrum $\mathbf{t}$ is selected from the U. S. Geological Survey (USGS) digital spectral library,[2] in which each kind of material contains a number of similar spectral signatures, and well illustrate the spectral variability. A sulfate mineral known as "jarosite" contains nine spectra and the fourth Jarosite mineral (Jarosite GDS101) is considered as the buried target $\mathbf{t}$ in our experiment. The other eight spectra of the jarosite mineral will be used to investigate the detection performance when inaccurate prior target spectrum is provided. Fig. 2(b) shows these spectra of the jarosite mineral.

In the simulated data, 20 target blocks distributed in five rows and four columns have been embedded. The sizes of the targets in each column are 2 $\times$ 2, 2 $\times$ 1, 1 $\times$ 2, and 1 $\times$ 1 from left to right, respectively. The fill fraction α remains unchanged for the same row and takes the value of 0.05, 0.1, 0.3, 0.5, and 0.8 from top to bottom, respectively. Finally, the white Gaussian noise is added to the simulated data with SNR $=$ 40 dB. The SNR is defined as follows:

$$SNR \equiv 10\log_{10}\frac{\mathbb{E}\left[\mathbf{y}^T\mathbf{y}\right]}{\mathbb{E}\left[\mathbf{e}^T\mathbf{e}\right]} \tag{18}$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator, $\mathbf{y}$ represents the pixel vector, and $\mathbf{e}$ is the additive noise. The false color image of the simulated data as well as the ground-truth map for the embedded targets is given in Fig. 2(c) and (d), respectively.

In the following subsection, the spectrum of Jarosite GDS101 is served as the *a priori* target dictionary ($q = 1$). The generated simulated data is first used to test the proposed BDL method and the DM. Then, the detection performance of the proposed DM-BDL is compared with several classical and state-of-the-art target detectors. Finally, the other eight jarosite spectra are used one by one to investigate the detection performance under the circumstance of inaccurate prior target information (spectral variability of target).

1) Analysis of the Proposed BDL Method: There are several parameters that need to be set in the proposed BDL method, including the diminishing step length η, the learning parameter γ, and the number of atoms m. Fortunately, a satisfactory learning result can be generally obtained with a series of slow descent η [27]. In our experiment, the initial η is set to 5, and we use $\eta \leftarrow \sigma\eta$ for every 50 iterations (pixels) to decrease the step length. The decay parameter σ is set adaptively to the data size such that the ultimate η is about 0.001 at the end of the iteration. Consequently, we set σ for the generated simulated data with $1e^4$ pixels to 0.96. The number of background atoms m is crucial to the learning algorithm. A small m cannot guarantee the completeness of the learned dictionary to comprise the whole scene, while a large m may result in some noise spectra being learned and also the computational burden in a later decomposition process. Fig. 3 illustrates the learning results when jointly taking m and γ into consideration. Here, the detection result obtained by DM-BDL in terms of the AUC score is used to evaluate the learning result, and the tradeoff parameters λ and β in the DM are, respectively, fixed at 0.01 and 0.1, which will be analyzed in the following section. It is apparent that the proposed BDL method can achieve stable and satisfactory results when m is larger than ten and γ greater than five. For simplicity, in the

[2][Online]. Available: http://speclab.cr.usgs.gov/spectral-lib.html

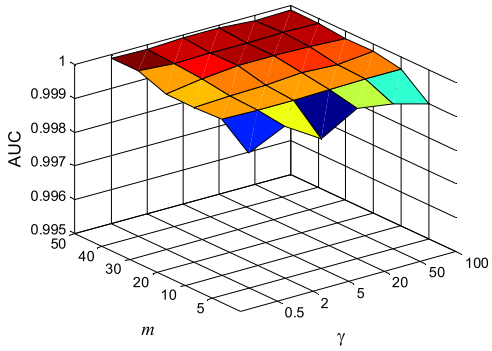


Fig. 3. Detection performance of the proposed DM-BDL under different combinations of m and γ in the BDL algorithm with the simulated dataset.

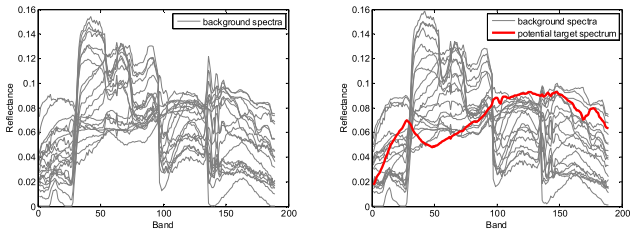


Fig. 4. Learned spectra obtained by the proposed BDL method. (a) With the target dictionary. (b) Without the target dictionary in the learning process.

following experiments, m and γ are empirically fixed at 20 and 20, respectively.

Then, we investigate the effect of the target dictionary that is incorporated in the learning process. To illustrate this, we plot the learned background spectra with and without the target dictionary, as shown in Fig. 4(a) and (b), respectively. It can be observed that a potential target spectrum is learned without the suppression of target contamination induced by the target dictionary. We utilize the spectral information divergence (SID) [45] to quantitatively describe how similar each learned spectrum is to the buried target spectrum $\mathbf{t}$. The closer the SID value is to zero, the higher the similarity of two spectra. The minimum SID values in regard to the results in Fig. 4(a) and (b) are 0.0310 and 0.0081, respectively. These facts indicate that the incorporated target dictionary in the learning process indeed can reduce the interference caused by target samples and thereby ensure the purity of the learned background dictionary.

2) Analysis of the Proposed DM: In this subsection, we first investigate the separated background and target components when the tradeoff parameters λ and β in the DM are set to 0.01 and 0.1, respectively. The false color shows for the rearranged background and target images are displayed in Fig. 5(a) and (b), respectively. It can be seen from Fig. 5(a) that the textures have been well preserved in the recovered background image. Several darker blocks appear in Fig. 5(a) since the filled target fractions have been removed, leaving the dimmer fraction of the background. In Fig. 5(b), the target fractions have been almost entirely extracted from the imagery and deposited in the target image with very few false alarms.

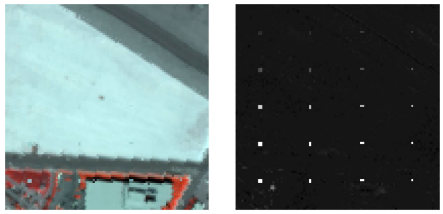


Fig. 5. False color shows for the separated (a) background image and (b) target image by the DM when $\lambda = 0.01$ and $\beta = 0.1$.

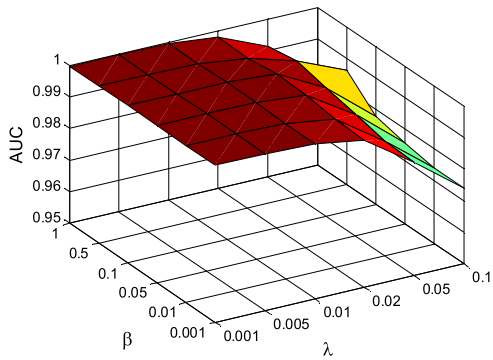


Fig. 6. Detection performance variation of the proposed DM-BDL under different combinations of tradeoff parameters β and λ with the simulated dataset.

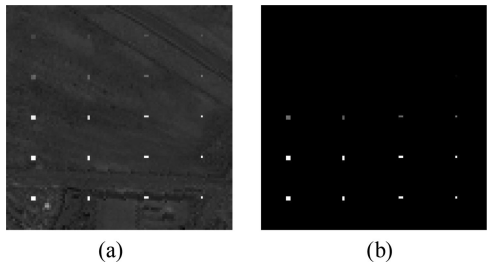


Fig. 7. False color shows for the separated target images by the DM with $\lambda = 0.01$ and (a) $\beta = 0.01$ and (b) $\beta = 1$.

Next, we analyze the influence of λ and β on the decomposition result. The detection result based on the target component is considered to evaluate the separation result. Fig. 6 illustrates that when λ takes relatively large values, the performance shows an apparent decline. This can be explained by the fact that the recovered background image is oversmoothing, and some textures may be smoothed out and deposited in the target image, thus resulting in a large number of false alarms. Another observation is that the separation result seems unchanged under different values of β. In order to better illustrate the role of sparsity that is imposed on the target component, Fig. 7(a) and (b) shows the separated target images obtained by the DM under $\beta = 0.01$ and $\beta = 1$ when λ is fixed at 0.01, respectively. There are some false alarms deposited in the target image caused by the background and noise when β is small. On the other hand, when β takes a large value, i.e., $\beta = 1$, some targets with low fractions, such as $\alpha = 0.05$ and $\alpha = 0.1$ may be further removed to favor sparsity

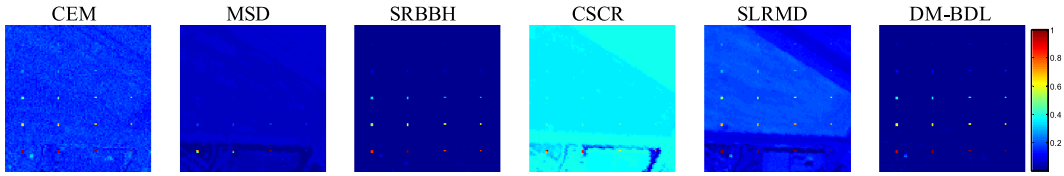


Fig. 8. 2-D plots of detection results obtained by different target detectors on the simulated data.

in the separated target image. It is worth noting that the detection result remains satisfactory since the false alarms have been fully suppressed, and thus the targets with very weak response in the separated target image can still be detected. Despite this fact, a moderate value of β is desirable to avoid the possible missed detections.

According to the above-mentioned analysis, we empirically set $\lambda = 0.01$ and $\beta = 0.1$ in the DM throughout the following experiments for the consideration of convenience and simplicity since the parameter tuning is usually a computationally expensive task.

3) Detection Performance Compared With Other Hyperspectral Target Detectors: Here, we compare the detection performance of the proposed DM-BDL for hyperspectral imagery with several classical and state-of-the-art target detectors: CEM, MSD, SRBBH, CSCR, and SLRMD. The CEM and MSD are two conventional detectors. The SRBBH and CSCR are two typical representation-based detectors, in which the background dictionary is generated via the dual windows for each test pixel. In the SLRMD, the separation of target and background components is established on the sparse and low-rank matrix decomposition.

For the compared detectors, the involved parameters are adjusted to achieve the best possible detection accuracy [17], [18], [21]. In the MSD implementation, the background and target subspaces are generated using the eigenvectors corresponding to the significant eigenvalues of the covariance matrices obtained from the atoms in background and target dictionaries $\mathbf{B}$ and $\mathbf{T}$, respectively [46]. In the case of SRBBH, the sizes of the dual windows $(w_{\text{out}}, w_{\text{in}})$ are set to $(7, 5)$ and the orthogonal matching pursuit [47] algorithm is utilized to recover the sparse vectors with the sparsity level of 10. For the CSCR, the setting for the dual windows $(w_{\text{out}}, w_{\text{in}})$ is consistent with that in SRBBH, and the balance parameters λ_1 for the sparsity under H_1 and λ_2 for the collaborative representation under H_0 are set to 0.1 and 1, respectively [18]. In SLRMD, the tradeoff parameters τ and λ that constrain the low rankness and sparsity are, respectively, set to 0.5 and 0.05 through global searching [21].

Fig. 8 shows the 2-D plots of the detection results obtained by the proposed DM-BDL as well as the compared target detectors. It is obvious that all the detectors can distinguish the target samples at the bottom of the scene, i.e., the targets with high fractions $\alpha = 0.8$. However, the effectiveness of a specific detector mainly relies on its ability to detect the targets with low fractions, such as $\alpha = 0.05$ and $\alpha = 0.1$, which is more challenging. The proposed DM-BDL provides a clear discrimination between all the embedded targets and the background. The SRBBH achieves a comparable detection result, which is owing to that the background dictionary generated via the dual windows is adaptive to the local statistics and therefore effective in suppressing the background. However, the SRBBH is unable to distinguish the targets with low fractions well. The response of all the targets in the detection maps of SLRMD is high, whereas that of background pixels also exhibits high values since the recovery of the background component is unsatisfactory without utilizing an appropriate background dictionary. The quantitative assessments of the detection results by means of ROC curves are presented in Fig. 9. Compared with the SLRMD, the proposed DM-BDL shows a slightly lower probability of detection when the false alarm rate is very low. Nevertheless, our method performs the best in terms of the overall detection accuracy and achieves a much lower false alarm rate than other detectors when the probability of detection reaches 100%, which should indicate the effectiveness of the proposed DM-BDL. The corresponding AUC scores are reported in Table I, from which a similar conclusion can be drawed.

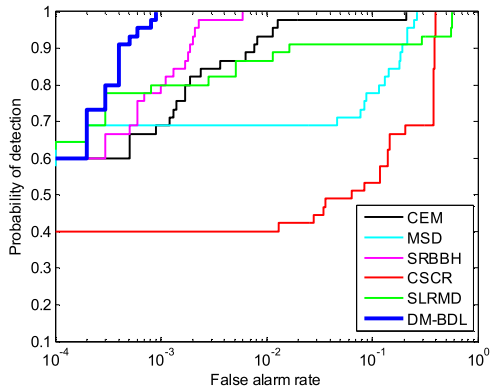


Fig. 9. ROC curves obtained by different target detectors on the simulated data.

TABLE I
AUC SCORES OF THE DETECTION RESULTS REPORTED IN FIG. 9

Algorithm	CEM	MSD	SRBBH	CSCR	SLRMD	DM-BDL
AUC	0.9939	0.9514	0.9993	0.8513	0.9656	**0.9998**

In the following, we randomly change the positions of the embedded targets and repeat the preceding experiment for 20 times under SNR = 40 dB to investigate the stability of all the compared detectors. Table II reports the average values and the standard deviations of AUC scores obtained by different target detectors. The algorithms based on local background dictionary, i.e., SRBBH and CSCR, exhibit a large deviation of the detection accuracy since the issue of target contamination

TABLE II
AUC SCORES AND EXECUTION TIME FOR DIFFERENT TARGET DETECTORS BASED ON 20 RUNS USING THE SIMULATED DATA WITH RANDOMLY EMBEDDED TARGETS

Algorithm	CEM	MSD	SRBBH	CSCR	SLRMD	DM-BDL
AUC	0.9892±0.0055	0.8940±0.0584	0.9783±0.0318	0.8618±0.0216	0.9234±0.0462	**0.9992±0.0008**
Time (s)	0.06±0.01	3.94±0.03	30.65±0.02	45.08±0.08	38.27±0.20	10.61±0.20

TABLE III
AUC SCORES OBTAINED BY DIFFERENT TARGET DETECTORS WITH EACH INACCURATE SPECTRUM OF JAROSITE MINERAL IN THE USGS SPECTRAL LIBRARY

Algorithm	CEM	MSD	SRBBH	CSCR	SLRMD	DM-BDL
AUC	0.9161±0.0332	0.9332±0.0411	0.9891±0.0076	0.7566±0.0905	0.9604±0.0096	**0.9995±0.0006**

in the background dictionary is inevitable when the targets reside in close proximity. It can be seen from Table II that the proposed DM-BDL shows a much more stable result, which can be attributed to two aspects: first, the recovery of background and target components is performed on the whole scene with the relationships of background pixels and targets being explicitly considered; second, compared with the local background dictionary generated by dual windows, the proposed BDL algorithm is not affected by the distribution of targets. Furthermore, the execution time of all the detectors is also listed in Table II. Note that the reported result for the proposed DM-BDL includes the time spent on the BDL (3.50 s) and that on the DM (7.11 s). It can be observed that the DM-BDL takes less time than the state-of-the-art target detectors. This efficiency results from two points: first, based on LLC, the BDL algorithm can be performed very fast; second, compared with the case of overcomplete background dictionary where a heavy computational cost is spent on the background recovery due to a large redundant dictionary, the learned background dictionary is complete and of small size (compact), and thus, accelerating the decomposition process.

4) Robustness to the Prior Target Information: In real scenarios, the accuracy of the prior target information cannot be guaranteed due to the inherent variability of spectral signatures. Therefore, it makes sense to investigate the robustness of the detection result of a specific target detector to an inaccurate prior target spectrum. As described above, there are another eight spectra corresponding to the jarosite mineral in the USGS spectral library. We repeat the preceding experiment for eight trials. In each trial, one out of eight jarosite spectra is taken as available target specturm to form the *a priori* target dictionary.

The average values and the standard deviations of AUC scores obtained by different detectors are given in Table III. Comparing the results of CEM in Tables I and III, it can be concluded that CEM can achieve a satisfactory result under exact prior knowledge of targets while shows a significant degradation when inaccurate prior target spetrum is provided. The performance of MSD is unsatisfactory, which may due to the generated dictionaries composed of eigenvectors cannot represent the background and target well. The CSCR exhibits a greater variation than SRBBH. It is expected since only target training samples are used to represent the test pixel under H_1 in CSCR, which makes the CSCR sensitive to the prior target information. The proposed DM-BDL and SLRMD are both based on the separation of background and target components from the whole scene. The SLRMD shows a large deviation under inaccurate prior target spectrum since the recovery of the background is not satisfactory. However, in the case of DM-BDL, the background estimation is achieved by the representation of the learned dictionary and the intrinsic characteristics of background and target parts have been properly exploited, making the recovery process of background and target components more stable and less affected by the spectral variability.

(a) (b) (c)

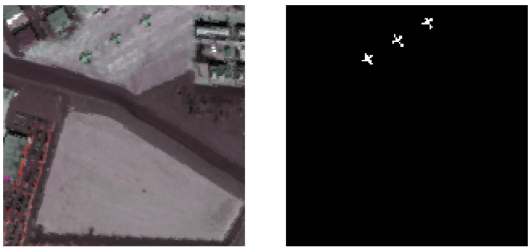
Fig. 10. HYDICE urban dataset. (a) False color image of the whole image. (b) False color image of the detection area. (c) Ground-truth map for the targets of interest.

B. Real Data Experiments

In this section, two widely used real-world hyperspectral datasets are utilized to evaluate the detection performance of the proposed DM-BDL in practical situations.

The first real dataset was collected by the hyperspectral digital imagery collection experiment (HYDICE) sensor over an urban area. The spatial resolution is about 1 m per pixel. The imagery has a spectral resolution of 10 nm and consists of 210 spectral bands in wavelengths ranging from 400 to 2500 nm. The low-SNR and water vapor absorption bands (1-4, 76, 87, 101-111, 136-153, and 198-210) are discarded, leaving 162 bands. The whole image contains 307 × 307 pixels and a region of interest with 80 × 100 pixels (rows 1–80 and columns 188–287) is cropped to perform the target detection. The whole image scene and the selected region are displayed in Fig. 10(a) and (b), respectively. In this dataset, the targets to be detected are the vehicles with 19 pixels, as shown in Fig. 10(c).

The second data used in the real experiments are also a portion of the aforementioned San Diego imagery. The region in the upper left corner with the size of 150 × 150 pixels is chosen as the test imagery. The targets of interest in this dataset are



(a) (b)

Fig. 11. Real San Diego dataset. (a) False color image of the detection region. (b) Ground-truth map for the targets of interest.

three airplanes with 62 pixels to be detected. There are full-pixel targets in the airplanes as well as some subpixel targets on the edge of the airplanes. The false color image and the ground-truth map for the targets are provided in Fig. 11(a) and (b), respectively.

For the HYDICE dataset, the *a priori* target dictionary is collected from two target samples with locations (21, 79) and (16, 87). In the BDL process, the decay parameter in $\eta \leftarrow \sigma\eta$ is set to 0.95 according to the adaptive setting mentioned in Section III-A1. The spectra of the targets of interest as well as those in the learned background dictionary are plotted in Fig. 12(a) and (c), respectively. In the CEM implementation, the single input target spectrum is obtained by averaging the prior target spectra. For the SRBBH, the dual windows $(w_{\text{out}}, w_{\text{in}})$ are set to (11, 7) and the sparsity level is 4. In the case of CSCR, the parameters λ_1 and λ_2 are 0.001 and 0.1, respectively. As for SLRMD, we set the parameters τ and λ to 1 and 0.01, respectively, after parameter tuning. The 2-D plots of the detection results obtained by all the detectors are shown in Fig. 13(a). Obviously, the response of targets exhibits a high contrast with the background in the detection result of the proposed DM-BDL, which indicates that the signals that reside in the subspace spanned by the target dictionary have been well separated and deposited in the separated target image. The false alarms in the detection result of DM-BDL are much less than those of SLRMD, which is owing to the learned dictionary is effective in representing the background. Note that the detection map of CSCR seems quite different from others. This is due to that the decision of CSCR is achieved by computing the difference between the residuals of CR and SR. Some pixels, which have very high response values in SR, will show extremely low response values (negative) in the final detection result. Thus, most pixels exhibit relatively high response values after a linear stretch to [0, 1]. Nevertheless, for the true targets, their response values are generally higher and the detection result is acceptable. To quantitatively compare the results of these detectors, the ROC curves and the AUC scores are provided in Fig. 14(a) and Table IV, respectively. The highest probability of detection is always achieved by our method over other target detectors under all values of the false alarm rate. The AUC scores also demonstrate that the proposed DM-BDL performs the best among the compared detectors.

For the real San Diego dataset, we select one pixel in the center of each plane as the available target spectra to form the *a priori* target dictionary with $q = 3$. The spectra of the targets to be detected and those in the learned background dictionary are displayed in Fig. 12(b) and (d), respectively. Note that the decay parameter σ is set to 0.982 for this dataset. The dual windows $(w_{\text{out}}, w_{\text{in}})$ and the sparsity level in SRBBH are set to (13, 11) and 10, respectively. For the CSCR, the parameters λ_1 and λ_2 are set to 0.001 and 0.1, respectively. In the case of SLRMD, the optimal setting for the tradeoff parameters τ and λ is 1 and 0.01, respectively. With these configurations of the involved parameters in each target detector, the detection results can be obtained in Fig. 13(b). The detection performance of SRBBH and CSCR that are based on dual windows is unsatisfactory due to that the targets in this dataset are irregular and reside closely to each other. The targets in the detection result of DM-BDL are more distinct compared with SLRMD since the background has been well represented and removed from the scene. The quantitative assessments in terms of the ROC curves and AUC scores are provided in Fig. 14(b) and Table IV, respectively. Clearly, there is a large margin between the ROC curve of the proposed DM-BDL and those of the compared detectors. When the probability of detection reaches 100%, the false alarm rate of our method is about 0.007, which is the lowest among all the detectors. The AUC scores also confirm that the proposed method is very promising in detecting the targets of interest in hyperspectral imagery.

TABLE IV
AUC SCORES AND EXECUTION TIME OF DIFFERENT TARGET DETECTORS ON THE TWO REAL DATASETS

Algorithm		CEM	MSD	SRBBH	CSCR	SLRMD	DM-BDL
HYDICE	AUC	0.9771	0.9442	0.9490	0.9649	0.9974	**0.9997**
	Time (s)	0.03	3.05	9.09	22.29	22.09	8.98
San Diego	AUC	0.9126	0.9702	0.9480	0.9916	0.9928	**0.9992**
	Time (s)	0.13	9.01	71.59	206.28	122.15	17.74

The execution time on the two real datasets is also provided in Table IV. The reported time cost for the DM-BDL comprises the portion on the BDL (2.78 s for HYDICE and 8.13 s for San Diego) and the rest on the DM. Compared with the representation-based detectors, i.e., SRBBH and CSCR, and SLRMD that is based on the decomposition of the whole scene, our method shows an improved efficiency, particularly when the observed data are of large size, that is, the real San Diego dataset.

C. Summary

The extensive experiments have been carried out on both simulated and real hyperspectral datasets. According to the experimental results, the main advantages of our proposed DM-BDL can be summarized as follows.

1) *Effectiveness*: The DM decomposes the observed data into background, target, and noise components. Based on the complete and pure background dictionary obtained with the proposed BDL algorithm, a satisfactory decomposition result can be achieved by DM with the characteristics of these components being explicitly considered. Then, based on the recovered target image, the proposed DM-BDL is effective in identifying the targets, especially for

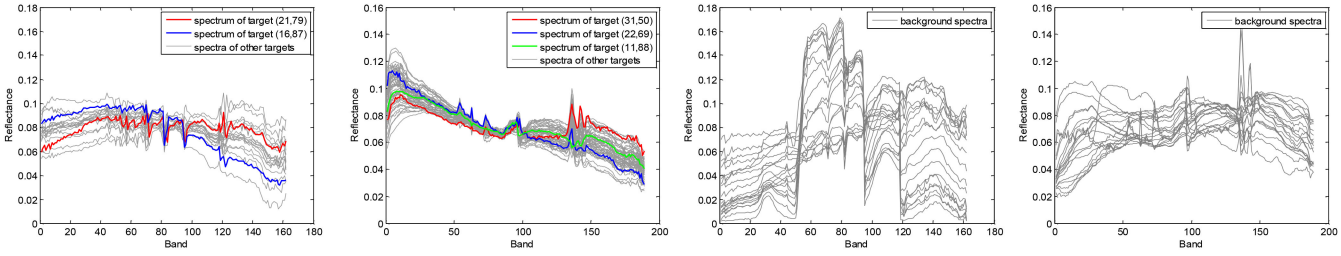


Fig. 12. Spectra of the targets of interest in the (a) HYDICE and (b) real San Diego datasets. The learned background spectra in (c) HYDICE and (d) real San Diego datasets.

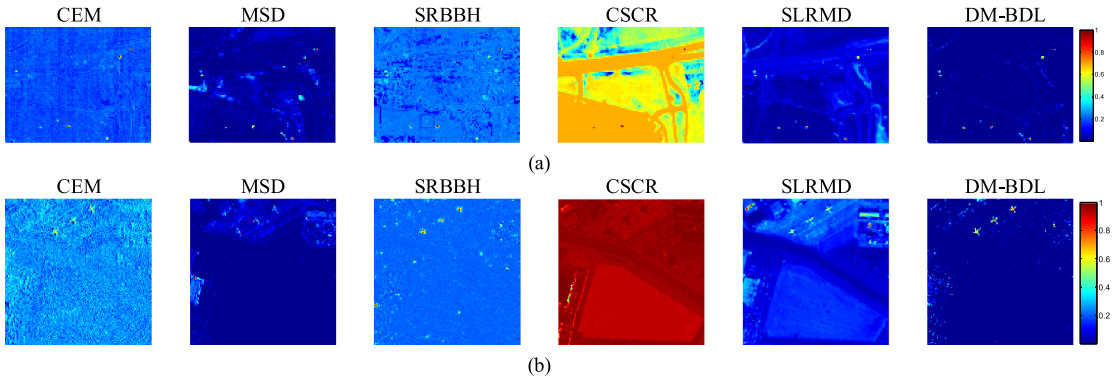


Fig. 13. 2-D plots of detection results obtained by different target detectors on (a) HYDICE and (b) real San Diego datasets.

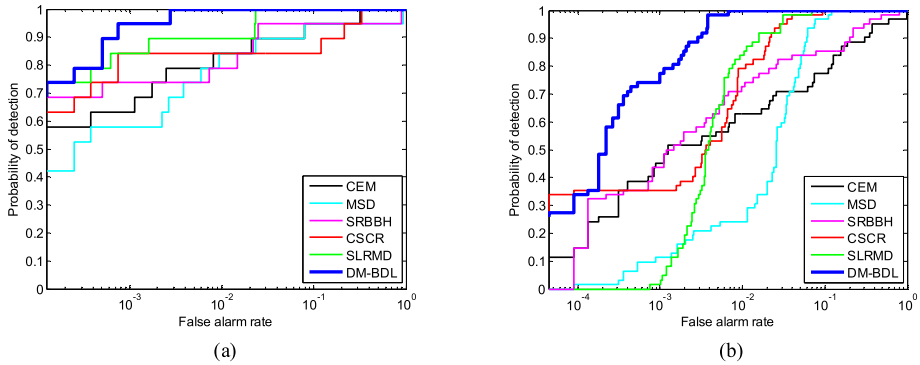


Fig. 14. ROC curves obtained by different target detectors on (a) HYDICE and (b) real San Diego datasets.

those with low fractions and maintains a very low false alarm rate.

2) *Efficiency*: The proposed BDL method based on LLC just requires to solve the ℓ_2-norm minimization problem, which can be implemented very fast. Besides, the learned background dictionary is complete and compact, thus saving the computational burden in the DM.
3) *Robustness to inaccurate prior target information:* The proposed detection algorithm is based on the separation of background and target components from the whole scene. The full exploitation of the characteristics of these components along with the learned background dictionary benefit for obtaining a satisfactory and stable decomposition result, and thus enables the detection result based on the recovered target image more robust to the prior target information with the inherent variability of spectral signatures.
4) *Convenience in selecting parameters:* The proposed method includes two parts, i.e., the BDL and DM. In the BDL process, the initial step length η is fixed at 5 and diminishes as $\eta \leftarrow \sigma\eta$ for every 50 iterations (pixels). The decay parameter σ is adaptive to the data size such that the ultimate η is about 0.001. The learning paramter γ and the number of atoms m are fixed at 20 and 20, respectively.

Moreover, the tradeoff parameters λ and β in the DM are, respectively, set to 0.01 and 0.1, and remain unchanged throughout the experiments. Satisfactory detection results have been achieved by the proposed DM-BDL under this setting for all the employed hyperspectral datasets.

IV. Conclusion

In this article, we have proposed a hyperspectral target detection algorithm based on the DM-BDL, the idea of which is built on the separation of background, target, and noise from the whole scene. In the DM, the relationships of background pixels and targets of interest, i.e., the spatial homogeneity and the spatial sparsity, are explicitly imposed in the DM, for achieving a satisfactory recovery. The detection is directly performed on the recovered target image. Since the background and noise have been well separated and removed from the scene, the proposed detector is effective in detecting the targets with low fractions and robust to inaccurate prior target information induced by the spectral variability. Furthermore, the proposed BDL algorithm based on LLC can learn a complete and compact background dictionary for better background estimation with a low computational cost. This efficiency increases the practical values of the proposed method in real scenarios. The *a priori* target dictionary is also incorporated into the learning process to mitigate the target interference, and a pure background dictionary is, therefore, learned. The extensive experimental results on both simulated and real hyperspectral datasets in comparison with several conventional and state-of-the-art target detection algorithms confirm that the proposed DM-BDL is very promising in the hyperspectral target detection.

There are two relevant topics deserving further research. First, the target dictionary is fixed in the proposed DBL method. In the future work, we will consider to learn an optimal target dictionary from the observed data in a reasonable manner given the inaccurate prior target information. The learned target dictionary may be beneficial to the subsequent recovery process. Second, the adaptive setting for the tradeoff parameters in the DM remains unsolved. Although the proposed method with a fixed setting has achieved satisfactory detection results for all the used hyperspectral datasets, the performance can be further improved with an optimal setting for each dataset. In future work, we will attempt to solve this issue with a recently proposed technique named multiobjective optimization [48].

References

[1] M. Borengasser, W. S. Hungate, and R. Watkins, *Hyperspectral Remote Sensing: Principles and Applications*. New York, NY, USA: Taylor & Francis, 2008.

[2] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.

[3] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 25, no. 7, pp. 5–28, Jul. 2010.

[4] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, and M. Brueggeman, "Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 24–33, Jan. 2014.

[5] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 58–69, Jan. 2002.

[6] M. T. Eismann, A. D. Stocker, and N. M. Nasrabadi, "Automated hyperspectral cueing for civilian search and rescue," *Proc. IEEE*, vol. 97, no. 6, pp. 1031–1055, Jun. 2009.

[7] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.

[8] N. M. Nasrabadi, "Regularized spectral matched filter for target recognition in hyperspectral imagery," *IEEE Signal Process. Lett.*, vol. 15, pp. 317–320, Mar. 2008.

[9] 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.

[10] D. Manolakis, D. Marden, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 79–116, 2003.

[11] S. Kraut, L. L. Scharf, and R. W. Butler, "The adaptive coherence estimator: A uniformly most-powerful-invariant adaptive detection statistic," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 427–438, Feb. 2005.

[12] C.-I. Chang, "Orthogonal subspace projection (OSP) revisited: A comprehensive study and analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 502–518, Mar. 2005.

[13] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.

[14] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Sparse representation for target detection in hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 629–640, Jun. 2011.

[15] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Simultaneous joint sparsity model for target detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 676–680, Jul. 2011.

[16] Y. Zhang, B. Du, Y. Zhang, and L. Zhang, "Spatially adaptive sparse representation for target detection in hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 1923–1927, Nov. 2017.

[17] Y. Zhang, B. Du, and L. Zhang, "A sparse representation-based binary hypothesis model for target detection in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1346–1354, Mar. 2015.

[18] W. Li, Q. Du, and B. Zhang, "Combined sparse and collaborative representation for hyperspectral target detection," *Pattern Recognit.*, vol. 48, no. 12, pp. 3904–3916, Dec. 2015.

[19] Y. Xu, Z. Wu, J. Li, A. Plaza, and Z. Wei, "Anomaly detection in hyperspectral images based on low-rank and sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 1990–2000, Apr. 2016.

[20] T. Cheng and B. Wang, "Graph and total variation regularized low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 391–406, Jan. 2020.

[21] A. W. Bitar, L.-F. Cheong, and J.-P. Ovarlez, "Sparse and low-rank matrix decomposition for automatic target detection in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5239–5251, Aug. 2019.

[22] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 663–670.

[23] D. Zhu, B. Du, and L. Zhang, "Single-spectrum-driven binary-class sparse representation target detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: 10.1109/TGRS.2020.2995775.

[24] N. Huyan, X. Zhang, H. Zhou, and L. Jiao, "Hyperspectral anomaly detection via background and potential anomaly dictionaries construction," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2263–2276, Apr. 2019.

[25] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[26] A. S. Charles, B. A. Olshausen, and C. J. Rozell, "Learning sparse codes for hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 963–978, Sep. 2011.

[27] Y. Niu and B. Wang, "Hyperspectral target detection using learned dictionary," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 7, pp. 1531–1535, Jul. 2015.

[28] Y. Niu and B. Wang, "Extracting target spectrum for hyperspectral target detection: An adaptive weighted learning method using a self-completed background dictionary," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1604–1617, Mar. 2017.

[29] X. Song, L. Zou, and L. Wu, "Detection of subpixel targets on hyperspectral remote sensing imagery based on background endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: 10.1109/TGRS.2020.3002461.
[30] X. Zhang, B. Hu, Z. Pan, and X. Zheng, "Dictionary learning based target detection for hyperspectral image," *Proc. SPIE*, vol. 11023, Mar. 2019, Art. no. 110232D.
[31] X. Lu, Y. Yuan, and X. Zheng, "Joint dictionary learning for multispectral change detection," *IEEE Trans. Cybern.*, vol. 47, no. 4, pp. 884–897, Apr. 2017.
[32] C. Yang, H. Liu, S. Liao, and S. Wang, "Small target detection in infrared video sequence using robust dictionary learning," *Infrared Phys. Technol.*, vol. 68, pp. 1–9, Jan. 2015.
[33] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained linear coding for image classification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, Jun. 2010, pp. 3360–3367.
[34] K. Yu, T. Zhang, and Y. Gong, "Nonlinear learning using local coordinate coding," in *Proc. 22nd Int. Conf. Neural Inf. Process. Syst.*, 2009, pp. 2223–2231.
[35] D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson, "Is there a best hyperspectral detection algorithm?," *Proc. SPIE*, vol. 7334, Apr. 2009, Art. no. 733402.
[36] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenom.*, vol. 60, no. 1/4, pp. 259–268, Nov. 1992.
[37] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 460–489, Jan. 2005.
[38] S. Yang and Z. Shi, "Hyperspectral image target detection improvement based on total variation," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2249–2258, May 2016.
[39] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.
[40] S. Yang and Z. Shi, "SparseCEM and SparseACE for hyperspectral image target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2135–2139, Dec. 2014.
[41] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
[42] Z. Lin, M. Chen, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," Univ. Illinois Urbana-Champaign, Champaign, IL, USA, Tech. Rep. UILU-ENG-09-2215, 2009.
[43] J. Kerekes, "Receiver operating characteristic curve confidence intervals and regions," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 251–255, Apr. 2008.
[44] S. Khazai, S. Homayouni, A. Safari, and B. Mojaradi, "Anomaly detection in hyperspectral images based on an adaptive support vector method," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 646–650, Jul. 2011.
[45] C.-I. Chang, "Spectral information divergence for hyperspectral image analysis," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 1, Jun. 1999, pp. 509–511.
[46] H. Kwon and N. M. Nasrabadi, "A comparative analysis of kernel subspace target detectors for hyperspectral imagery," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 1, Jan. 2007, Art. no. 029250.
[47] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[48] Z. Cai and Y. Wang, "A multiobjective optimization-based evolutionary algorithm for constrained optimization," *IEEE Trans. Evol. Comput.*, vol. 10, no. 6, pp. 658–675, Dec. 2006.

Tongkai Cheng received the B.S. degree in information display and optoelectronic technology from South China University of Technology, Guangzhou, China, in 2016. He is currently working toward the Ph.D. degree in circuits and systems with the Department of Electronic Engineering, Fudan University, Shanghai, China.

His main research interests include hyperspectral target detection, machine learning, and pattern recognition.

Bin Wang (Senior Member, IEEE) received the B.S. degree in electronic engineering and the M.S. degree in communication and electronic systems from Xidian University, Xi'an, China, in 1985 and 1988, respectively, and the Ph.D. degree in system science from Kobe University, Kobe, Japan, in 1999.

After his graduation in 1988, he was a Teacher with Xidian University. From 1999 to 2000, he was with the Communications Research Laboratory, Ministry of Posts and Telecommunications, Kobe, Japan, as a Research Fellow, working on magnetoencephalography signal processing and its application for brain science. Then, as a Senior Supervisor, he was with the Department of Etching, Tokyo Electron Ltd., Tokyo, Japan, from 2000 to 2002, dealing with the development of advanced process control systems for etching semiconductor equipment. Since September 2002, he has been with the Department of Electronic Engineering, Fudan University, Shanghai, China, where he is currently a Full Professor and Leader of the Image and Intelligence Laboratory. He has authored or coauthored more than 120 scientific papers in important domestic and international periodicals. He is the holder of several patents. His main research interests include multispectral/hyperspectral image analysis, automatic target detection and recognition, pattern recognition, signal detection and estimation, and machine learning.

Dr. Wang is an Associate Editor for the IEEE *Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (JSTARS).