

# Hierarchical Suppression Method for Hyperspectral Target Detection

Zhengxia Zou and Zhenwei Shi\*, *Member IEEE*

## Abstract

Target detection is an important application in hyperspectral image processing field and several detection algorithms have been proposed in the past decades. Some traditional detectors are built based on the statistical information of the target and background spectra, and their performances tend to be affected by the spectral quality. Some previous methods cope with this problem by refining the target spectra to make the detector robust. In this paper, instead of doing like this, we propose a new hierarchical method to suppress the backgrounds while preserving the target spectra, with the purpose of boosting the performance of traditional hyperspectral target detector. The proposed method consists of different layers of classical Constrained Energy Minimization (CEM) detectors. In each layer of detection, the CEM's output of each spectrum is transformed by a nonlinear suppression function and then considered as a coefficient to impose on this spectrum for the next round of iteration. To our knowledge, such hierarchical structure is proposed for the first time. Theoretically, we prove the convergency of the proposed algorithm, and we also give a theoretically explanation that why we can obtain the gradually increasing detection performance through the hierarchical suppression process. Experimental results on two real hyperspectral images and one synthetic image suggest our method significantly improves the performance of the original CEM detection algorithm and also outperforms other classical and recently proposed hyperspectral target detection algorithms.

## Index Terms

Hyperspectral target detection, Hierarchical structure, Nonlinear suppression function, Constrained Energy Minimization.

## I. INTRODUCTION

The hyperspectral imaging sensor collects digital images with very densely sampled radiance spectra in the scene. A hyperspectral image can be seen as a three-dimensional image cube. Each pixel in this cube serves as a spectral vector. The elements of the vector represent the radiance or reflectance values of each spectral band. As

The work was supported by the National Natural Science Foundation of China under the Grants 61273245 and the Beijing Natural Science Foundation under the Grant 4152031. (*Corresponding author: Zhenwei Shi.*)

Zhengxia Zou (e-mail: zhengxiazou@buaa.edu.cn) and Zhenwei Shi (Corresponding Author, e-mail: shizhenwei@buaa.edu.cn) are with State Key Laboratory of Virtual Reality Technology and Systems, School of Astronautics, Beihang University, Beijing 100191, China and with Beijing Key Laboratory of Digital Media, Beihang University, Beijing 100191, China and also with Image Processing Center, School of Astronautics, Beihang University, Beijing 100191, China.

each material is characterized by a unique deterministic spectrum, it is able to discriminate the materials based on the spectral characteristics. The hyperspectral image processing has received considerable interest in the past two decades. Among its wide range of applications, hyperspectral image target detection is one of the most important applications due to its both civil and military use.

Several algorithms for hyperspectral target detection have been proposed in the past decades such as Matched Filter (MF) [1], Adaptive Coherence Estimator (ACE) [1], [2], Adaptive Matched Filter (AMF) [3], Spectral Angle Mapper (SAM) [4], Adaptive Subspace Detector (ASD) [5], and the Orthogonal Subspace Projection (OSP) detector [6]. These algorithms aim at suppressing the background spectra and highlighting the target spectra at the same time. MF, ACE and AMF see the target detection as hypothesis test problems. ASD and OSP are subspace based methods. Other approaches such as Constrained Energy Minimization (CEM) [7] and Target-Constrained Interference-Minimized Filter (TCIMF) [8] impose a constraint on the targets or backgrounds and build a Finite Impulse Response (FIR) filter which minimizes the filter's output energy. All these detectors have been widely used, and several improved versions have been developed in recent years [9], [10], [11]. Recently the machine learning algorithms have been introduced to hyperspectral target detection problems. These methods include 1) sparse representation based detectors such as sparsity-based target detector [12] and simultaneous joint sparsity based detector [13]. 2) Metric learning based detectors such as supervised metric learning based detectors [14] and random forest based metric learning detector [15]. 3) Kernel based detectors such as kernel spectral matched filter [16], kernel matched subspace detectors [17] and kernel OSP detector [18]. 4) Manifold learning based detectors such as sparse transfer manifold embedding based detectors [19]. These target detection methods treat spectral data in sample space or even in the high dimensional kernel space, and benefit from the characteristics of the machine learning algorithms. Other recently proposed detection methods include the hybrid structured and unstructured detectors [20], the robust high-order statistics target detectors [21], difference measured function based matched filter [22], and etc.

Performances of the hyperspectral target detection methods usually rely on the quality of the prior target signature [2][23]. Some of methods are also required to estimate the statistical information of the pure background spectra. However, pure background and target spectral data are difficult to obtain or even unavailable. Once the quality of the spectral data decreases, the statistical characteristics of spectra will change, and the detector's performance will be affected. The quality of the spectral data is mainly influenced by the uncompensated errors in the sensor, uncompensated atmospheric and environmental effects, surface contaminants, variation in the material, etc [24]. Besides, the hyperspectral imagery usually has relatively low spatial resolution, so that one pixel sometimes is composed of several individual spectra of different materials [25]. Usually these pixels cannot represent the characteristics of the target material and background materials accurately, thus some hyperspectral unmixing algorithms have been developed [26], [27].

Some hyperspectral target detection methods have directly taken the above-mentioned problems into consideration. In [23], an iteratively reweighted method which iteratively refines the target spectrum to generate the "optimal" spectrum is proposed. An ACE detector generates the final detection results by using this "optimal" target spectrum. In [28], the authors construct the target detector based on the combined spectral signatures, which can tolerate the

spectral variations of different pixels in the same object. In this paper, instead of refining the target spectrum directly like [23] and [28], we build a new hierarchical architecture to suppress the variational background spectra while preserving the targets. A simple and effective algorithm, Hierarchical CEM (hCEM) algorithm, is proposed with the purpose of improving the performance of traditional CEM detector. Since the classical CEM detector in some special cases cannot completely push out the targets and suppress the backgrounds in one round of filtering process, we filter the data for several times to solve this problem. In our method, the CEM detectors of different layers are linked in series. After each layer's detection, some background spectra are suppressed by a nonlinear function based on the output of the detector. Then the transformed spectra are forward sent to the next layer's detector, until the CEM detector's output converges to a constant. Suppressing the undesired backgrounds makes the CEM detector better concentrate on the hard-detected targets. In this way, the performance of the detector will be gradually enhanced layer-by-layer. The contributions of our work are summarized as bellow:

1) A new hierarchical structure for hyperspectral target detection is proposed. To our knowledge, it is the first time to apply such hierarchical structure to hyperspectral target detection problems.

2) We theoretically prove the convergence of the proposed algorithm. After several layers' detection, the detection outputs will converge to a constant (final output). We also theoretically prove that in each layer we learn a better detector than the previous layers, and the detection performance will be gradually enhanced.

3) Experimental results on one synthetic hyperspectral image and two real hyperspectral images suggest the proposed method significantly enhances the traditional CEM and outperforms than other classical detection algorithms.

The rest of this paper is organized as follows. In Section II, we will introduce the original CEM detector and our hCEM. In section III, we give some theoretical analyses. Some experimental results are given in section IV and conclusions are drawn in Section V.

## II. HIERARCHICAL SUPPRESSION METHOD

In this section, we will first introduce the traditional CEM algorithm and then we will introduce our hCEM algorithm.

### A. A Brief Introduction to CEM

Consider a hyperspectral image with  $N$  spectral vectors and  $L$  bands:  $\mathbf{x}_i \in \mathbb{R}^{L \times 1}$ ,  $i = 1, 2, \dots, N$ . All spectra of the hyperspectral image can be arranged in an  $L \times N$  matrix as  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ . The aim of CEM algorithm is to design an optimal FIR filter specified by the vector  $\mathbf{w} = [w_1, w_2, \dots, w_L]^T$ . The average output energy from all the pixel vectors can be represented as

$$\begin{aligned} \frac{1}{N} \|\mathbf{y}\|_2^2 &= \frac{1}{N} \|\mathbf{w}^T \mathbf{X}\|_2^2 \\ &= \frac{1}{N} \mathbf{w}^T \mathbf{X} \mathbf{X}^T \mathbf{w} \\ &= \mathbf{w}^T \mathbf{R} \mathbf{w} \end{aligned} \tag{1}$$

where  $\mathbf{R} = \frac{1}{N}\mathbf{X}\mathbf{X}^T$  represents the correlation matrix,  $\mathbf{y} = [y_1, y_2, \dots, y_N] \in \mathbb{R}^{1 \times N}$  represents the output of the detector. The CEM designs an FIR filter which minimizes the total output energy subject to a constraint that the filter's response to  $\mathbf{d}$  is a constant (e.g.  $\mathbf{w}^T \mathbf{d} = 1$ ):

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^T \mathbf{R} \mathbf{w} \\ \text{s.t.} \quad & \mathbf{w}^T \mathbf{d} = 1 \end{aligned} \quad (2)$$

where  $\mathbf{d}$  is a pre-chosen target spectrum and can be obtained by averaging different target spectral vectors of a certain material in one hyperspectral image. The solution of the above optimization problem is given by [7] which is

$$\mathbf{w}^* = \frac{\mathbf{R}^{-1} \mathbf{d}}{\mathbf{d}^T \mathbf{R}^{-1} \mathbf{d}}. \quad (3)$$

Thus the output of the CEM filter is given by

$$\begin{aligned} \mathbf{y} &= (\mathbf{w}^*)^T \mathbf{X} \\ &= \frac{\mathbf{d}^T \mathbf{R}^{-1}}{\mathbf{d} \mathbf{R}^{-1} \mathbf{d}^T} \mathbf{X}. \end{aligned} \quad (4)$$

Usually, the target pixels will get large value of outputs while the background pixels will get small ones. Finally, each element of  $\mathbf{y}$  is compared with a fix threshold. If the output value is higher than the threshold, we decide the target present in the corresponding pixel, else we decide there is a target absent.

### B. Hierarchical CEM

In this paper, we believe that a transformation on the spectra is beneficial for target detection problem. The major improvements of the proposed method can be summarized as the following 3 points.

- 1) The traditional CEM detector is a single layer detector, while the proposed hCEM detector consists of different layers of traditional CEM detectors, and the detectors of different layers are linked in series.
- 2) After each layer of detection, the background spectra are suppressed (reduce its magnitude while keeping its direction in the spectral space) based on the current layers output score.
- 3) The CEM detector is constructed based on the correlation matrix  $\mathbf{R}$  while the hCEM detector is constructed based on the corresponded revised correlation matrix. Since the revised correlation matrix contains more information of the hard detected spectra, the hCEM could have better concentration on those hard detected pixels.

Now consider the  $k$ th layer. The CEM output of this layer can be represented as

$$\mathbf{y}^k = \frac{\mathbf{R}_k^{-1} \mathbf{d}}{\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{d}} \mathbf{X}^k, \quad (5)$$

where  $\mathbf{X}^k$  and  $\mathbf{R}^k$  represent the spectral matrix and the correlation matrix of the  $k$ th layer, respectively. Then each spectral vector  $\mathbf{x}_i^k$  is transformed by multiplying a non-negative number  $q(y_i^k)$  based on its output score:

$$\mathbf{x}_i^{k+1} = q(y_i^k) \mathbf{x}_i^k \quad (6)$$

where a nonlinear function  $q(t) : \mathbb{R}^1 \rightarrow \mathbb{R}^1$  is used to impose on the spectral vector  $\mathbf{x}_i^k$ . We consider this function as a ‘‘soft-threshold’’ operation: hold the spectrum  $\mathbf{x}_i$  whose output score is large, while suppress the spectrum  $\mathbf{x}_j$

whose output score is small. In this way, the undesired background spectra are gradually suppressed after each layer's detection while the target spectra will keep unchanged. In this paper, the nonlinear suppression function is defined as follows

$$q(t) = \begin{cases} q(t) = 1 - e^{-\lambda t} & t \geq 0 \\ 0 & t < 0, \end{cases} \quad (7)$$

where  $\lambda$  is a positive parameter to adjust the shape of the function (7). Fig. 1 shows shape of function (7) under different choices of  $\lambda$ . Finally, the target spectra and the transformed background spectra will be used to construct the new CEM detector in the  $(k + 1)$ th layer. The above steps will be repeated until the output  $\mathbf{y}^k$  converges to a constant. In this paper, we calculate  $\delta_k$ , the error of the average output energy of the current layer and the previous layer:

$$\delta_k = \frac{1}{N} \|\mathbf{y}^k\|_2^2 - \frac{1}{N} \|\mathbf{y}^{k-1}\|_2^2. \quad (8)$$

If  $\delta_k < \varepsilon$  ( $\varepsilon$  refers to a small positive number), the iteration will be stopped. The outline of the proposed hCEM algorithm for hyperspectral target detection is given as follows:

---

**Algorithm 1** Hierarchical CEM Algorithm

---

**Input:**

1. spectral matrix  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \in \mathbb{R}^{L \times N}$ , target spectrum  $\mathbf{d} \in \mathbb{R}^{L \times 1}$ , set tolerance  $\varepsilon > 0$ ,

**Initialization:**

2.  $k = 1$ ,  $\mathbf{X}^1 = \mathbf{X}$ ,

**Hierarchical Suppression:**

3.  $\mathbf{R}_k = \frac{1}{N} \mathbf{X}^k \mathbf{X}^{kT}$ ,
4.  $\mathbf{y}^k = (\mathbf{R}_k^{-1} \mathbf{d}) / (\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{d}) \mathbf{X}^k$ ,
5.  $\mathbf{x}_i^{k+1} = (1 - e^{-\lambda y_i^k}) \mathbf{x}_i^k$ ,
6.  $k \leftarrow k + 1$ ,

**Stop criterion:**

- $\delta_k = \frac{1}{N} \|\mathbf{y}^k\|_2^2 - \frac{1}{N} \|\mathbf{y}^{k-1}\|_2^2$ , if  $\delta_k > \varepsilon$ , go back to step 3; else, go to step 7,

**Output:**

7. final outputs:  $\mathbf{y}^k = [y_1^k, y_2^k, \dots, y_N^k] \in \mathbb{R}^{1 \times N}$ .
- 

### III. THEORETICAL ANALYSIS

In this section, we will give some theoretical analyses of our hCEM algorithm. We will first theoretically analyze the convergence of the proposed hCEM algorithm. Then we will explain why the background suppression process of each layer makes it a better detection. Finally, we will give some guidance on how to set the parameter  $\lambda$ .

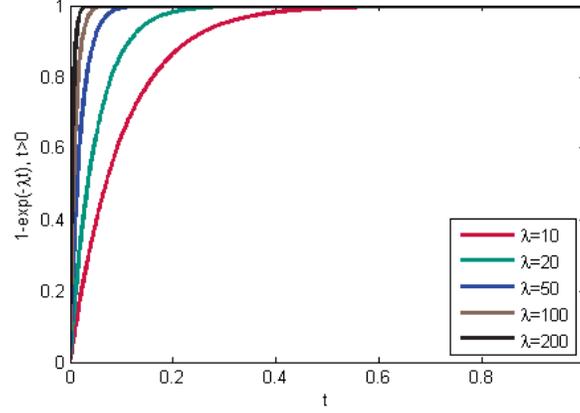


Fig. 1. Shape of the nonlinear suppression function  $q(t) = 1 - e^{-\lambda t}$ ,  $t \geq 0$ , with different choices of  $\lambda$ .

### A. Convergence Analysis

In this subsection, we will prove that after several layers' filtering operations, the output energy of the hCEM will converge to a constant. This proof also provides the theoretical basis of the “stop criterion” used in our algorithm. First, we give a lemma.

**Lemma 1.** The inverse of correlation matrices of the  $k$ th layer and the  $(k-1)$ th layer have the following relationship

$$\mathbf{R}_k^{-1} - \mathbf{R}_{k-1}^{-1} = \sum_{i=1}^N \frac{\mathbf{R}_k^{-1} \mathbf{u}_i \mathbf{u}_i^T \mathbf{R}_k^{-1}}{1 + \mathbf{u}_i^T \mathbf{R}_k^{-1} \mathbf{u}_i} \quad (9)$$

where  $\mathbf{R}_k = \frac{1}{N} \mathbf{X}_k \mathbf{X}_k^T$  represents the correlation matrix of  $k$ th layer,  $\mathbf{u}_i = \sqrt{(1 - q_i^2)/N} \mathbf{x}_i^k \in \mathbb{R}^{D \times 1}$ ,  $q_i$  is the suppression coefficient which is imposed on the corresponding spectrum.

**Proof of Lemma 1.** The correlation matrices of the  $k$ th layer and the  $(k-1)$ th layer have the following relationship

$$\begin{aligned} \mathbf{R}_k - \mathbf{R}_{k-1} &= \sum_{i=1}^N (q_i^2 \mathbf{x}_i^{k-1} \mathbf{x}_i^{k-1T} - \mathbf{x}_i^{k-1} \mathbf{x}_i^{k-1T}) / N \\ &= \mathbf{U} \mathbf{U}^T, \end{aligned} \quad (10)$$

where  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N] \in \mathbb{R}^{D \times N}$ . Suppose  $\mathbf{R}_k$  and  $(\mathbf{R}_k + \mathbf{u}_i \mathbf{u}_i^T)$  are all invertible, then based on the Sherman-Morrison-Woodbury Formula [29], we have

$$(\mathbf{R}_k + \mathbf{u}_i \mathbf{u}_i^T)^{-1} = \mathbf{R}_k^{-1} - \frac{\mathbf{R}_k^{-1} \mathbf{u}_i \mathbf{u}_i^T \mathbf{R}_k^{-1}}{1 + \mathbf{u}_i^T \mathbf{R}_k^{-1} \mathbf{u}_i}. \quad (11)$$

Thus we have

$$\begin{aligned} \mathbf{R}_k^{-1} - \mathbf{R}_{k-1}^{-1} &= \mathbf{R}_k^{-1} - (\mathbf{R}_k + \mathbf{U} \mathbf{U}^T)^{-1} \\ &= \sum_{i=1}^N \frac{\mathbf{R}_k^{-1} \mathbf{u}_i \mathbf{u}_i^T \mathbf{R}_k^{-1}}{1 + \mathbf{u}_i^T \mathbf{R}_k^{-1} \mathbf{u}_i}. \quad \square \end{aligned} \quad (12)$$

**Theorem 1.** The average output energy  $\frac{1}{N} \|\mathbf{y}^k\|_2^2$  of hCEM algorithm will converge to a constant.

*Proof of Theorem 1.* The CEM detector in  $k$ th and  $(k-1)$ th layer can be viewed as two FIR filters specified by the vectors  $\mathbf{w}_{k-1}$  and  $\mathbf{w}_k$ , which can be obtained by solving the optimization problem (2)

$$\mathbf{w}_{k-1} = \frac{\mathbf{R}_{k-1}^{-1} \mathbf{d}}{\mathbf{d}^T \mathbf{R}_{k-1}^{-1} \mathbf{d}}, \quad \mathbf{w}_k = \frac{\mathbf{R}_k^{-1} \mathbf{d}}{\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{d}}. \quad (13)$$

$$\|\mathbf{y}^k\|_2^2 = \|\mathbf{w}^{kT} \mathbf{X}^k\|_2^2 = \frac{N}{\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{d}}, \quad (14)$$

$$\|\mathbf{y}^{k-1}\|_2^2 = \|\mathbf{w}^{k-1T} \mathbf{X}^{k-1}\|_2^2 = \frac{N}{\mathbf{d}^T \mathbf{R}_{k-1}^{-1} \mathbf{d}}.$$

From Lemma 1. we have

$$\begin{aligned} \mathbf{d}^T (\mathbf{R}_k^{-1} - \mathbf{R}_{k-1}^{-1}) \mathbf{d} &= \sum_{i=1}^N \frac{\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{u}_i \mathbf{u}_i^T \mathbf{R}_k^{-1} \mathbf{d}}{1 + \mathbf{u}_i^T \mathbf{R}_k^{-1} \mathbf{u}_i} \\ &= \sum_{i=1}^N \frac{\|\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{u}_i\|_2^2}{1 + \mathbf{u}_i^T \mathbf{R}_k^{-1} \mathbf{u}_i}. \end{aligned} \quad (15)$$

Since  $\|\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{u}_i\|_2^2 \geq 0$  and  $\mathbf{u}_i^T \mathbf{R}_k^{-1} \mathbf{u}_i \geq 0$ , thus we have the relationship  $\mathbf{d}^T (\mathbf{R}_k^{-1} - \mathbf{R}_{k-1}^{-1}) \mathbf{d} \geq 0$ . Thus,

$$\frac{1}{\mathbf{d}^T \mathbf{R}_{k-1}^{-1} \mathbf{d}} \geq \frac{1}{\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{d}}. \quad (16)$$

Therefore,

$$\frac{1}{N} \|\mathbf{y}^k\|_2^2 \leq \frac{1}{N} \|\mathbf{y}^{k-1}\|_2^2. \quad (17)$$

The average output energy will converge to a constant.  $\square$

### B. Performance Analysis

In this subsection, we will show why in each layer we can learn a better CEM detector than previous layers, and how is the detection performance gradually enhanced.

**Theorem 2.** The residual error of hierarchical CEM detector  $R(k)$  in  $k$ th layer is not larger than the residual error  $R(k-1)$  in  $(k-1)$ th layer:

$$R(k) \leq R(k-1), \quad (18)$$

where  $R(k) = \frac{1}{N} \|\mathbf{y}^k - \mathbf{z}\|_2^2$ ,  $\mathbf{z} = [z_1, z_2, \dots, z_N] \in \mathbb{R}^{1 \times N}$  represents the label of the input pixel  $\mathbf{x}_i$ : if  $\mathbf{x}_i$  belongs to target class,  $z_i = 1$ , else  $z_i = 0$ .

*Proof of Theorem 2.*

$$\begin{aligned} R(k) &= \frac{1}{N} \|\mathbf{y}^k - \mathbf{z}\|_2^2 \\ &= \frac{1}{N} (\|\mathbf{y}^k\|_2^2 - 2\mathbf{y}^k \mathbf{z}^T + \|\mathbf{z}\|_2^2) \\ &= \frac{1}{N} \|\mathbf{y}^k\|_2^2 - \frac{2}{N} \mathbf{w}_k^T \mathbf{X}_k \mathbf{z}^T + \frac{1}{N} \|\mathbf{z}\|_2^2 \end{aligned} \quad (19)$$

$R(k)$  is composed of three terms. The first term can be obtained from Theorem 1:

$$\frac{1}{N} \|\mathbf{y}^k\|_2^2 = \frac{1}{\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{d}}. \quad (20)$$

In the second term, suppose all the target spectra will not be suppressed during the hierarchically filtering process, thus we have

$$\mathbf{X}_k \mathbf{z}^T = N_t \mathbf{d}, \quad (21)$$

where  $\mathbf{d}$  is the mean vector of the target spectra and  $N_t$  is the number of target spectra. Thus the second term becomes

$$\frac{2}{N} \mathbf{w}_k^T \mathbf{X}_k \mathbf{z}^T = \frac{2N_t}{N} \mathbf{w}_k^T \mathbf{d} = \frac{2N_t}{N}. \quad (22)$$

The third term

$$\frac{1}{N} \|\mathbf{z}\|_2^2 = \frac{N_t}{N} \quad (23)$$

is a constant. Substitute (20) (22) and (23) into (19), we have

$$\begin{aligned} R(k) &= \frac{1}{\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{d}} - \frac{2N_t}{N} + \frac{N_t}{N} \\ &= \frac{1}{\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{d}} - \frac{N_t}{N} \end{aligned} \quad (24)$$

Since  $\mathbf{d}^T \mathbf{R}_k^{-1} \mathbf{d} \geq \mathbf{d}^T \mathbf{R}_{k-1}^{-1} \mathbf{d}$ , the residual error have the relationship  $R(k) \leq R(k-1)$ .  $\square$

Actually  $R(k)$  can be seen as a measurement of the similarity between the output  $\mathbf{y}^k$  and the label  $\mathbf{z}$ . Clearly, the smaller  $R(k)$  is, a better detector we obtained. When  $R(k) = 0$ , we have  $\mathbf{y}^k = \mathbf{z}$ , which represents the ideal detection results. The above theorem indicates that the performance of the CEM detector in the  $k$ th layer is at least not worse than that in the  $k-1$ th layer. In fact, by using the nonlinear suppression operation behind each layer, the magnitude of some background spectra will be suppressed to zero. In this way, the detectors in a deeper layer of the structure will become less concerned about these undesired backgrounds, and can better concentrate on the hard-detected targets.

### C. Parameter Analysis

The performance of the proposed hCEM algorithm depends on the choice of the parameter  $\lambda$  of the nonlinear suppression function (7). Before the hCEM detection process, we have to guarantee the nonlinear transformation will not ‘‘hurt the real target spectra, since any miss-suppressed targets cannot be recovered in the subsequence layers. Actually, every algorithm has its limitation, and no one can guarantee their methods work well under any extreme conditions, so does our hCEM. In this subsection, we theoretically analyze the influence of the parameter  $\lambda$  in the proposed hCEM algorithm in a probabilistic way. We also introduce a criterion on how to select the parameter  $\lambda$  in this subsection. By using this criterion, we can make sure that the nonlinear transformation will not ‘‘hurt’’ any target spectra. In the end, the failure conditions of hCEM are briefly analyzed.

Suppose that the background spectra and target spectra can be modeled by two multivariate normal distributions with two mean vectors  $\boldsymbol{\mu}$ ,  $\mathbf{d}$  and the same covariance matrix  $\mathbf{C}$  [2],

$$\begin{aligned}\mathbf{x}_B &\sim N(\boldsymbol{\mu}, \mathbf{C}) \\ \mathbf{x}_T &\sim N(\mathbf{d}, \mathbf{C}).\end{aligned}\tag{25}$$

Since any linear combinations of the components of multivariate normal random variable still obey normal distribution, the probability distribution of the background output  $y_B$  and the target output  $y_T$  can be determined as

$$\begin{aligned}y_B &= \mathbf{w}^T \mathbf{x}_B \sim N(\mathbf{w}^T \boldsymbol{\mu}, \mathbf{w}^T \mathbf{C} \mathbf{w}) \\ y_T &= \mathbf{w}^T \mathbf{x}_T \sim N(\mathbf{w}^T \mathbf{d}, \mathbf{w}^T \mathbf{C} \mathbf{w}),\end{aligned}\tag{26}$$

where  $\mathbf{w} = \frac{\mathbf{R}^{-1} \mathbf{d}}{\mathbf{d}^T \mathbf{R}^{-1} \mathbf{d}}$  and  $\mathbf{w}^T \mathbf{d} = 1$ .

Given a small positive number  $0 < \xi < 1$ , we define

$$\begin{aligned}P_\xi(\lambda) &= P\{q(y_T) \leq 1 - \xi\} \\ &= P\{e^{-\lambda y_T} \geq \xi\}\end{aligned}\tag{27}$$

as the probability of a target spectrum  $\mathbf{x}_i$  miss-suppressed by the nonlinear function (1).  $P_\xi(\lambda)$  can be determined based on the normal distribution (26) that

$$P_\xi(\lambda) = F\left(\frac{1}{\lambda} \log\left(\frac{1}{\xi}\right); \mathbf{w}^T \mathbf{d}, \mathbf{w}^T \mathbf{C} \mathbf{w}\right),\tag{28}$$

where  $F(t; \mu, \sigma^2)$  refers to the cumulative distribution function of a gaussian distribution variable  $t$  with expectation  $\mu = \mathbf{w}^T \mathbf{d} = 1$  and variance  $\sigma^2 = \mathbf{w}^T \mathbf{C} \mathbf{w}$ . It can be seen that, under certain value of  $\xi$ , the suppression probability  $P_\xi(\lambda)$  is monotonous decreasing as  $\lambda$  increases. The larger value of  $\lambda$  we choose, the smaller probability a target spectrum will be miss-suppressed. In this paper, we choose relatively large value of  $\lambda$  to control the probability  $P_\xi(\lambda)$  small. The choice of  $\lambda$  should be satisfied with the following criterion:

$$P_\xi(\lambda) < \eta,\tag{29}$$

where  $\eta > 0$  gives the safety range of  $P_\xi(\lambda)$ , say,  $\eta = 10^{-9}$ . In this way, we can make sure that the nonlinear transformation will not ‘‘hurt’’ any target pixels (even if this does happen, the probability will become quite small). For the worst case, if the CEM outputs of target pixels are much smaller than the most background pixels, say, the detection outputs of targets and backgrounds are totally inversed. Under this case, hCEM will fail to detect the target. Nevertheless, in most case, if we choose a reasonable parameter, our hCEM can always work well with a large probability.

#### IV. EXPERIMENTS RESULTS AND ANALYSIS

In this section, we use one synthetic hyperspectral image and two real hyperspectral images to demonstrate the efficiency of our hCEM algorithm. In all our experiments, we use the same parameters  $\lambda = 200$  and  $\varepsilon = 10^{-6}$  for our hCEM algorithm. We compare our hCEM algorithm with several single layer detection algorithms: CEM, ACE,

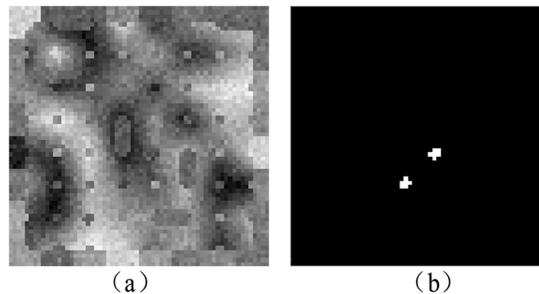


Fig. 2. (a) First band of the synthetic hyperspectral image. (b) Truth distribution of the targets (groundtruth).

MF, AMF, SAM, difference measured function based matched filter (DFMF) [22] and robust high-order matched filter (RHMF) [21]. The former five algorithms are classical detection algorithm, while the last two algorithms are recent improved version of linear matched detection algorithm. In DFMF and RHMF, the authors use the nonlinear measured functions to design the objective function, and apply the gradient descent method to find the optimal projection vector  $\mathbf{w}$ . We also compare with another multi-layer detection algorithm recently published in [23], which is named as “reweighted ACE (rACE)” detector. In their method, an ACE detector is used as a basic detector in each layer, and the target spectrum is revised iteratively based on the last layer’s detection outputs to generate the “optimal” target spectrum. Finally the authors use the “optimal” target spectrum to obtain the final detection result. In DFMF, RHMF and rACE, the parameters we used are provided by their authors. The matlab code of our hCEM algorithm can be free download at [http://levir.buaa.edu.cn/code/hCEM\\_Demo.rar](http://levir.buaa.edu.cn/code/hCEM_Demo.rar).

#### A. Experiment on Synthetic Data

The spectral data we used in this experiment is provided by the United States Geological Survey (USGS) [30] digital spectral library, where 15 endmember signatures are used to generate our synthetic data, including Labradorite HS17.3B, Rhodochrosite HS67, and etc. The above 15 spectra are collected in 224 bands uniformly spanning from 0.4 to 2.5  $\mu\text{m}$ . In this experiment, Labradorite HS17.3B is used as the target spectrum. We use the target implantation method introduced by Chang et al. [31] to generate the synthetic data. We first divide the synthetic map, whose size is  $s^2 \times s^2$  ( $s = 8$ ), into  $s \times s$  regions. Each region is initialized with the same type of ground cover which is randomly selected from the above 15 kinds of spectra. We implant clean target into the backgrounds by replacing their corresponding pixels. To evaluate the detector’s performance on mixed spectral data, we first mix the synthetic map through a  $(s + 1) \times (s + 1)$  spatial low-pass filter to generate the mixed pixels, and then both targets and backgrounds are corrupted by a Gaussian white noise with 30dB SNR at the same time. Fig. 2 (a) shows the first band of the synthetic image and Fig. 2 (b) shows the groundtruth.

The deviation  $\delta_k$  of our hCEM algorithm converges to 0 (less than  $10^{-6}$ ) after 8 layers of iterations, and we obtain the final detection results. Fig. 3 shows the average output energy of hCEM in different layers. As we expected, the output energy decreases with the layer number increases. It drops quickly at the 1st, 2nd and 3rd layer, and it

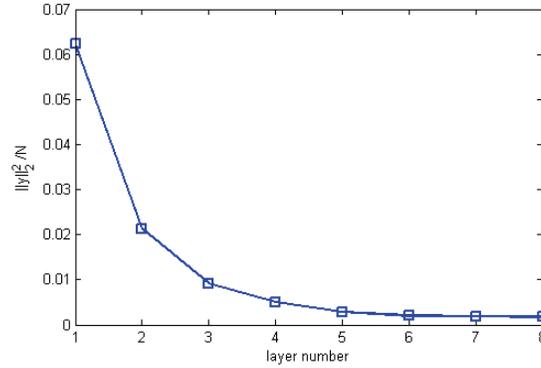


Fig. 3. The average output energy of hCEM algorithm with different layers on synthetic data.

tends to converge to a constant after the 5th layer. The first row of Fig. 4 shows the detection results of our hCEM of the 1st-5th layers. The second row of Fig. 4 shows the detection results of rACE of the 1st-5th layers. The third and the fourth row of Fig. 4 shows the detection results of ACE, AMF, CEM, MF, SAM, DFMF and RHMF. The first layer result of rACE and hCEM also represents the original single layer ACE and CEM algorithm. All results in Fig. 4 are normalized to [0,1] for comparison. We can see the gradually increasing performance of our hCEM algorithm with the increasing number of layers. The rACE algorithm reaches its best detection performance at the 2nd layer, but after the 2nd layer, the performance is somewhat instable. It should be noticed that although rACE converges and stops at the 3rd layer, for better illustration, we also list its subsequent 4-5th layers detection results.

To further illustrate the effectiveness of our hCEM algorithm, the Receiver Operating Characteristics (ROC) [32] curves are used. The ROC curves describe the varying relationship of detection probability and false alarm rate, and provide performance comparison of the different detectors [2], [32]. Based on the groundtruth image, the ROC curve gives the relationship between the false alarm rate (Fa) and the probability of detection (Pd) by changing different thresholds on a detector's output. Fa and Pd are defined as follows:

$$Fa = \frac{N_f}{N_b}, \quad Pd = \frac{N_c}{N_t}, \quad (30)$$

where  $N_f$  is the number of false alarm pixels,  $N_b$  is the total number of background pixels,  $N_c$  is the number of correct detection target pixels and  $N_t$  is the number of total true target pixels. Clearly, if an algorithm gets a higher detection probability on the same false alarm rate than other algorithms, it means this algorithm performs better than others. Fig. 5 shows the ROC curves of the detection results of all above mentioned methods. The ROC curves further prove that our hCEM algorithm largely improves the detection performance of the original CEM algorithm, and outperforms than some of the other classical detection algorithms (ACE, CEM, SAM and RHMF) and the multi-layer detection algorithm rACE. In this experiment, AMF, MF, DFMF and our hCEM get the best detection results, and their ROC curves become four straight lines. To better discriminate their performances, both targets and backgrounds are added with stronger Gaussian white noise with 20dB SNR. Fig. 6 plots the ROC curves of all above mentioned methods. This time our hCEM algorithm outperforms than all the other methods.

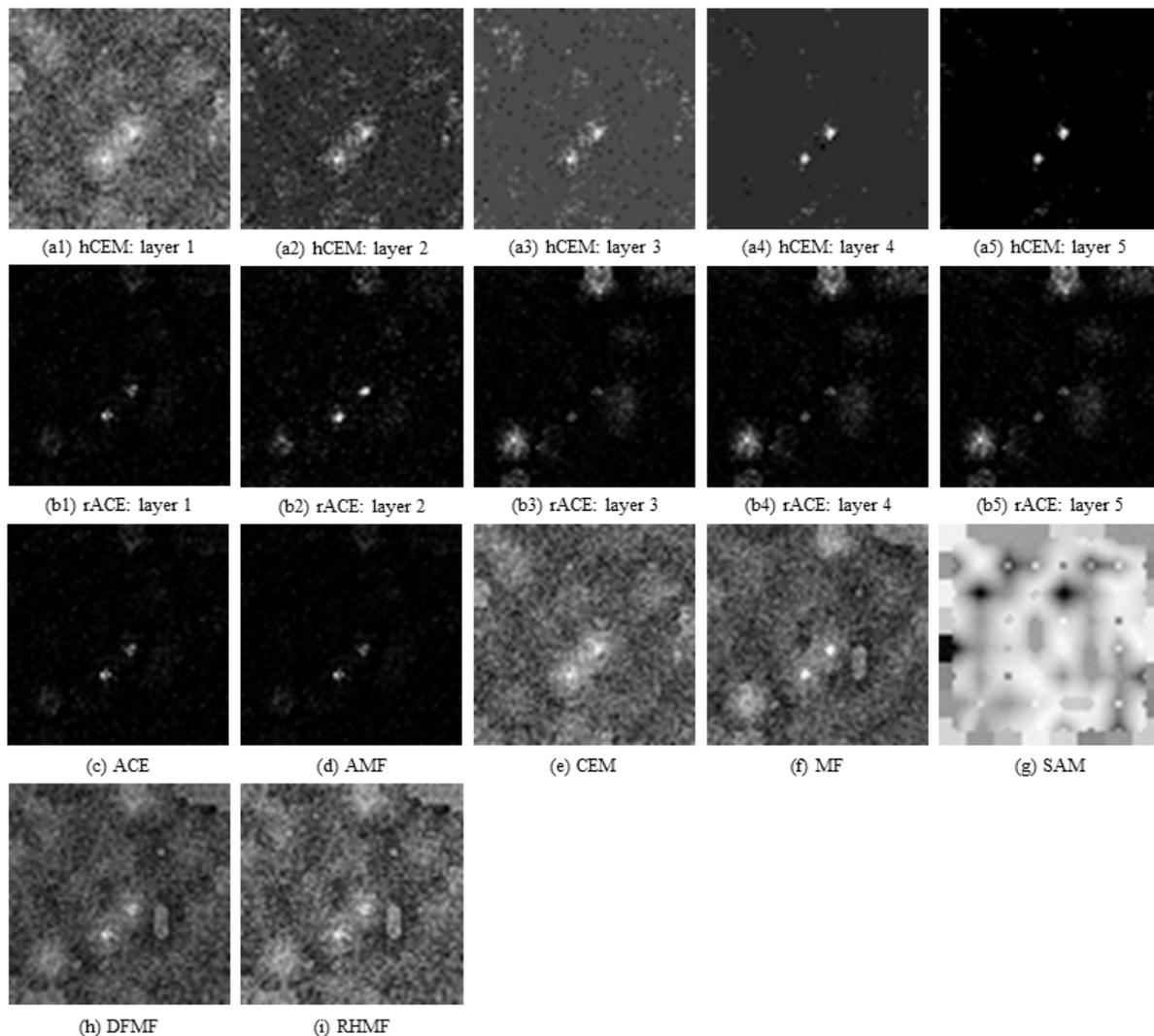


Fig. 4. Detection results of the synthetic hyperspectral data (with Gaussian white noise 30dB SNR). First row (a1)-(a5): results of our hCEM of 1st-5th layers. Second row (b1)-(b5): results of rACE of 1st-5th layers. Third row and fourth row (c)-(i): results of ACE, AMF, CEM, MF, SAM, DFMF and RHMF. The first layer results of rACE and hCEM also represent the results of the original ACE and CEM algorithm. It should be noticed that although rACE converges and stops at the 3rd layer, for better illustration, we also list its subsequent 4-5th layer's detection results. All the detection results are normalized to [0,1] for comparison.

It should be noticed that performances of the recently proposed multi-layer detection algorithm rACE are not so good, even worse than its basic single layer algorithm ACE. This is mainly because that, in each layer, rACE algorithm utilizes the ACE's detection score as the weights to revise the target signature. However, in our experiment, the target spectra only occupy a very small population of the whole spectral data. In this case, reweighted target signature in each layer will gradually diverge from the "true" target signature due to the interferences of undesired background spectra. If the target spectra have a larger proportion in our experiment data, rACE may have a better

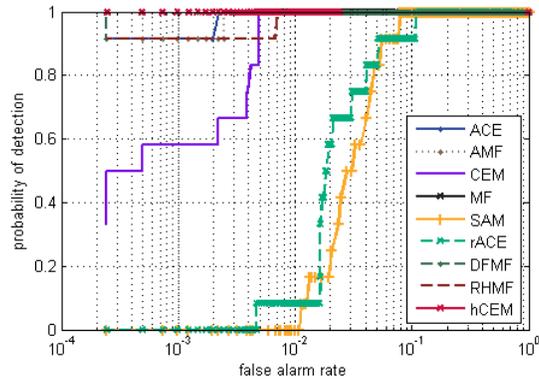


Fig. 5. The ROC curves of ACE, AMF, CEM, MF, SAM, DFMF, RHMF, rACE and our hCEM on the synthetic hyperspectral image with 30dB SNR Gaussian white noise. AMF, MF DFMF, and our hCEM get the best detection results.

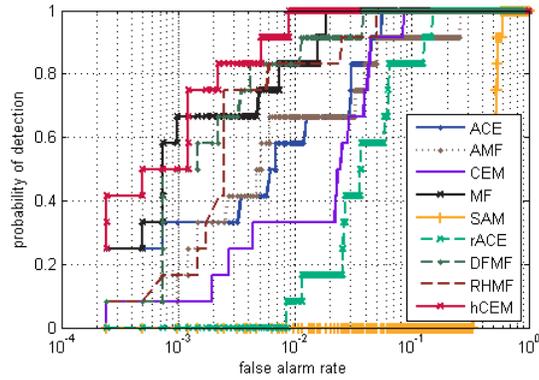


Fig. 6. The ROC curves of ACE, AMF, CEM, MF, SAM, DFMF, RHMF, rACE and our hCEM on the synthetic hyperspectral image with 20dB SNR Gaussian white noise. Our hCEM get the best detection result.

detection result.

### B. Experiment on AVIRIS Data I

In this subsection, we test our hCEM algorithm on the well-known Cuprite data [33] set which is collected by the Airborne Visible/InfraRed Imaging Spectrometer (AVIRIS). The data were captured in the Cuprite mining district of Nevada in 1997. The AVIRIS sensor collects the spectral data in 224 bands spanning from 0.4 to 2.5  $\mu\text{m}$ . There are about 14 kinds of mineral in this hyperspectral scene. The data we used in this experiment is part of the AVIRIS image, which has  $250 \times 191$  pixels and 188 bands (low SNR and water absorption bands were removed). We use the data to detect the buddingtonite target, which occupies about 35-40 pixels. Fig. 8 (a) shows the first band of the Cuprite data. Fig. 8 (b) shows the distribution of the buddingtonite targets produced by the Tricorder software. Fig. 7 shows the minerals map [34] which is produced by the Tricorder 3.3 software product.

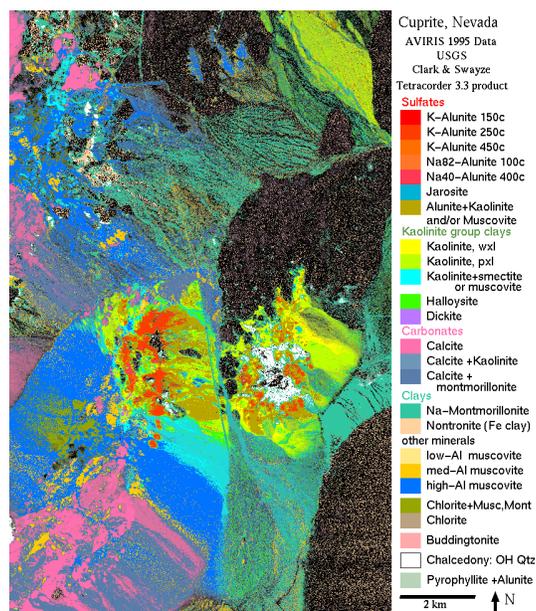


Fig. 7. Distribution of different minerals in the Cuprite mining district in Nevada.

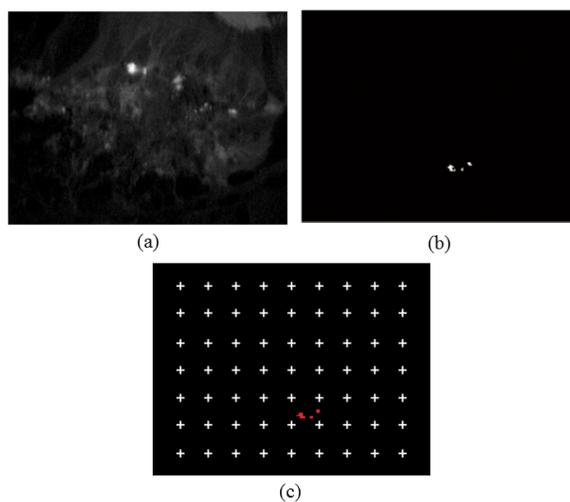


Fig. 8. (a) The 1st band of AVIRIS Cuprite data. (b) Distribution of the buddingtonite targets produced by the Tricorder software. (c) Observation point distribution: 63 background points (white "+") and 40 target points (red).

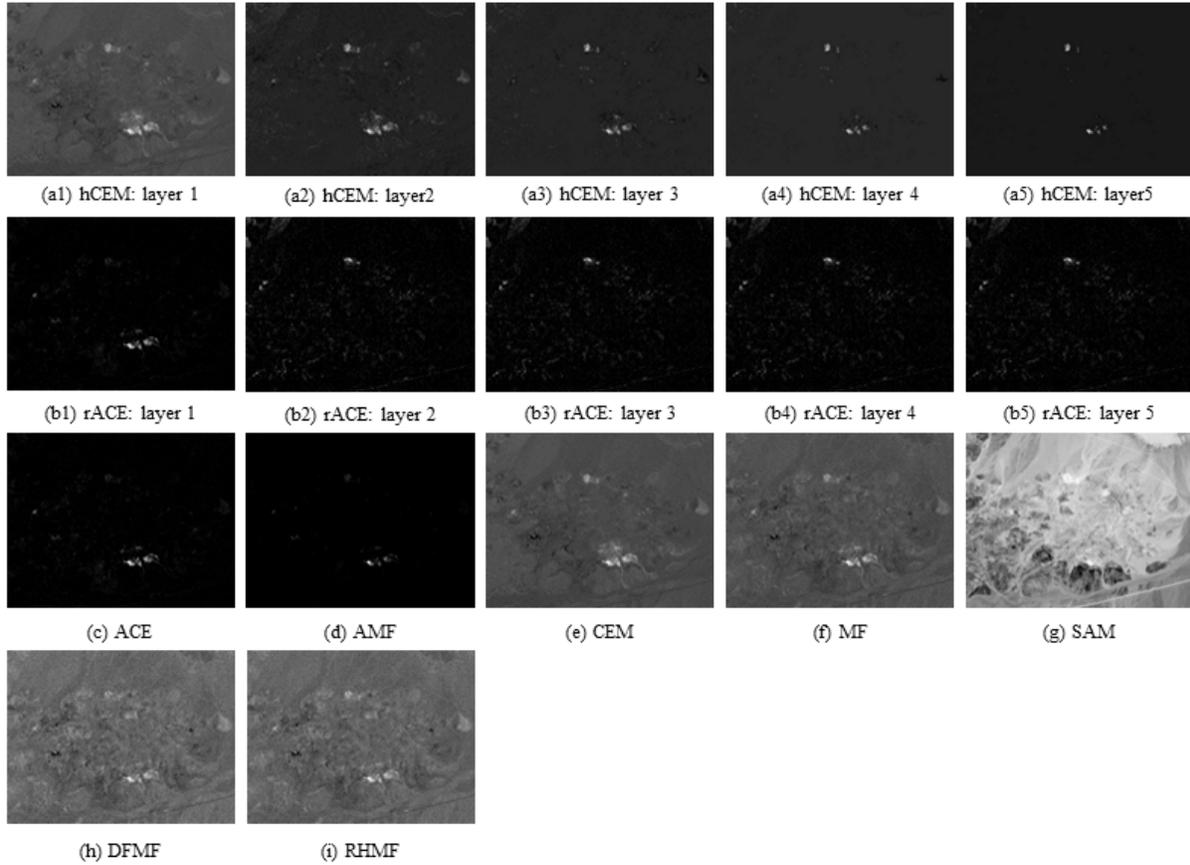


Fig. 9. Detection results of AVIRIS Cuprite data. First row (a1)-(a5): results of our hCEM of 1st-5th layers. Second row (b1)-(b5): results of rACE of 1st-5th layers. Third row and fourth row (c)-(i): results of ACE, AMF, CEM, MF, SAM, DFMF and RHMF. The first layer results of rACE and hCEM also represent the results of the original ACE and CEM algorithm. All the detection results are normalized to  $[0,1]$  for comparison. Notice although rACE converges and stops at the 3rd layer, for better illustration, we also list its subsequent 4-5th layer's detection results. rACE fails to detect the targets mainly because the proportion of the target pixels in this data is too small and accumulative error of other undesired background pixels decrease its performance.

In this experiment, the hCEM iteration stops after 6 layers of filtering, and we obtain the final detection results. Fig. 9 shows the detection results on Cuprite data. The first row of Fig. 9 shows the detection results of our hCEM of 1st-5th layers. The second row of Fig. 9 shows the detection results of rACE of 1st-5th layers. The third row of Fig. 9 shows the detection results of ACE, AMF, CEM, MF, SAM, DFMF and RHMF. It should be noticed that the publicly available Cuprite data were captured by AVIRIS in 1997 while the corresponding target location map was produced by Tricorder software in 1995 [27], thus we can only make a qualitative analysis of the performances of different target detection algorithms based on this target map. Although the Tricorder map is not exactly the same as the groundtruth, it still can be observed in Fig. 9 that the highest detection scores of the buddingtonite by our hCEM algorithms generally correspond with those pixels belong to the Tricorder map. We can clearly see that our hCEM significantly outperforms than rACE algorithm and its performance is enhanced with the increasing number

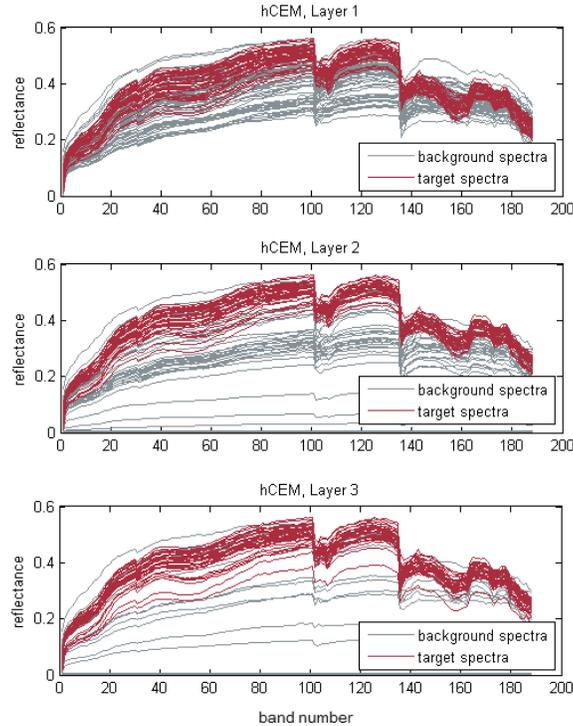


Fig. 10. Spectral suppression process under different layers. We select 63 background pixels (marked as white “+” in Fig. 8 (c)) and all target pixels (marked as red points in Fig. 8 (c)) as observation points. Spectral curves of these pixels in the 1st-3rd layers of hCEM are plotted. After the 3rd layers’ filtering process, most background spectra are suppressed while almost all the target spectra have not been changed.

of layers. In this experiment, rACE fails to detect the buddingtonite targets mainly because the proportion of the target pixels in this data is too small and accumulative error of other undesired background pixels decrease its performance. To better observe the changes of target and background spectra over the iterative process, we select 63 background pixels (marked as white “+” in Fig. 8 (c)), and all target pixels (marked as red points in Fig. 8 (c)) as observation points. The spectral curves of these points under different hCEM layers are plotted in Fig. 10. We can see after the 3rd iteration, most background spectral curves have been suppressed closed to zero while the all of the target spectrum almost have not changed.

### C. Experiment on AVIRIS Data II

In this experiment, another real hyperspectral image is used which is also collected by the AVIRIS. The scene is part of the airport in San Diego, America. The size of each band is  $200 \times 200$  pixels, and the band number of each pixel is 189 (low SNR and water absorption bands were removed). The first band of the hyperspectral image is shown in Fig. 11 (a) and Fig. 11 (b) shows the groundtruth. We can see that three airplanes are located in the left half of the image. All the above-mentioned algorithms and our hCEM algorithm are tested in this experiment. The target spectrum  $\mathbf{d}$  used in this experiment is obtained by averaging all the target spectra of this image.

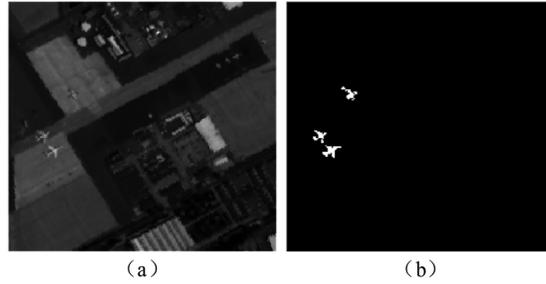


Fig. 11. (a) First band of the AVIRIS image. (b) Truth distribution of the targets (groundtruth).

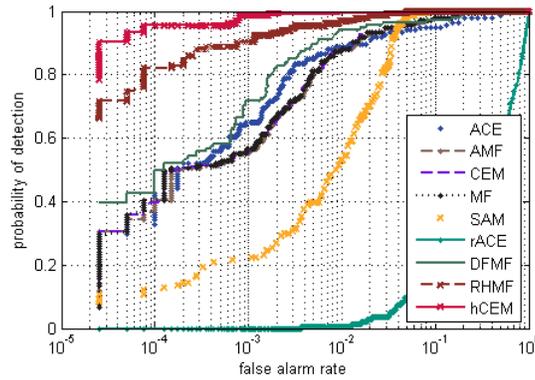


Fig. 12. The ROC curves of ACE, AMF, CEM, MF, SAM, DFMF, RHMF, rACE and our hCEM on the AVIRIS San Diego data. Our hCEM gets the best detection results on this data set.

hCEM algorithm stops after 14 layers of filtering, and we obtain the final detection results. Fig. 13 shows the detection results of different methods. The first row of Fig. 13 shows the detection results of our hCEM of its 1st-5th layers. The second row of Fig. 13 shows the detection results of rACE of its 1st-5th layers. The third row of Fig. 13 shows the detection results of ACE, AMF, CEM, MF, SAM, DFMF and RHMF. We again see the gradually increasing performance of our hCEM algorithm with the increasing number of layers. Fig. 12 shows the ROC curves of the detection results of all above mentioned methods. Our hCEM algorithm obtains the best detection results on this data set, which significantly outperforms than all the other methods.

We also compare the time performances of different detection methods, as shown in Table I. We test all the above mentioned methods on an Intel i5 PC with 6G RAM. The programming environment is Matlab 2010b. The test image is AVIRIS San Diego data with pixel number  $N = 200 \times 200$  and band number  $D = 224$ . Although hCEM is a multi-layer detection algorithm, it still reaches a satisfactory time performance. In this experiment, hCEM takes 14 layers to stop. Thus hCEM takes only about  $\frac{3.927}{14} = 0.281$  seconds to complete one round of detection.

TABLE I  
TIME PERFORMANCES OF DIFFERENT ALGORITHMS ON AVIRIS SAN DIEGO DATA WITH PIXEL NUMBER  $N = 200 \times 200$  AND BAND  
NUMBER  $D = 189$

Algorithm	Layers	Total Time (s)	Time / Layer (s)
ACE	–	0.881	0.881
AMF	–	0.542	0.542
CEM	–	0.268	0.268
MF	–	0.460	0.460
SAM	–	0.105	0.105
DFMF	–	0.822	0.822
RHMF	–	96.35	96.35
rACE	4	1.644	0.411
hCEM	14	3.927	0.281

#### D. Comparing on Different Parameters

$\lambda$  is an important parameter in our hCEM algorithm. In Section III(C), we have introduced a method to determine an appropriate value of  $\lambda$  by controlling the probability  $P_\xi(\lambda)$ .  $P_\xi(\lambda)$  represents the probability of a target spectrum being miss-transformed by the non-linear function  $q(y_T)$  during the filtering process. Clearly,  $P_\xi(\lambda)$  should be small enough to guarantee the transformation not to damage the target spectra. In AVIRIS San Diego hyperspectral data experiment, we have  $\mathbf{w}^T \mathbf{d} = 1$  and  $\mathbf{w}^T \mathbf{C} \mathbf{w} = 0.127$ . Thus we have

$$\begin{aligned}
 P_\xi(\lambda) &= F\left(\frac{1}{\lambda} \log\left(\frac{1}{\xi}\right); \mathbf{w}^T \mathbf{d}, \mathbf{w}^T \mathbf{C} \mathbf{w}\right) \\
 &= F\left(\frac{1}{\lambda} \log\left(\frac{1}{\xi}\right); 1, 0.127\right).
 \end{aligned} \tag{31}$$

Fig. 14 plots the shape of  $P_\xi(\lambda)$  with different  $\xi$  on AVIRIS San Diego hyperspectral data. We can see that the probability  $P_\xi(\lambda)$  is monotonous decreasing as  $\lambda$  increases, which means if we choose a larger value of  $\lambda$ , we will get a smaller probability that a target spectrum is suppressed. We can also see  $P_\xi(\lambda)$  drops to 0 quickly when  $\lambda$  is greater than 20, which means the nonlinear operation will no longer change the target spectra when  $\lambda$  is set greater than 20.

In order to further test the robustness of our hCEM algorithm, the ROC curves of with five different  $\lambda$  are plotted (See Fig. 15), where the value of  $\lambda$  is set to 200, 100, 50, 20 and 10 respectively. The iteration stops after 14, 14, 12, 11 and 5 layers of filtering respectively. As we can see, hCEM improves the original CEM detection results when  $\lambda$  is set to 200, 100, 50 and 20. However, when we further reduce  $\lambda$  to 10, the detection performance begins to drop. This is mainly because the probability  $P_\xi(\lambda)$  rises quickly when  $\lambda$  is greater than 20 (see Fig. 14). Nevertheless, Fig. 15 still suggests that our hierarchical approach works under a wide range of its parameter settings. As long as we keep  $\lambda > 20$ , satisfactory results can be obtained. However, setting larger value of  $\lambda$  does not mean the better. Fig. 16 plots the average output energy with different layers and different  $\lambda$  on AVIRIS San Diego hyperspectral data. We can see if we set larger value of  $\lambda$ , the convergence speed will be slower, and hCEM need more steps

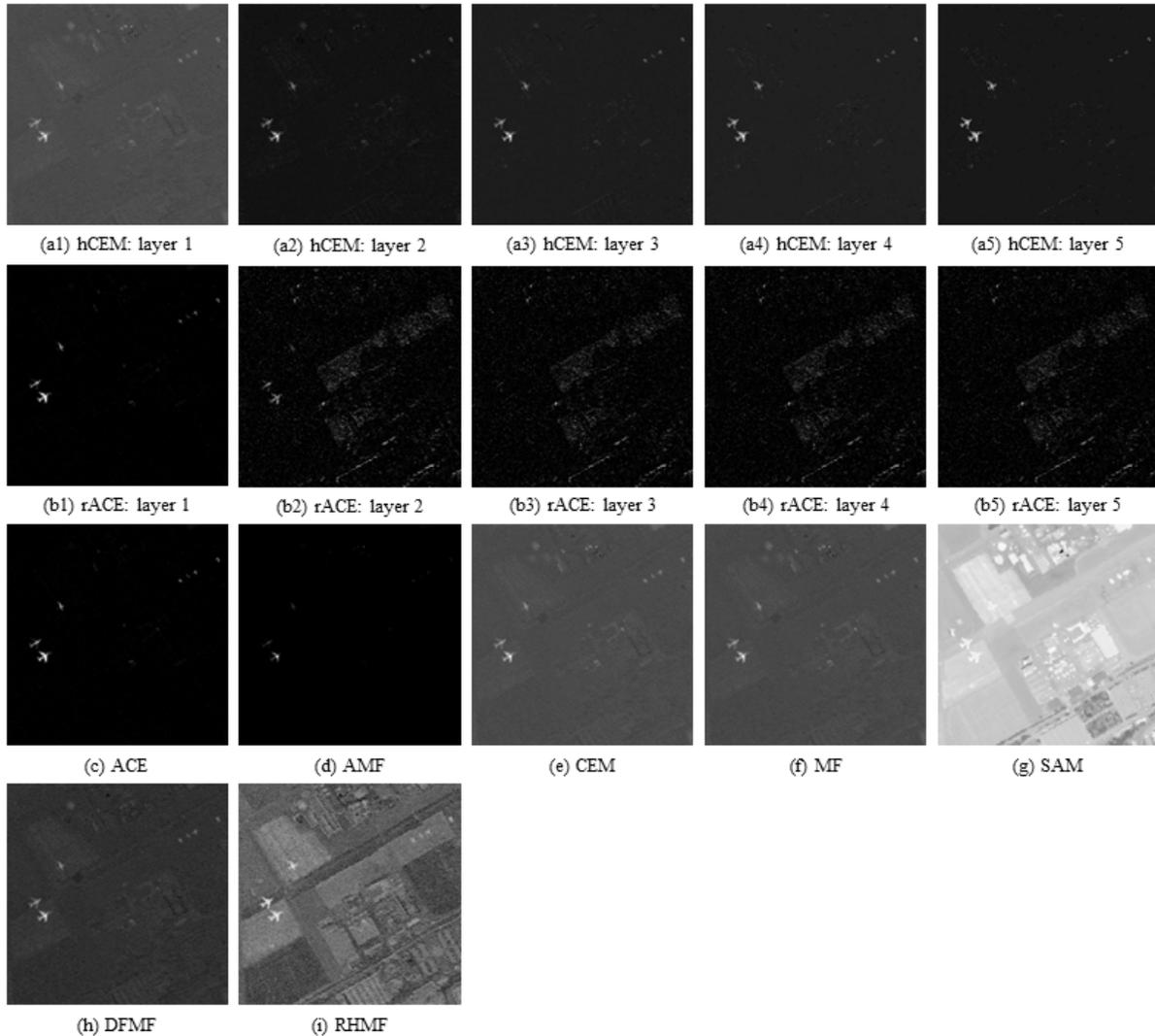


Fig. 13. Detection results of different methods on the AVIRIS San Diego data. First row (a1)-(a5): results of our hCEM of 1st-5th layers. Second row (b1)-(b5): results of rACE of 1st-5th layers. Third row and fourth row (c)-(i): results of ACE, AMF, CEM, MF, SAM, DFMF and RHMF. The first layer results of rACE and hCEM also represent the results of the original ACE and CEM algorithm. All the detection results are normalized to  $[0,1]$  for comparison. Notice rACE converges and stops at the 4th layer. For better illustration, we also list its subsequent 5th layer's detection results. rACE and SAM fail to detect the targets and similar reasons have been explained in the last subsection.

to get the final detection result. Therefore, the practical choice of  $\lambda$  should be a trade-off between the detection performance and the time efficiency.

## V. CONCLUSION

In this paper, we propose a new hyperspectral target detection algorithm, the hCEM algorithm, which suppresses undesired background spectra and holds the target spectra through a layer-by-layer filtering procedure. In each layer we construct a better detector than previous layers, and we theoretically prove the convergency of our algorithm.

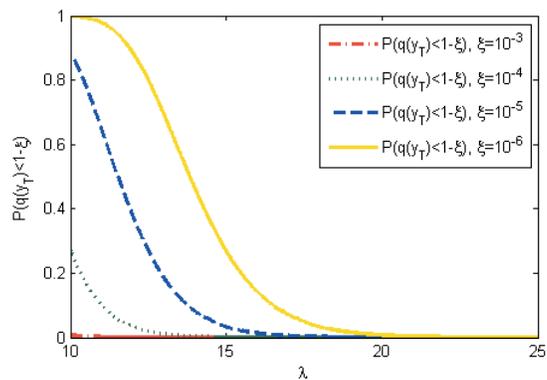


Fig. 14. Shape of the probability  $P_{\xi}(\lambda)$  with different  $\xi$  of AVIRIS San Diego hyperspectral data.

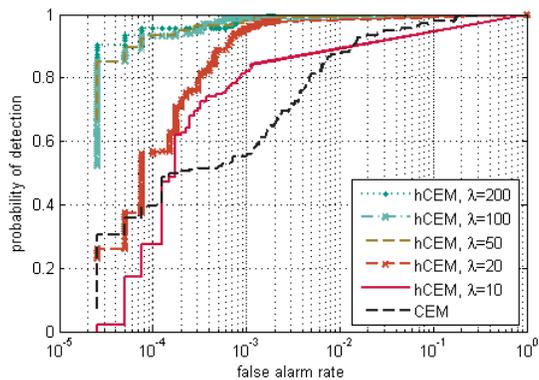


Fig. 15. The ROC curves of the original CEM algorithm and hCEM algorithm on the AVIRIS hyperspectral image with different values of parameter  $\lambda$ .

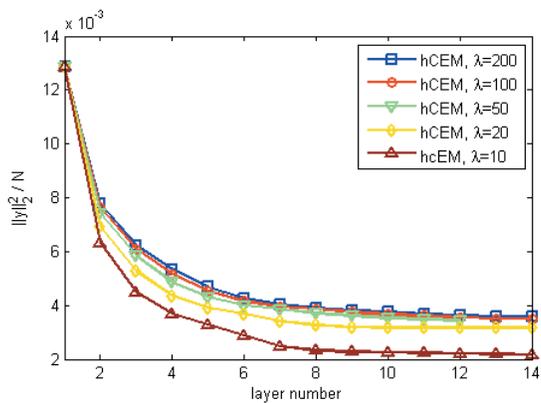


Fig. 16. The average output energy of the hCEM algorithm on the AVIRIS hyperspectral image of different layers with different values of parameter  $\lambda$ .

Experimental results on two real hyperspectral images and one synthetic image suggest that our hCEM algorithm is robust, and has significantly improved the classical CEM detectors. hCEM also outperforms other detectors like ACE, MF, AMF, SAM, DFMF, RHMF and another multi-layer detector rACE.

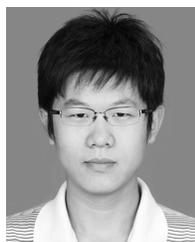
#### ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the three anonymous reviewers for their very useful comments and suggestions which greatly improve the quality of this paper.

#### REFERENCES

- [1] D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson, "Is there a best hyperspectral detection algorithm?" in *Proc SPIE*, Orlando, USA, 2009.
- [2] 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.
- [3] 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.
- [4] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz, "The spectral image processing system (SIPS) - Interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*, vol. 44, pp. 145-163, 1993.
- [5] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1-16, Jan. 2001.
- [6] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779-785, Jul. 1994.
- [7] 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.
- [8] H. Ren and C.-I. Chang, "Target-constrained interference-minimized approach to subpixel target detection for hyperspectral images," *Opt. Eng.*, vol. 39, pp. 3138-3145, Jul. 2000.
- [9] 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, May. 2014.
- [10] T.-M. Tu, C.-H. Chen, and C.-I. Chang, "A noise subspace projection approach to target signature detection and extraction in an unknown background for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 1, pp. 171-181, Jan. 1998.
- [11] D. Manolakis, C. Siracusa, and G. Shaw, "Hyperspectral subpixel target detection using the linear mixing model," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 7, pp. 1392-1409, Jul. 2001
- [12] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Sparse representation for target detection in hyperspectral imagery," *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 3, pp. 629-640, Jun. 2011.
- [13] 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.
- [14] L. Zhang, L. Zhang, D. Tao, X. Huang and B. Du, "Hyperspectral Remote Sensing Image Subpixel Target Detection Based on Supervised Metric Learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4955-4965, Aug. 2014
- [15] Y. Dong, B. Du and L. Zhang, "Target Detection Based on Random Forest Metric Learning," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, Article in Press, Apr. 2015
- [16] H. Kwon and N. M. Nasrabadi, "Kernel Spectral Matched Filter for Hyperspectral Imagery," *Int. J. Comput. Vis.*, vol. 71, no. 2, pp. 127-141, Feb. 2007.
- [17] H. Kwon, and N. M. Nasrabadi, "Kernel Matched Subspace Detectors for Hyperspectral Target Detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 2, Feb. 2006.

- [18] L. Capobianco, A. Garzelli, and G. Camps-Valls, "Target detection with semisupervised kernel orthogonal subspace projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3822-3833, Nov. 2009.
- [19] L. Zhang, L. Zhang, D. Tao and X. Huang, "Sparse Transfer Manifold Embedding for Hyperspectral Target Detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1030-1043, Feb. 2014.
- [20] J. Broadwater and R. Chellappa, "Hybrid detectors for subpixel targets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 11, pp. 1891-1903, Nov. 2007.
- [21] Z. Shi and S. Yang, "Robust high-order matched filter for hyperspectral target detection," *Electron. Lett.*, vol. 46, no. 15, pp. 1065-1066, Jul. 2010.
- [22] Z. Shi, S. Yang and Z. Jiang "Target detection using difference measured function based matched filter for hyperspectral imagery," *Optik*, vol. 124, no. 17, pp. 3017-3021, Sep. 2013.
- [23] T. Wang, B. Du, and L. Zhang, "An Automatic Robust Iteratively Reweighted Unstructured Detector for Hyperspectral Imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 7, no. 6, pp. 2367-2382, Jun. 2014.
- [24] G. A. Shaw and H.-h. K. Burke, "Spectral imaging for remote sensing," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 3-28, 2003.
- [25] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44-57, Jan. 2002.
- [26] J.M. Bioucas-Dias et al., "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354-379, Apr. 2012.
- [27] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014-2039, Jun. 2011.
- [28] X. Fan, Y. Zhang, F. Li, Y. Chen, T. Shao and S. Zhou, "A robust spectral target recognition method for hyperspectral data based on combined spectral signatures," *IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, pp. 4328-4331, 2011.
- [29] J. Sherman, W. J. Morrison, "Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix," *Ann. Math. Statist.* vol. 21, pp.124-127, 1950.
- [30] R. N. Clark, G. A. Swayze, A. J. Gallagher, T. V. V. King, and W. M. Calvin, "The U.S. geological survey digital spectral library: Version 1: 0.2 to 3.0 microns," *U.S. Geological Survey, Denver, CO, Open File Rep.*, pp. 93-592, 1993.
- [31] Y.-C. C. Chang, H. Ren, C.-I. Chang, and R. S. Rand, "How to design synthetic images to validate and evaluate hyperspectral imaging algorithms," in *Proc. SPIE Int. Soc. Opt. Eng.*, Orlando, FL, United States, 2008.
- [32] C.-I. Chang, "Multiparameter receiver operating characteristic analysis for signal detection and classification," *IEEE Sensors J.*, vol. 10, no. 3, pp. 423-442, Mar. 2010.
- [33] AVIRIS Cuprite data set. [Online]. Available: <http://aviris.jpl.nasa.gov/html/aviris.freedata.html>.
- [34] USGS map of the distribution of different minerals in the Cuprite mining district in Nevada. [Online]. Available: <http://speclab.cr.usgs.gov/PAPERS/tetracorder/>.



**Zhengxia Zou** received the B.S. degree from the School of Astronautics, Beihang University, Beijing, China, in 2013. He is currently working toward the Ph.D. degree in the Image Processing Center, School of Astronautics, Beihang University. His research interests include hyperspectral target detection, and pattern recognition.



**Zhenwei Shi** Zhenwei Shi (M13) received his Ph.D. degree in mathematics from Dalian University of Technology, Dalian, China, in 2005. He was a Postdoctoral Researcher in the Department of Automation, Tsinghua University, Beijing, China, from 2005 to 2007. He was Visiting Scholar in the Department of Electrical Engineering and Computer Science, Northwestern University, U.S.A., from 2013 to 2014. He is currently a professor in the Image Processing Center, School of Astronautics, Beihang University. His current research interests include remote sensing image processing and analysis, computer vision, pattern recognition, and machine learning.

He has authored or co-authored over 100 scientific papers in refereed journals and proceedings, including the IEEE Transactions on Pattern Analysis and Machine Intelligence, the IEEE Transactions on Neural Networks, the IEEE Transactions on Geoscience and Remote Sensing, the IEEE Geoscience and Remote Sensing Letters and the IEEE Conference on Computer Vision and Pattern Recognition.