Entropy-Based Convex Set Optimization for Spatial–Spectral Endmember Extraction From Hyperspectral Images

Dharambhai Shah , *Student Member, IEEE*, Tanish Zaveri , *Senior Member, IEEE*, Yogesh N. Trivedi , and Antonio Plaza , *Fellow, IEEE*

Abstract—Spectral unmixing is an important problem for remotely sensed hyperspectral data exploitation. Automatic spectral unmixing can be viewed as a three-stage problem, where the first stage is subspace identification, the next one is endmember extraction, and the final one is abundance estimation. In this sequence, endmember extraction is the most challenging problem. Many researchers have attempted to extract endmembers from hyperspectral images using spectral information only. However, it is well known that the inclusion of spatial information can improve the endmember extraction task. In this article, we introduce a new endmember extraction algorithm that exploits both spectral and spatial information. A main innovation of the proposed algorithm is that spatial information is exploited using entropy, while spectral information is exploited using convex set optimization. In the literature, none of the spatial–spectral algorithms has used entropy as spatial information. The inclusion of this entropy-based spatial information improves the accuracy of the endmember extraction process. The results obtained by the proposed algorithm are compared (using a variety of metrics) with those obtained by other state-of-the-art methods, using both synthetic and real datasets. Our experimental results demonstrate that the proposed algorithm outperforms many available algorithms.

Index Terms—Convex set optimization, endmember extraction, entropy, hyperspectral imaging, spectral unmixing.

I. INTRODUCTION

REMOTE sensing is used in various applications of Earth science, geography, land surveying, and Mars exploration [1]. Hyperspectral sensors have opened up new avenues in the field of remote sensing by collecting information in hundreds of (narrow) bands from the electromagnetic spectrum. Hyperspectral sensors provide precise and robust information in the analysis of geological features, soil, vegetation, and the

environment. Spaceborne and airborne systems use hyperspectral sensors for many different purposes, including target detection, material mapping, material identification, and surface property identification. However, the accuracy of these tasks strongly depends upon the spatial resolution of the captured image. Due to the (generally low) spatial resolution of hyperspectral sensors, many pixels are mixed in nature, i.e., they consist of more than one pure spectral material. Other reasons for the formation of mixed pixels include multiple scattering and intimate mixtures of materials [2]. Such mixing can be linear or nonlinear, depending on how pure spectral signatures (called endmembers in hyperspectral imaging terminology) are combined in a mixed pixel. Most works assume the linear mixing model, as it is a simple approximation to real-world applications [2]. In this model, the concept of endmember is a key aspect, since endmembers are spectrally distinct signatures of pure materials that can be used to model (linearly or nonlinearly) the mixed pixels in the scene.

Spectral unmixing decomposes mixed pixels into a combination of endmembers, weighted by their corresponding (subpixel) abundance fractions. From an operational point of view, unmixing can be either supervised or unsupervised. In unsupervised unmixing, the hyperspectral image cube is the only input [3]. In the supervised approach, there is manual consideration of the number of materials and their respective spectra. As shown in Fig. 1, an unsupervised hyperspectral unmixing problem can be divided into three subproblems [3]. The first one is the hyperspectral data subspace estimation, which finds the number of endmembers (Q). The second block is the endmember extraction itself, which finds a matrix ($\mathbf{M}$) of pure endmember spectra from the image or from a library. Finally, the third block is an abundance estimation step, which finds the abundance ($\boldsymbol{\alpha}$) of all the individual endmembers in each mixed pixel. The most important block (and the one we specifically address in this article) is the endmember extraction one, which provides prior information of pure materials for target detection [4], abundance mapping [5], change detection [6], and object classification [7]. As a result, proper extraction of pure endmembers is very important in hyperspectral data exploitation [8].

There are mainly three types of approaches in the literature for endmember extraction [2]. The first one is the statistical approach [9], which formulates the unmixing problem as a statistical inference one [10]. This approach is generally
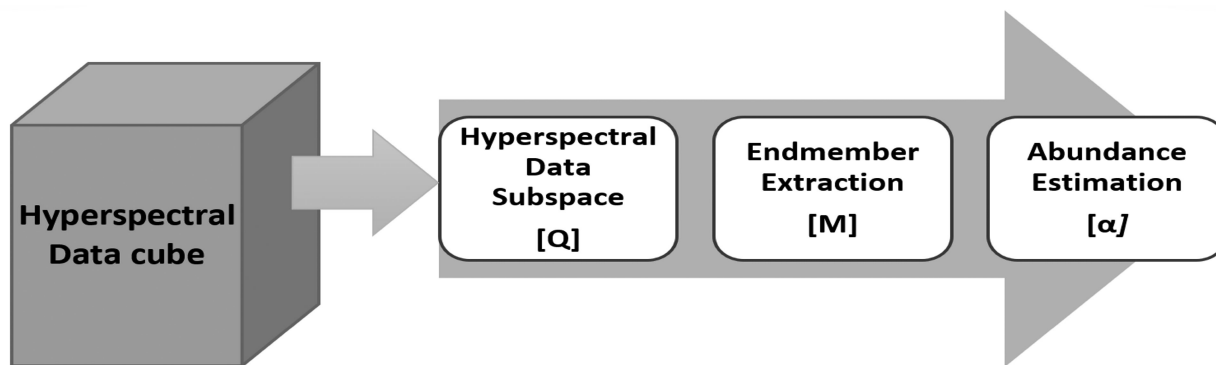
Fig. 1. Unsupervised spectral unmixing chain.

expensive from a computational viewpoint. The second approach is sparse regression [11], which requires a detailed spectral library containing instances of the endmembers present in the scene. The third approach is the geometrical one, which assumes that there are pure pixels present in the image. Geometrical approaches are actually very popular in the hyperspectral imaging literature, perhaps due to their clear conceptual meaning and low computational complexity. Recently, endmember extraction has also been approached by deep-learning-based methods [12], [13].

Simplex volume maximization by multidimensional geometry and similarity of spectral signatures are the two fundamental criteria [14] of geometrical endmember extraction approaches. Simplex volume maximization-based algorithms are mostly based on the concept that the volume of the simplex formed by any combination of pixels is always less than the volume contained by the simplex formed by the purest pixels in the hyperspectral image [15]. Many algorithms such as NFINDR [16], simplex growing algorithm (SGA) [17], successive volume maximization (SVMAX) [18], and alternating volume maximization (AVMAX) [18] have been developed based on this fact. The second criterion considers the similarity between different spectral signatures to find pure pixels. Two popular algorithms in this category are vertex component analysis (VCA) [19] and TRIple-P:P-norm based pure pixel identification (TRIP) [20], which are based on similarity measures. The pixel purity index (PPI) algorithm [21] generates random skewers. All pixels are projected onto these skewers to find the associated projection scores. The purity of pixels can be found using maximum and minimum thresholds. Independent component analysis (ICA) [22], a method for separating a multivariate signal into additive subcomponents, has also been used for the purpose of endmember extraction from hyperspectral images.

All the aforementioned algorithms focus only on exploiting the spectral information of the data alone. However, hyperspectral sensors are designed to capture spatial as well as high spectral ground information. This conflicts with the fact that most techniques available in the literature were designed from a spectroscopic viewpoint, neglecting the spatial features present in the image. The earliest attempt that uses both spatial and spectral information was the automatic morphological endmember extraction (AMEE) [23]. AMEE used the concept of mathematical morphology to combine the spatial information contained in the data together with the spectral one. Three popular algorithms that improve the endmember extraction process using spatial and spectral information are spatial preprocessing for endmember extraction (SPEE) [24], region-based spatial preprocessing (RBSPP) [25], and spatial–spectral preprocessing (SSPP) [26]. The SPEE algorithm spatially weighs the spectral information related to each pixel for endmember extraction. The RBSPP exploits spectral information more effectively. The algorithm guides the endmember finding process to image zones, which are both spatially homogeneous and spectrally pure. The SPEE algorithm primarily accounts for the spatially homogeneous areas in the scene (regardless of their spectral purity). The SSPP fuses spatial and spectral information (at the preprocessing level) for improving the extraction process. Several other endmember extraction algorithms have been presented in the literature that incorporate spatial information along with spectral information [27]–[39].

In this article, we develop a new algorithm for endmember extraction that combines both spectral and spatial information. The main innovation of the proposed algorithm is that spatial information is exploited using entropy, while spectral information is exploited using convex set optimization. The concept of entropy [40] has been used in many image processing techniques, including registration, reconstruction, segmentation, classification, and compression. Few researchers have also applied entropy in hyperspectral compression [41], band selection [42], and unmixing [22]. Bayliss *et al.* [22] developed an ICA-based algorithm for unmixing based on the entropy between spectral signatures. ICA was selected under the assumption that components are statistically independent. However, in the data acquisition process of a hyperspectral sensor, the sum of abundance fractions associated with each pixel adds to one under the abundance sum-to-one constraint (ASC). As a result, the sources (endmembers) are not statistically independent [43]. ICA uses the entropy of various spectra as a spectral feature, while the proposed algorithm uses the entropy of each band as a spatial feature.

At this point, it is important to emphasize that many hyperspectral endmember extraction algorithms [16]–[18], [24]–[26] have used the concept of convex geometry optimization. The proposed algorithm also uses this concept but incorporating the

idea of entropy to characterize the spatial information in the scene. Specifically, our algorithm uses low- and high-entropy bands in the convex set optimization, which represents a new concept of entropy-based convex set optimization. Here, the entropy characterizes the spatial heterogeneity of each band (which is useful for extracting rare or anomalous endmembers in the scene). Due to this feature, our algorithm can accurately extract rare and anomalous endmembers, which ultimately increases the accuracy of the endmember extraction stage.

Endmember extraction algorithms can be implemented in a parallel or sequential manner [14]. Parallel implementation determines all endmembers simultaneously, while sequential implementation determines the endmembers one by one. The SGA is the sequential version of the original NFINDR algorithm. The computational complexity of sequential implementations is low compared to that of parallel implementations. However, parallel implementations may be beneficial when extracting distinctive pixels effectively. In this regard, the proposed approach extracts pure and distinct endmembers using a parallel approach.

The remainder of this article is organized as follows. Section II presents a problem statement along with our newly proposed endmember extraction algorithm. Section III describes the synthetic and real image dataset used in the experimental evaluation, and the performance metrics adopted to compare endmember extraction algorithms. In Section IV, various endmember extraction algorithms are compared based on the considered metrics, and the quality of the extracted endmembers in terms of abundance estimation is also tested. Section V concludes this article with some remarks and hints at plausible future research lines.

II. PROBLEM STATEMENT AND PROPOSED ALGORITHM

Let us denote a mixed pixel in the hyperspectral image as an $(L \times 1)$-dimensional vector

$$\mathbf{y} = \mathbf{M}\boldsymbol{\alpha} + \mathbf{n} \qquad (1)$$

where $\mathbf{M}$ is an $L \times Q$ matrix, with Q and L, respectively, denoting the number of endmembers and the number of bands in the original hyperspectral image. In (1), $\mathbf{n}$ is a noise vector of size $L \times 1$, which is assumed to be Gaussian in nature. $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_Q]^T$ denotes the abundance vector of size $Q \times 1$, which satisfies the following two constraints:

1) Abundance Non-negativity Constraint (ANC):

$$\alpha_i \geq 0, i = 1, 2, \ldots, Q. \qquad (2)$$

2) Abundance Sum-to-one Constraint (ASC):

$$\sum_{i=1}^{Q} \alpha_i = 1. \qquad (3)$$

Let us denote the hyperspectral image as $\mathbf{Y} \equiv [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_L]^T \equiv [\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^Z]$, which contains Z mixed

TABLE I
NOTATIONS USED IN THIS ARTICLE

Symbol	Description
L	Number of bands in the hyperspectral image
Q	Hyperspectral subspace dimension
$\mathbf{y}$	Mixed pixel vector of size $L \times 1$
$\mathbf{M}$	Endmember matrix of size $L \times Q$
$\boldsymbol{\alpha}$	Abundance vector $[\alpha_1, \alpha_2, \ldots, \alpha_Q]^T$
$\mathbf{n}$	Noise vector of size $L \times 1$
p	Probability of occurrence
p_i	Probability of ith grey shade in single band
p_i^k	Probability of ith grey shade in k^{th} band
U	Height of the hyperspectral image (in pixels)
V	Width of the hyperspectral image (in pixels)
W_i	Frequency of ith grey value in a single band
Z	Number of pixels $(U \times V)$ in the hyperspectral image
$\mathbf{I}$	Grey image of size $U \times V$ pixels
$\mathbf{Y}$	Hyperspectral image of size $L \times Z$ pixels
$\bar{\mathbf{Y}}$	Band-normalized hyperspectral image
G	Maximum shade value in the image
S_G	Set of grey scale values $\{0, 1, \ldots, G-1\}$
$\mathbf{y}_i$	i^{th} bands in the hyperspectral image
$\mathbf{y}^j$	j^{th} mixed pixel vectors in the hyperspectral image
y	Single pixel value
$\bar{y}$	Normalized pixel value
H_k	Entropy of kth band
$\mathbf{S}$	Band entropy-based ordered image (in ascending order)
$\mathbb{R}$	Set of real numbers
$\mathbb{R}_+$	Set of positive real numbers
C	Convex set
T_i	Number of convex set points for $(\mathbf{s}_i, \mathbf{s}_{L-i-1})$
$\mathbf{T}$	Set of numbers $\{T_1, T_2, \ldots, T_{L/2}\}$
N	Number of optimized convex set points
$\mathbf{s}_l$	Low entropy band of normalized image
$\mathbf{s}_h$	High entropy band of normalized image
$D(\mathbf{x})$	Euclidean distance of point $\mathbf{x}$
$\hat{\mathbf{M}}$	Extracted endmember matrix of size $L \times Q$

pixels of length L. $\mathbf{Y}$ is defined as follows:

$$\mathbf{Y} \equiv \begin{bmatrix} y_1^1 & y_1^2 & y_1^3 & \cdots & y_1^Z \\ y_2^1 & y_2^2 & y_2^3 & \cdots & y_2^Z \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L^1 & y_L^2 & y_L^3 & \cdots & y_L^Z \end{bmatrix} \qquad (4)$$

where each band $\mathbf{y}_i$ of size $1 \times Z$ is represented as a row vector $\mathbf{y}_i \equiv [y_i^1, y_i^2, y_i^3, \ldots, y_i^Z]$, and each mixed pixel vector $\mathbf{y}^i$ of size $L \times 1$ is represented as a column vector $\mathbf{y}^i \equiv [y_1^i, y_2^i, y_3^i, \ldots, y_L^i]^T$. Here, $Z = U \times V$ is the number of pixels in the original image. The height and width of each band are, respectively, U and V. Various notations used in this article are shown in Table I.

The proposed algorithm takes as an input $\mathbf{Y}$ and the subspace dimension (number of endmembers) Q. The result of the proposed algorithm is a matrix of endmembers $\hat{\mathbf{M}}$. In order to extract such a matrix, the proposed algorithm follows three steps. First, it uses entropy of each normalized band and finds a new matrix using it. This first step explores the spatial information contained in the scene. The second step solves a convex set optimization problem using the spatial information obtained in the first step. This second step explores the spectral information contained in

the scene. Finally, our algorithm removes unnecessary spectral signatures, if any. We discuss each step in detail in the following.

Many hyperspectral data analysis techniques use dimensionality reduction as a preprocessing step, which aims to remove the redundant spectral information while preserving only critical information for subsequent processing. Hyperspectral dimensionality reduction can be achieved through feature extraction or band selection [44], [45]. Feature extraction methods such as principal component analysis and minimum noise fraction transform the original data into reduced feature spaces by means of different criteria, whereas band selection aims to select a small subset of hyperspectral bands to reduce the burden of heavy computations [46]. Regardless of whether dimensionality reduction is performed using band selection or feature extraction, some spectral information in the original image will be lost. The proposed algorithm uses a different approach and exploits all the spectral bands in the original image. In other words, we avoid dimensionality reduction and simply use band normalization as a preprocessing step. The advantage of using band normalization is that it can allow us to find the entropy of each band. Since different bands have a different dynamic range (max value–min value), our band normalization strategy is intended to make all bands similar in terms of dynamic range. In our case, band normalization for each ith band ($\mathbf{y}_i$) is conducted as follows:

$$\bar{y}_i^j = \frac{\left(y_i^j - \min(\mathbf{y}_i)\right)}{(\max(\mathbf{y}_i) - \min(\mathbf{y}_i))} \quad \forall y_i^j \in \mathbf{y}_i. \quad (5)$$

Each pixel value y_i^j of band ($\mathbf{y}_i$) is normalized as per (5) and converted to a new normalized value $\bar{y}_i^j$, which is in the range of [0, 1]. The aforementioned band normalization process [shown in (5)] is repeated for all the bands of the hyperspectral image. A band-normalized hyperspectral image $\bar{\mathbf{Y}} \equiv [\bar{\mathbf{y}}_1, \bar{\mathbf{y}}_2, \ldots, \bar{\mathbf{y}}_L]^T \equiv [\bar{\mathbf{y}}^1, \bar{\mathbf{y}}^2, \ldots, \bar{\mathbf{y}}^Z]$ is defined as

$$\bar{\mathbf{Y}} \equiv \begin{bmatrix} \bar{y}_1^1 & \bar{y}_1^2 & \bar{y}_1^3 & \cdots & \bar{y}_1^Z \\ \bar{y}_2^1 & \bar{y}_2^2 & \bar{y}_2^3 & \cdots & \bar{y}_2^Z \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{y}_L^1 & \bar{y}_L^2 & \bar{y}_L^3 & \cdots & \bar{y}_L^Z \end{bmatrix}. \quad (6)$$

A. Entropy as Spatial Information

For a k-state system, Shannon [40] defined the entropy as

$$H = -\sum_{i=1}^{k} p_i \log p_i \quad (7)$$

where p_i is the probability of occurrence for the ith event, and $\sum p_i = 1, 0 \leq p_i \leq 1$. Shannon's entropy is very popular in the field of communications. Many researchers have extended the concept of entropy for image processing purposes [47].

Let $\mathbf{I}$ be a grayscale image of size $U \times V$ and $S_G \in \{0, 1, \ldots, G-1\}$ be the set of associated grayscale values. Let G be maximum shade value in the image. Image $\mathbf{I}$ contains $Z = U \times V$ pixels. Let W_i be the frequency of the ith grayscale value, where $i \in S_G$. The entropy for $\mathbf{I}$ (grayscale image) is

defined as

$$H = -\sum_{i=0}^{G-1} p_i \log p_i, \quad p_i = W_i/Z. \quad (8)$$

We can extend the entropy definition in (8) for hyperspectral images and define the kth band entropy (H_k) as

$$H_k = -\sum_{i=0}^{G-1} p_i^k \log p_i^k; \quad p_i^k = W_i^k/Z, k = 1, 2, \ldots, L \quad (9)$$

Here, p_i^k is probability of having an ith gray shade in the kth band. The entropy of each band in $\bar{\mathbf{Y}}$ is calculated using (9) and denoted as $\{H_1, H_2, \ldots, H_L\}$. Entropy can be interpreted as a measure of order (or randomness) or as a measure of homogeneity [48]. Instead of looking at various interpretations, we can look at it as an expression of the number of states of a system. Lin [49], Jost [50], and other researchers have used Shannon's entropy concept for information-theoretic divergence between two probability distributions. In the proposed algorithm, this entropy concept is used to measure the divergence between the probability distribution of two bands. A system with many states has high entropy and a system with few states has low entropy. A band with low entropy exhibits fewer variations, and a band with high entropy exhibits more variations. Bands can be rearranged in ascending order based on their values of entropy. A new matrix $\mathbf{S}$ can be, thus, obtained from $\bar{\mathbf{Y}}$ in such a manner that low-entropy bands come first, and high-entropy bands come last. The matrix $\mathbf{S}$ is generated such that $H_k < H_{k+1}$ for each value of k, we have

$$\mathbf{S} \equiv [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_L]^T \equiv [\mathbf{s}^1, \mathbf{s}^2, \ldots, \mathbf{s}^Z]. \quad (10)$$

B. Spectral Information for Convex Set Optimization

In convex analysis [51], the affine hull of a set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_Q\}$ is defined as

$$\text{aff}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_Q\} = \left\{ \sum_{i=1}^{Q} \theta_i \mathbf{x}_i \Big| \boldsymbol{\theta} \in \mathbb{R}, \mathbf{1}_Q^T \boldsymbol{\theta} = 1 \right\}. \quad (11)$$

The immediate implication of (11) to the linear mixing model in (1) is that every mixed pixel vector $\mathbf{y}$ can be in the affine hull $\text{aff}\{x_1, x_2, \ldots, x_Q\}$. The affine hull follows the ASC constraint in (3), but not the ANC constraint in (2). A special case of an affine hull set is the convex hull set, which follows both (2) and (3). The convex hull set C is defined as

$$\text{conv}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_Q\} = \left\{ \sum_{i=1}^{Q} \theta_i \mathbf{x}_i \Big| \boldsymbol{\theta} \in \mathbb{R}_+, \mathbf{1}_Q^T \boldsymbol{\theta} = 1 \right\}. \quad (12)$$

The affine space subset that is closed under convex combinations is called a convex set C [51]. The convex set has the property that a convex combination is a linear combination of vectors, where all coefficients are nonnegative and sum up to 1. This property of a convex set describes well the linear mixture model in (1), including both the ANC in (2) and the ASC in (3). Coifman and Wickerhauser [52] developed algorithms based on entropy for best basis selection. In the proposed approach, we use this

entropy feature of bands for best basis (band) selection to find the best convex set. We define an optimization problem for best basis selection as follows:

Optimization Problem:

$$\operatorname*{minimize}_{N \in \mathbf{T}} \quad N - Q$$

$$\text{subject to} \quad N \geq Q. \tag{13}$$

Here, $\mathbf{T} = \{T_1, T_2, \ldots, T_{L/2}\}$ is a set of numbers, which represent the number of points involved in making a convex set C [using (12)] for two bands ($\mathbf{s}_l$ and $\mathbf{s}_h$) of the data. T_i is the number of convex set points for two-band ($\mathbf{s}_l = \mathbf{s}_i$ and $\mathbf{s}_h = \mathbf{s}_{L-i-1}$) data. $\mathbf{s}_l = \mathbf{s}_i$ is a low-entropy band and $\mathbf{s}_h = \mathbf{s}_{L-i-1}$ is a high-entropy band. The high-entropy band contains many heterogeneous objects, while the low-entropy band contains many homogeneous objects. The proposed model computes the convex set points from low- and high-entropy bands. Hence, this combination of homogeneity and heterogeneity features explores the variety of object information between two bands for extracting rare endmembers. The two-band combination from $\mathbf{S}$ results in $L/2$ values in $\mathbf{T}$. When the value of L is even, the set $\mathbf{T}$ exactly has $L/2$ values. When the value of L is odd, the middle band from $\mathbf{S}$ is removed to make it even. N is an outcome of optimization problem, as described in (13). This optimization problem selects the minimum number N from the available $L/2$ values in the set $\mathbf{T}$.

As shown in Fig. 1, the expected number of extracted endmembers is Q, which is also known as the hyperspectral subspace dimension. Generally, the value of Q is found by many hyperspectral subspace estimation techniques [3], [53], before extracting the endmembers from the image. The optimization problem in (13) thus relates to finding the optimum value of N such that the difference between N and Q is the absolute minimum, constrained by $N \geq Q$. The algorithm is of no use if $N < Q$ (this may happen in multispectral data) because it will extract fewer materials than subspace dimensions.

C. Removing Extra Points

As a result of the previous step, we obtain a number N, which is greater than or equal to Q. At this point, there are two possibilities.

1) $N = Q$: N is the number of materials extracted from the previous step. The hyperspectral subspace dimension is Q. N and Q should be the same. If this condition is true, then all the extracted materials are endmembers.
2) $N > Q$: The main goal of the proposed algorithm is to extract exactly Q endmembers. If the above step gives some extra points $(N - Q)$, then we have to remove $(N - Q)$ points. Preservation of pure signatures/materials needs to be taken care of while removing these extra points.

Convex set points can be arranged in a clockwise or anticlockwise manner. The advantage of arranging them in order is that every point/endmember will have two proper neighbors. If two points are very near, it means that they are very similar. If we remove one of them, it is likely that we are not losing any pure signature/spectra. The proposed algorithm eliminates materials

---

Algorithm 1: Proposed Algorithm.

1: **Inputs : $\mathbf{Y}, Q$** ▷ Inputs to the algorithm
2: %% *Step-1: Entropy as spatial information* %%
3:      **For** k=1 to L
4:          $\bar{\mathbf{y}}_k \leftarrow \mathbf{y}_k$ ▷ Band normalization for the kth band
5:          $H_k =$ entropy $(\bar{\mathbf{y}}_k)$ ▷ Entropy of the kth band
6:      **EndFor**
7:      $[\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_L]^T \leftarrow [\bar{\mathbf{y}}_1, \bar{\mathbf{y}}_2, \ldots, \bar{\mathbf{y}}_L]^T$ ▷ Entropy-based band sorting
8:      $\mathbf{S} \leftarrow \bar{\mathbf{Y}}$ ▷ Creation of new matrix from normalized hyperspectral image
9:      %% *Step-2: Spectral information for convex set optimization* %%
10:      **For** $i = 1$ to $L/2$
11:          $C_i = conv(\mathbf{s}_i, \mathbf{s}_{L-i-1})$ ▷ Calculate convex set for two-band data
12:          $T_i =$ count (C_i) ▷ Calculate the number of points in the convex set
13:      **EndFor**
14:      $\mathbf{T} = \{T_1, T_2, \ldots, T_{L/2}\}$ ▷ Set of number of convex set points
15:      $\operatorname*{minimize}_{N \in \mathbf{T}} N - Q$ subject to $N \geq Q$ ▷ Find N using optimization problem
16:      %% Step-3: Removing extra points %%
17:      **If** $N > Q$
18:          **For** k=1 to N
19:              $D(\mathbf{x}(i, j), \mathbf{x}(i_1, j_1))$ ▷ Calculate Euclidean distance for each point $\mathbf{x} \in C$
20:          **EndFor**
21:          Remove $(N - Q)$ points from C ▷ Removal of points having lowest $D(\mathbf{x}(i, j), \mathbf{x}(i_1, j_1))$
22:      **EndIf**
23:      $\hat{\mathbf{M}} =$ Convex spectral signatures from C
24: **Output : $\hat{\mathbf{M}}$** ▷ Output of the algorithm

---

that are very close to each other by taking into account the above concept.

To do this, every point in the ordered convex set C is assumed to have two neighbors: clockwise neighbor and anticlockwise neighbor. Let point $\mathbf{x}(i, j)$ have one neighbor $\mathbf{x}(i_1, j_1)$ in the anticlockwise direction. In our approach, we use anticlockwise direction (it does not make much difference whether we take clockwise or anticlockwise direction because similar types of points are close to each other). The second-order Minkowski's distance is the Euclidean distance (ED). The ED $D(\mathbf{x})$ for each point $\mathbf{x}(i, j)$ is defined as

$$D(\mathbf{x}(i, j), \mathbf{x}(i_1, j_1)) = \sqrt{(i - i_1)^2 + (j - j_1)^2}. \tag{14}$$

Our algorithm removes $(N - Q)$ points based on the smallest $D(\mathbf{x})$. The algorithm returns Q points from the original set of N points based on their ED. As a result, the algorithm returns a matrix $\hat{\mathbf{M}}$ of size $L \times Q$. $\hat{\mathbf{M}}$ contains Q pure endmembers extracted from the original image $\mathbf{Y}$. The algorithm output ($\hat{\mathbf{M}}$) is an extracted endmember matrix (second block of Fig. 1). A pseudocode of the proposed algorithm is given in Algorithm 1.

Fig. 2. Real datasets. (a) Urban. (b) Cuprite.

TABLE III
CUPRITE GT ENDMEMBERS

Notation	Endmember
C_1	Alunite
C_2	Andradite
C_3	Buddingtonite
C_4	Dumortierite
C_5	Kaolinite1
C_6	Kaolinite2
C_7	Muscovite
C_8	Montmorillonite
C_9	Nontronite
C_{10}	Pyrope
C_{11}	Sphene
C_{12}	Chalcedony

TABLE II
URBAN GT ENDMEMBERS

Notation	Endmember
U_1	Asphalt
U_2	Grass
U_3	Tree
U_4	Roof
U_5	Metal
U_6	Dirt

III. DATASETS AND PARAMETER SETTINGS

A. Datasets

The real and synthetic datasets used in our experiments are described in the following. Both types of datasets are in reflectance units.

1) Real Datasets: Two hyperspectral images (Urban and Cuprite) shown in Fig. 2 are used to validate the proposed algorithm. The Urban dataset was captured by the Hyperspectral Digital Imagery Collection Experiment sensor and is one of the most popular datasets in the spectral unmixing research community [54]. The Urban image has 307 lines, and each line has 307 pixels. The dataset contains a total of 210 bands, which cover the range from 400 to 2500 nm with a spectral resolution of 10 nm. We processed 162 bands after removing 48 noisy and water absorption bands (1–4, 76, 87, 101–111, 136–153, and 198–210). Each pixel is 2×2 m^2. As mentioned by Zhu [54], six ground-truth (GT) endmembers are selected from the United States Geological Survey (USGS) spectral library [55], as shown in Table II.

The Cuprite dataset is the benchmark dataset [54] for the hyperspectral unmixing research community. It was collected by the airborne visible infrared imaging spectrometer over the Cuprite mining district in Nevada, USA. The dataset covers the range from 370 to 2480 nm with 224 channels. There are six noisy bands (1, 2, and 221–224) and 30 water absorption bands (104–113 and 148–167). After removing these 36 bands, we process a total of 188 bands. A subscene with 250×190 pixels is taken to validate the proposed algorithm. We have assumed only

12 GT endmembers, as suggested in [54]. These endmember types, along with their notations used in this article, are shown in Table III.

2) Synthetic Datasets: We have used a set of synthetic hyperspectral images generated by the Hyperspectral Imagery Synthesis Toolbox [56]. All these synthetic images have been generated using five endmembers (asphalt—gds367, brick—gds350, fiberglass—gds374, sheetmetal—gds352, and vinylplastic—gds372) extracted from the USGS spectral library [55]. Each simulated image comprises 128×128 pixels and 431 spectral bands, ranging from 350 to 2500 nm (see Fig. 3). We have generated five images, Legendre (S1), Matern Gaussian (S2), Exponential Gaussian (S3), Rational Gaussian (S4), and Spheric Gaussian (S5), using the aforementioned tool. To evaluate the robustness of the proposed algorithm, each synthetic dataset is corrupted with additive Gaussian noise at different noise levels, to achieve signal-to-noise ratios (SNRs) of 20, 40, 60, 80, and 100 dB.

B. Comparison Parameters

In this section, each L-dimensional vector endmember ($\hat{\mathbf{m}} \equiv [\hat{m}_1, \hat{m}_2, \ldots \hat{m}_L]^T$) is compared to the corresponding GT endmember ($\mathbf{m} \equiv [m_1, m_2, \ldots m_L]^T$) using different metrics: spectral angle mapper (SAM), spectral information divergence (SID), Euclidean distance (ED), and normalized cross correlation (NXC). The root-mean-square error (RMSE) is also used to compare various algorithms in terms of abundance estimation accuracy. These metrics are summarized as follows.

1) Spectral Angle Mapper: The SAM between two spectral vectors $\mathbf{m}$ and $\hat{\mathbf{m}}$ of length L is defined as

$$\text{SAM} = \cos^{-1}\left(\frac{\mathbf{m} \cdot \hat{\mathbf{m}}}{|\mathbf{m}||\hat{\mathbf{m}}|}\right). \tag{15}$$

The total SAM (TSAM) for Q endmembers is given by

$$\text{TSAM} = \sum_{i=1}^{Q} \cos^{-1}\left(\frac{\mathbf{m}_i \cdot \hat{\mathbf{m}}_i}{|\mathbf{m}_i| \cdot |\hat{\mathbf{m}}_i|}\right). \tag{16}$$

2) Spectral Information Divergence: The SID [19] is an information-theoretic criterion for spectral discriminability
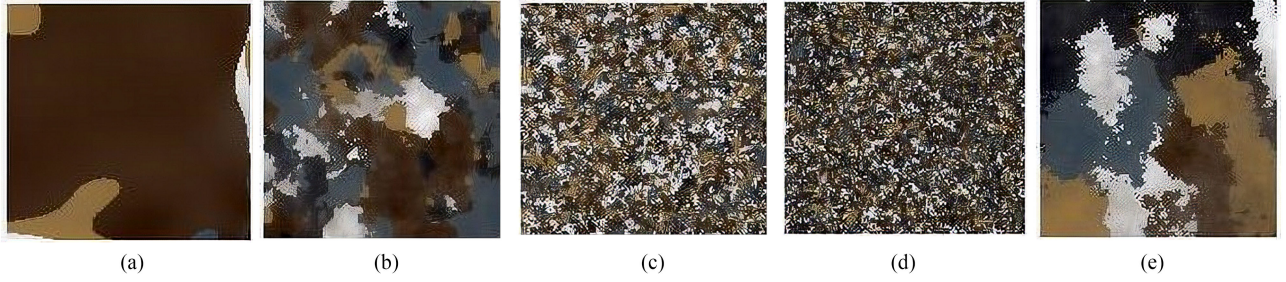
Fig. 3.   Synthetic datasets. (a) S1-Legendre. (b) S2-Matern. (c) S3-Exponential. (d) S4-Rational. (e) S5-Spheric.

and similarity. The probability vectors $\mathbf{p} = [p_1, p_2, \ldots p_L]^T$ and $\mathbf{q} = [q_1, q_2, \ldots q_L]^T$ (for two L-dimensional spectra $\mathbf{m} = [m_1, m_2, \ldots m_L]^T$ and $\hat{\mathbf{m}} = [\hat{m}_1, \hat{m}_2, \ldots \hat{m}_L]^T$) are, respectively, defined as

$$p_i = \frac{m_i}{\sum_{i=1}^{L} m_i}, \quad q_i = \frac{\hat{m}_i}{\sum_{i=1}^{L} \hat{m}_i}. \tag{17}$$

The SID between two spectra $\mathbf{m}$ and $\hat{\mathbf{m}}$ of length L is defined as

$$\text{SID} = \sum_{i=1}^{L} p_i \log\left(\frac{p_i}{q_i}\right) + \sum_{i=1}^{L} q_i \log\left(\frac{q_i}{p_i}\right). \tag{18}$$

The term $\sum_{i=1}^{L} p_i \log\left(\frac{p_i}{q_i}\right)$ is the Kullback–Leibler information function, which shows the relative entropy of $\mathbf{m}$ with respect to $\hat{\mathbf{m}}$.

The total SID (TSID) for Q endmembers is defined as

$$\text{TSID} = \sum_{j=1}^{Q} \left(\sum_{i=1}^{L} p_{ji} \log\left(\frac{p_{ji}}{q_{ji}}\right) + \sum_{i=1}^{L} q_{ji} \log\left(\frac{q_{ji}}{p_{ji}}\right) \right). \tag{19}$$

3) Euclidean Distance: The ED between two spectral signatures $\mathbf{m}$ and $\hat{\mathbf{m}}$ of length L is defined as

$$\text{ED} = \sqrt{\sum_{i=1}^{L} (m_i - \hat{m}_i)^2}. \tag{20}$$

The total Euclidean distance (TED) for Q endmembers is defined as

$$\text{TED} = \sum_{j=1}^{Q} \sqrt{\sum_{i=1}^{L} (m_{ij} - \hat{m}_{ij})^2}. \tag{21}$$

4) Normalized Cross Correlation: The NXC between two spectra $\mathbf{m}$ and $\hat{\mathbf{m}}$ of length L is defined as

$$\text{NXC} = \frac{\sqrt{\frac{\sum_{i=1}^{L}(m_i - \mu_m)(\hat{m}_i - \mu_{\hat{m}})}{\sigma_m \times \sigma_{\hat{m}}}}}{(L-1)}. \tag{22}$$

Here, μ_m and $\mu_{\hat{m}}$ are the mean of spectra $\mathbf{m}$ and $\hat{\mathbf{m}}$, respectively. σ_m and $\sigma_{\hat{m}}$ denote the standard deviation of spectra $\mathbf{m}$ and $\hat{\mathbf{m}}$, respectively.

Total normalized cross correlation (TNXC) for Q endmembers is defined as

$$\text{TNXC} = \frac{\sum_{j=1}^{Q} \sqrt{\frac{\sum_{i=1}^{L}(m_{ij} - \mu_{mj})(\hat{m}_{ij} - \mu_{\hat{m}j})}{\sigma_{mj} \times \sigma_{\hat{m}j}}}}{(L-1)}. \tag{23}$$

5) Root-Mean-Square Error: The RMSE is defined as

$$\text{RMSE} = \sum_{i=1}^{Q} \sqrt{\frac{\sum_{j=1}^{Z} (\mathbf{A}_j^i - \mathbf{G}_j^i)^2}{Z}}. \tag{24}$$

Here, $\mathbf{A}^i$ is the abundance map of the ith endmember, and $\mathbf{G}^i$ is an abundance map of the ith GT endmember. Z is the total number of pixels in the image. $\mathbf{A}_j^i$ is the jth pixel in the abundance map associated to the ith endmember. $\mathbf{G}_j^i$ is the jth pixel in the abundance map associated with the ith GT endmember.

IV. EXPERIMENTAL RESULTS

In this section, the results obtained by the proposed algorithm are compared to those obtained by ten well-known endmember extraction algorithms (SVMAX, AVMAX, VCA, TRIP, PPI, ICA, AMEE, SPEE, RBSPP, and SSPP) on both synthetic and real datasets. All these well-known algorithms are carefully optimized using empirical parameter settings of that respective algorithms. We empirically found the optimized values for both the real datasets in our experiments. We have used the optimized parameter values given in Table IV. In the hyperspectral unmixing chain, the number of endmembers Q is estimated before endmember extraction. Many subspace identification algorithms [53] exist in the literature. We have used the well-known HYperspectral Signal Identification be Minimum Error (HYSIME), [57] algorithm as a popular subspace identification method in the hyperspectral unmixing community. The proposed algorithm is implemented using MATLAB 2019b tool.

A. Experiments With Real Datasets

Tables V–VIII and Tables IX–XII report our experimental results on the Urban and Cuprite datasets, respectively. The compared algorithms are depicted in the first column of Tables V–VIII and Tables IX–XII. SAM, SID, ED, and NXC values are computed using (15), (18), (20), and (22), respectively. The second column in Tables V–VIII shows a comparison (using different metrics) between the first GT signature (U_1) and the

TABLE IV
PARAMETER SETTINGS FOR EACH SCENE

Algorithm	Parameter	Cuprite	Urban
TRIP	P^{th} norm	2	2
PPI	Number of skewers	1500	600
AMEE	Maximum kernel size	15	15
	Maximum iterations	100	100
SPEE	Window size	5	7
	Spectral algorithm	VCA	VCA
RBSPP	Clustering algorithm	ISODATA	ISODATA
	Spectral algorithm	VCA	VCA
SSPP	α	70	30
	β	60	50
	σ	1.6	2.2
	Spectral algorithm	VCA	VCA

TABLE V
SAM FOR URBAN REAL DATASET

Algorithm	U_1	U_2	U_3	U_4	U_5	U_6	TSAM	Mean	Std
SVMAX	0.1	0.123	0.095	0.512	0.116	0.084	1.029	0.171	0.167
AVMAX	0.1	0.097	0.095	0.669	0.116	0.084	1.16	0.193	0.233
VCA	0.1	0.097	0.095	0.08	0.554	0.084	1.009	0.168	0.189
TRIP	0.1	0.123	0.095	0.194	0.121	0.084	0.717	0.119	0.04
PPI	0.1	0.123	0.095	0.403	0.121	0.084	0.926	0.154	0.123
ICA	0.63	0.077	0.109	0.744	0.446	0.088	2.095	0.349	0.298
AMEE	0.482	0.131	0.187	0.060	0.201	0.195	1.255	0.209	0.144
SPEE	0.232	0.138	0.151	0.273	0.123	0.112	1.029	0.171	0.065
RBSPP	0.168	0.071	0.437	0.387	0.130	0.093	1.287	0.214	0.158
SSPP	0.200	0.069	0.174	0.205	0.324	0.128	1.101	0.183	0.086
Proposed	0.100	0.084	0.095	0.065	0.149	0.116	**0.609**	**0.101**	**0.029**

TABLE VI
SID FOR URBAN REAL DATASET

Algorithm	U_1	U_2	U_3	U_4	U_5	U_6	TSID	Mean	Std
SVMAX	0.011	0.018	0.02	0.314	0.014	0.007	0.384	0.064	0.122
AVMAX	0.011	0.013	0.02	0.603	0.014	0.007	0.669	0.111	0.241
VCA	0.011	0.013	0.02	0.008	0.385	0.007	0.445	0.074	0.152
TRIP	0.011	0.018	0.02	0.042	0.014	0.007	0.113	0.019	0.012
PPI	0.011	0.018	0.02	0.182	0.014	0.007	0.253	0.042	0.069
ICA	0.5	0.008	0.03	0.845	0.225	0.008	1.616	0.269	0.341
AMEE	0.257	0.030	0.240	0.004	0.083	0.042	0.655	0.109	0.111
SPEE	0.056	0.027	0.268	0.084	0.018	0.013	0.467	0.078	0.097
RBSPP	0.443	0.007	0.056	0.041	0.138	0.021	0.705	0.118	0.166
SSPP	0.028	0.006	0.304	0.164	0.022	0.011	0.536	0.089	0.121
Proposed	0.011	0.011	0.020	0.005	0.022	0.013	**0.081**	**0.014**	**0.006**

TABLE VII
ED FOR URBAN REAL DATASET

Algorithm	U_1	U_2	U_3	U_4	U_5	U_6	TED	Mean	Std
SVMAX	0.771	0.397	0.981	2.192	1.08	1.024	6.445	1.074	0.602
AVMAX	0.771	1.448	0.981	1.811	1.08	1.024	7.115	1.186	0.377
VCA	0.771	1.448	0.981	0.336	1.45	1.024	6.01	1.002	0.424
TRIP	0.771	0.397	0.981	1.937	3.572	1.024	8.683	1.447	1.159
PPI	0.771	0.397	0.981	2.358	1.572	1.024	7.104	1.184	0.691
ICA	2.273	0.252	0.288	2.458	1.186	2.157	8.615	1.436	1.004
AMEE	1.698	1.157	0.992	0.891	1.100	1.612	7.449	1.242	**0.334**
SPEE	1.662	0.444	1.816	2.229	0.375	0.783	7.310	1.218	0.784
RBSPP	0.810	0.169	1.376	1.782	0.737	3.610	8.483	1.414	1.211
SSPP	1.579	0.213	0.524	1.302	0.983	2.962	7.562	1.260	0.971
Proposed	0.771	0.353	0.981	0.764	2.092	0.959	**5.921**	**0.987**	0.587

TABLE VIII
NXC FOR URBAN REAL DATASET

Algorithm	U_1	U_2	U_3	U_4	U_5	U_6	TNXC	Mean	Std
SVMAX	0.876	0.98	0.995	-0.277	0.845	0.937	4.355	0.726	0.495
AVMAX	0.876	0.988	0.995	-0.025	0.845	0.937	4.615	0.769	0.393
VCA	0.876	0.988	0.995	0.976	0.462	0.937	5.233	0.872	0.206
TRIP	0.876	0.98	0.995	0.632	0.85	0.937	5.269	0.878	0.133
PPI	0.876	0.98	0.995	-0.422	0.85	0.937	4.215	0.703	0.554
ICA	0.122	0.991	0.989	-0.048	0.523	0.929	3.505	0.584	0.461
AMEE	-0.405	0.977	0.886	0.980	0.859	0.864	4.162	0.694	0.541
SPEE	0.618	0.980	0.772	0.041	0.834	0.882	4.127	0.688	0.339
RBSPP	0.673	0.991	0.973	-0.410	0.835	0.940	4.002	0.667	0.540
SSPP	0.122	0.992	0.982	0.777	0.657	0.893	4.423	0.737	0.327
Proposed	0.876	0.991	0.995	0.984	0.785	0.878	**5.509**	**0.918**	**0.086**

algorithms) from the Urban dataset. Similarly, in Tables IX–XII, C_1–C_{12} columns report a comparison (using different metrics) between 12 GT spectra in the Cuprite scene and the corresponding endmembers extracted by different algorithms. The TSAM, TSID, TED, and TNXC values, calculated using (16), (19), (21), and (23), are reported in Tables V–XII. The mean and standard deviation of all evaluation metrics (SAM, SID, ED, and NXC) are reported in the last two columns of Tables V–XII. The bold entries in the Tables V-XIV represent the best performing algorithm. All values of SAM, SID, TSAM, and TSID are in radians. The results obtained using the considered metrics are discussed in the following.

1) Spectral Angle Mapper: It can be observed from Tables V and IX that the SAM values obtained for the proposed algorithm are close to zero, which indicates that the extracted endmembers are very similar to the corresponding GT spectra for the Urban and Cuprite datasets. The SAM values of a few algorithms are close to $\pi/2$, which indicates that the extracted endmembers are different than the GT ones. It can be seen that the TSAM values of the proposed algorithm are consistently lower than those obtained by other algorithms. In addition, the mean and the standard deviation values obtained for the proposed algorithm are lower than those obtained for the other algorithms in the two considered real datasets.

2) Spectral Information Divergence: As shown in Tables VI and X, the SID values obtained for the proposed algorithm on the Urban and Cuprite datasets are close to zero, which again indicates that the extracted endmembers are very similar to the corresponding GT ones. The TSID values in Tables VI and X also reveal that the values obtained for the proposed algorithm are low when compared to those of other algorithms. As shown in the last two columns of Tables VI and X, the mean and standard deviation values obtained for the proposed algorithm are lower than those obtained for the other algorithms.

3) Euclidean Distance: The ED values of Tables VII and XI indicate that the values obtained for the proposed algorithm are low when compared to those of other algorithms. The mean values obtained for the proposed algorithm are lower than those obtained for the other algorithms, and the standard deviation values are comparable.

4) Normalized Cross Correlation: The NXC values computed for the proposed algorithm on both real datasets are positive for all endmembers, as shown in Tables VIII and XII. It can be seen that some values are negative in both tables. It can also be noted that negative values are coming mostly for those

corresponding extracted endmember (i.e., the one that is most similar to U_1 among the set of extracted endmembers, according to each metric). The remaining columns in Tables V–VIII report a comparison between the remaining GT spectra (U_2–U_5) and their corresponding endmembers (extracted by different

TABLE IX
SAM FOR CUPRITE REAL DATASET

Algorithm	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	TSAM	Mean	Std
SVMAX	0.597	0.06	0.201	0.108	0.085	0.065	0.151	0.056	0.12	0.09	0.137	0.258	1.927	0.161	0.15
AVMAX	0.674	0.054	0.446	0.104	0.098	0.064	0.142	0.054	0.105	0.083	0.149	0.149	2.121	0.177	0.188
VCA	0.546	0.06	0.335	0.108	0.106	0.053	0.133	0.054	0.116	0.169	0.137	0.465	2.281	0.19	0.166
TRIP	0.338	0.057	0.158	0.108	0.098	0.065	0.085	0.052	0.115	0.125	0.137	0.151	1.489	0.124	0.076
PPI	0.374	0.064	0.163	0.108	0.109	0.053	0.115	0.05	0.106	0.145	0.084	0.089	1.46	0.122	0.086
ICA	0.474	0.048	0.185	0.108	0.124	0.063	0.125	0.056	0.112	0.092	0.102	0.076	1.564	0.13	0.114
AMEE	0.517	0.062	0.126	0.135	0.114	0.069	0.113	0.047	0.105	0.123	0.184	0.145	1.739	0.145	0.123
SPEE	0.122	0.060	0.446	0.103	0.098	0.074	0.130	0.053	0.112	0.122	0.160	0.465	1.946	0.162	0.140
RBSPP	0.133	0.011	0.192	0.142	0.120	0.051	0.170	0.041	0.419	0.183	0.101	0.428	1.993	0.166	0.133
SSPP	0.090	0.065	0.313	0.568	0.147	0.081	0.479	0.060	0.118	0.099	0.081	0.474	2.575	0.215	0.190
Proposed	0.251	0.064	0.136	0.114	0.108	0.053	0.108	0.050	0.106	0.108	0.152	0.151	**1.401**	**0.117**	**0.054**

TABLE X
SID FOR CUPRITE REAL DATASET

Algorithm	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	TSID	Mean	Std
SVMAX	0.336	0.007	0.054	0.021	0.019	0.008	0.031	0.007	0.021	0.012	0.026	0.09	0.632	0.053	0.092
AVMAX	0.396	0.007	0.182	0.019	0.018	0.009	0.028	0.007	0.018	0.011	0.031	0.035	0.761	0.063	0.115
VCA	0.306	0.007	0.129	0.021	0.016	0.007	0.026	0.007	0.02	0.042	0.026	0.122	0.73	0.061	0.088
TRIP	0.147	0.007	0.033	0.021	0.018	0.008	0.013	0.006	0.018	0.024	0.026	0.034	0.356	0.03	0.038
PPI	0.171	0.008	0.037	0.02	0.018	0.007	0.022	0.007	0.017	0.032	0.016	0.017	0.37	0.031	0.045
ICA	0.274	0.006	0.047	0.02	0.023	0.009	0.023	0.007	0.018	0.014	0.017	0.014	0.472	0.039	0.075
AMEE	0.242	0.007	0.023	0.031	0.020	0.010	0.020	0.006	0.017	0.022	0.045	0.033	0.476	0.040	0.065
SPEE	0.027	0.007	0.182	0.020	0.018	0.010	0.025	0.007	0.017	0.022	0.035	0.122	0.494	0.041	0.054
RBSPP	0.017	0.011	0.114	0.022	0.032	0.020	0.021	0.009	0.020	0.018	0.016	0.190	0.489	0.041	0.055
SSPP	0.192	0.020	0.271	0.060	0.012	0.024	0.047	0.013	0.070	0.042	0.020	0.094	0.864	0.072	0.080
Proposed	0.088	0.008	0.025	0.022	0.017	0.007	0.018	0.007	0.017	0.017	0.029	0.034	**0.290**	**0.024**	**0.022**

TABLE XI
ED FOR CUPRITE REAL DATASET

Algorithm	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	TED	Mean	Std
SVMAX	8.481	7.065	4.06	4.99	1.792	4.236	5.512	4.286	1.772	4.758	1.412	5.23	53.594	4.466	2.108
AVMAX	8.952	6.178	5.108	4.278	2.447	3.803	5.706	4.024	2.529	4.654	1.476	4.144	53.3	4.442	1.967
VCA	7.847	6.848	5.921	4.99	2.406	3.122	5.676	4.024	2.057	4.475	1.412	6.468	55.246	4.604	2.048
TRIP	7.364	6.587	3.743	4.99	2.447	4.236	4.769	4.321	1.777	5.186	1.412	4.502	51.334	4.278	**1.773**
PPI	7.775	7.057	3.968	4.874	2.51	3.122	4.959	4.734	2.177	6.704	1.221	2.97	52.068	4.339	2.064
ICA	7.676	6.016	4.616	5.084	1.977	4.123	5.531	4.286	1.083	4.845	0.942	3.399	50.579	4.215	2.009
AMEE	8.195	6.834	4.103	4.689	3.247	3.704	1.775	4.360	2.218	5.233	3.466	4.296	52.120	4.343	1.797
SPEE	4.919	7.065	5.108	4.494	2.447	2.794	5.447	4.010	1.694	5.379	1.733	6.468	51.558	4.297	1.785
RBSPP	4.979	7.085	5.874	4.543	2.965	2.578	5.976	4.956	1.075	4.976	1.708	6.371	53.088	4.424	1.917
SSPP	5.298	8.365	5.560	5.609	3.316	3.168	5.005	5.896	1.518	6.842	0.936	6.250	57.762	4.814	2.180
Proposed	6.208	7.057	3.426	5.118	2.182	3.122	5.273	4.734	2.177	5.597	0.722	4.502	**50.117**	**4.177**	1.872

TABLE XII
NXC FOR CUPRITE REAL DATASET

Algorithm	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	TNXC	Mean	Std
SVMAX	-0.475	0.944	0.595	0.902	0.962	0.947	0.86	0.95	0.9	0.922	0.864	0.391	8.761	0.73	0.417
AVMAX	-0.401	0.968	-0.061	0.897	0.944	0.949	0.578	0.955	0.908	0.934	0.835	0.866	8.374	0.698	0.452
VCA	-0.578	0.945	0.336	0.902	0.929	0.967	0.738	0.955	0.892	0.916	0.864	-0.151	7.716	0.643	0.509
TRIP	0.054	0.949	0.731	0.902	0.944	0.947	0.852	0.958	0.89	0.847	0.864	0.836	9.774	0.815	0.248
PPI	-0.27	0.937	0.746	0.887	0.936	0.967	0.774	0.96	0.914	0.792	0.95	0.932	9.526	0.794	0.344
ICA	-0.625	0.965	0.739	0.89	0.918	0.951	0.681	0.95	0.918	0.918	0.924	0.887	9.115	0.76	0.445
AMEE	-0.324	0.942	0.741	0.827	0.923	0.939	0.653	0.965	0.918	0.850	0.760	0.678	8.872	0.739	0.352
SPEE	0.900	0.944	-0.061	0.917	0.944	0.935	0.771	0.958	0.898	0.847	0.812	-0.151	8.714	0.726	0.393
RBSPP	-0.239	0.984	0.245	0.934	0.991	0.901	0.698	0.905	0.864	0.950	0.927	-0.017	8.143	0.679	0.431
SSPP	0.910	0.946	0.232	0.077	0.868	0.941	-0.298	0.943	0.895	0.903	0.961	0.604	7.983	0.665	0.426
Proposed	0.304	0.937	0.713	0.889	0.940	0.967	0.758	0.960	0.914	0.891	0.825	0.836	**9.933**	**0.828**	**0.183**

TABLE XIII
OVERALL COMPARISON ON URBAN DATASET

Algorithm	SAM			SID			ED			NXC			RMSE	Running Time
	Total	Mean	std	Total	Mean	std	Total	Mean	std	Total	Mean	std		
SVMAX	1.029	0.171	0.167	0.384	0.064	0.122	6.445	1.074	0.602	4.355	0.726	0.495	2.571	11.393
AVMAX	1.160	0.193	0.233	0.669	0.111	0.241	7.115	1.186	0.377	4.615	0.769	0.393	2.453	**10.325**
VCA	1.009	0.168	0.189	0.445	0.074	0.152	6.01	1.002	0.42	5.233	0.872	0.206	2.646	11.141
TRIP	0.717	0.119	0.040	0.113	0.019	0.012	8.683	1.447	1.16	5.269	0.878	0.133	2.472	14.102
PPI	0.926	0.154	0.123	1.203	0.401	0.455	7.104	1.184	0.691	4.215	0.703	0.554	2.601	16.672
ICA	2.095	0.349	0.298	1.616	0.269	0.341	8.615	1.436	1	3.505	0.584	0.461	2.271	22.888
AMEE	1.255	0.209	0.144	1.608	0.536	0.624	7.449	1.242	**0.334**	4.162	0.694	0.541	2.254	66.740
SPEE	1.029	0.171	0.065	1.265	0.422	0.528	7.310	1.218	0.784	4.127	0.688	0.339	2.171	188.033
RBSPP	1.287	0.214	0.158	1.659	0.553	0.636	8.483	1.414	1.211	4.002	0.667	0.540	2.634	82.328
SSPP	1.101	0.183	0.086	1.370	0.457	0.560	7.562	1.260	0.971	4.423	0.737	0.327	2.388	77.211
Proposed	**0.609**	**0.101**	**0.029**	**0.081**	**0.014**	**0.006**	**5.921**	**0.987**	0.587	**5.509**	**0.918**	**0.086**	**2.106**	70.183

TABLE XIV
OVERALL COMPARISON ON CUPRITE DATASET

Algorithm	SAM			SID			ED			NXC			RMSE	Running Time
	Total	Mean	std	Total	Mean	std	Total	Mean	std	Total	Mean	std		
SVMAX	1.927	0.161	0.150	0.632	0.053	0.092	53.594	4.466	2.108	8.761	0.73	0.417	3.357	13.233
AVMAX	2.121	0.177	0.188	0.761	0.063	0.115	53.300	4.442	1.967	8.374	0.698	0.452	3.631	**11.483**
VCA	2.281	0.190	0.166	0.730	0.061	0.088	55.246	4.604	2.048	7.716	0.643	0.509	3.549	11.560
TRIP	1.489	0.124	0.076	0.356	0.030	0.038	51.334	4.278	**1.773**	9.774	0.815	0.248	3.527	13.546
PPI	1.460	0.122	0.086	0.370	0.031	0.045	52.068	4.339	2.064	9.526	0.794	0.344	3.801	30.686
ICA	1.564	0.130	0.114	0.472	0.039	0.075	50.579	4.215	2.009	9.115	0.76	0.445	3.503	28.907
AMEE	1.739	0.145	0.123	0.476	0.040	0.065	52.120	4.343	1.797	8.872	0.739	0.352	3.402	91.546
SPEE	1.946	0.162	0.140	0.494	0.041	0.054	51.558	4.297	1.785	8.714	0.726	0.393	3.411	191.297
RBSPP	1.993	0.166	0.133	0.489	0.041	0.055	53.088	4.424	1.917	8.143	0.679	0.431	3.646	123.048
SSPP	2.575	0.215	0.190	0.864	0.072	0.080	57.762	4.814	2.180	7.983	0.665	0.426	3.383	107.650
Proposed	**1.401**	**0.117**	**0.054**	**0.290**	**0.024**	**0.022**	**50.117**	**4.177**	1.872	**9.933**	**0.828**	**0.183**	**3.185**	105.520

materials, which are very rare in a scene. For the Urban scene, materials (U_2, U_3, U_5, and U_6) cover more area in comparison with materials (U_1 and U_4). Similarly, materials C_1, C_3, C_7, and C_{12} correspond to rare minerals in the Cuprite scene. Many algorithms cannot extract these rare materials. Using the heterogeneous features from high-entropy bands (and homogeneous features from low-entropy bands) in the second step of the proposed algorithm leads to high NXC values for these rare endmembers, as opposed to other methods. The TNXC values reveal that the TNXC values obtained for the proposed algorithm are high compared to those obtained for other algorithms. As the last two columns of Tables VIII and XII show, the mean values obtained for the proposed algorithm are higher than those obtained for the other algorithms, and the standard deviation values are lower.

5) Root Mean Square Error: The third stage in the unsupervised spectral unmixing chain is abundance estimation, as described in Fig. 1. There are mainly three types of abundance estimation methods: fully constrained, nonnegatively constrained, and unconstrained. If a method uses the constraints in (2) and (3) for finding an abundance vector α, a method is called fully constrained. If a method uses only the constraint mentioned in (2) for finding an abundance vector α, it is called nonnegatively constrained. If a method does not use any constraint for finding an abundance vector α, it is called unconstrained.

In our experiment, we have used the fully constrained least square (FCLS) [58] method to estimate fractional abundances and compare various algorithms, as it follows the ANC and ASC constraints of the linear mixing model described in (1). The RMSE, as described in (24), is used to establish a comparison between the abundance maps obtained by various algorithms with those obtained by the proposed algorithm. The RMSE values obtained for all tested algorithms with the Urban and Cuprite real datasets are shown in Tables XIII and XIV, respectively. As shown in the figure, the proposed algorithm outperforms all other algorithms in terms of abundance estimation.

6) Running Time: The experiments are performed on a computer with Intel i7-7500 CPU at 2.7 GHz, 8-GB RAM, and windows 10 (64-bit) operating system. The running times of the Urban and Cuprite real datasets for all algorithms are shown in Tables XIII and XIV, respectively. It can be observed that the running time of the proposed algorithm is competent with other algorithms.

7) Others: The abundance maps for the Urban and Cuprite dataset are shown in Figs. 4 and 5, respectively. The comparison of GT and extracted endmembers by the proposed algorithm for the Urban and Cuprite datasets are shown in Figs. 6 and 7, respectively. In these figures, brown and blue color curves represent GT and extracted endmembers, respectively. In these graphs, the Y-axis represents the reflectance value and the X-axis represents
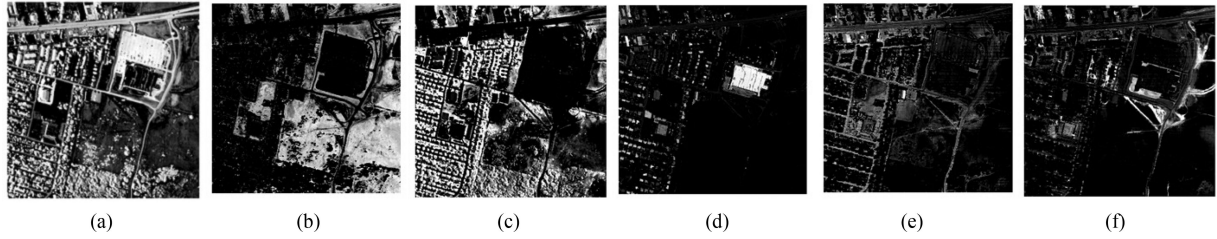
Fig. 4.   Abundance maps estimated by FCLS from the Urban dataset. (a) Ashphalt road. (b) Grass. (c) Tree. (d) Roof. (e) Metal. (f) Dirt.
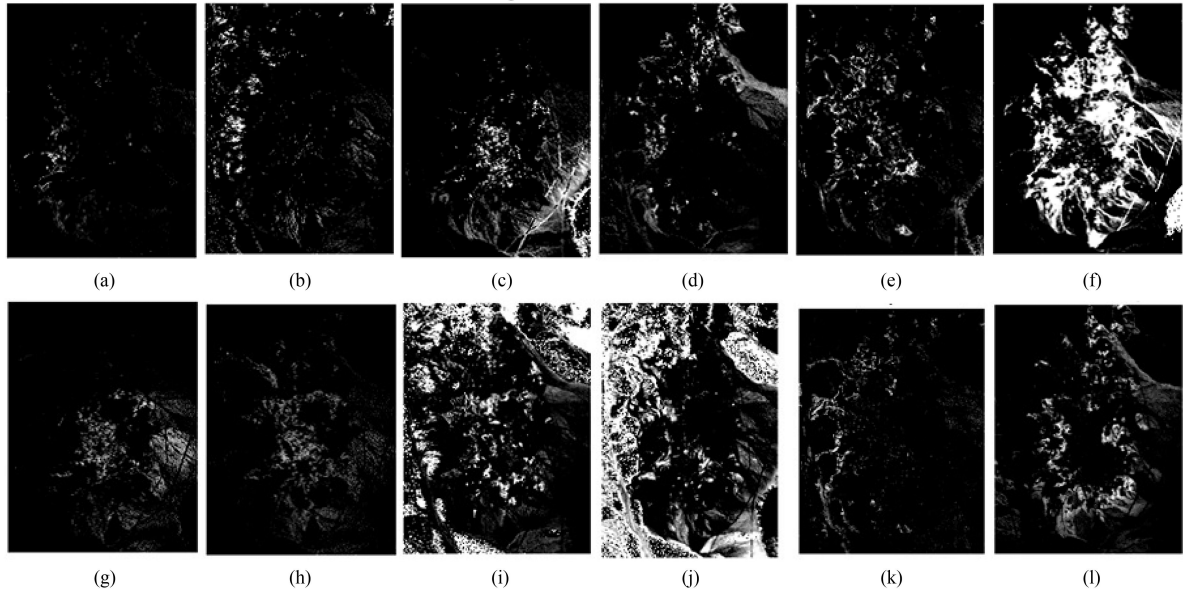

Fig. 5.   Abundance maps estimated by FCLS from the Cuprite dataset. (a) Alunite. (b) Andradite. (c) Buddingtonite. (d) Dumortierite. (e) Kaolinite1. (f) Kaolinite2. (g) Muscovite. (h) Montmorillonite. (i) Nontronite. (j) Pyrope. (k) Sphene. (l) Chalcedony.
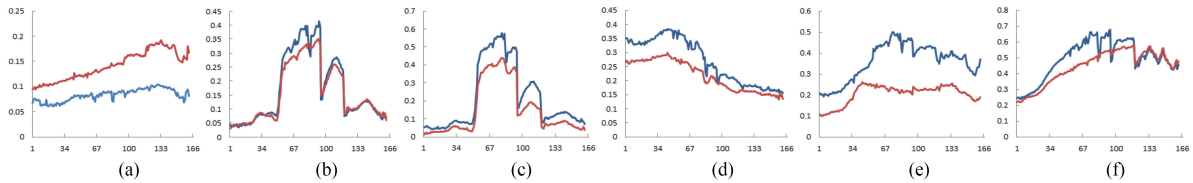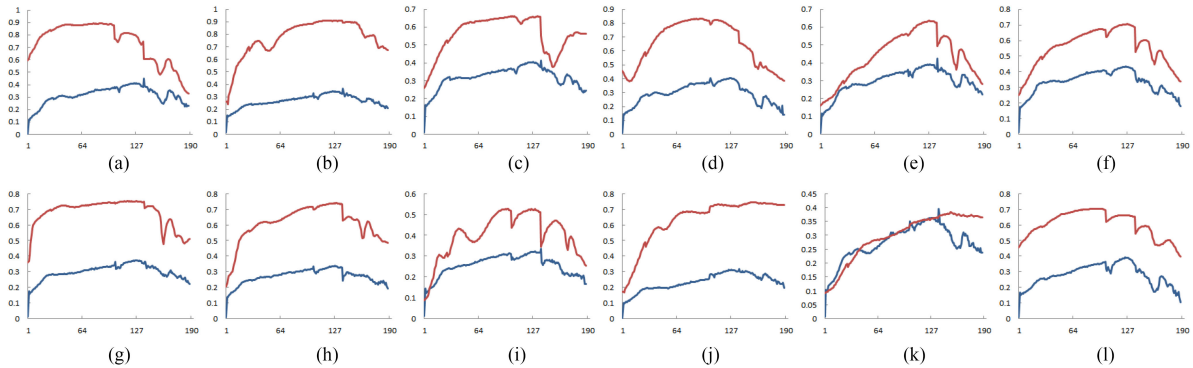

Fig. 6.   Extracted endmembers from the Urban dataset. (a) Ashphalt road. (b) Grass. (c) Tree. (d) Roof. (e) Metal. (f) Dirt.

Fig. 7.   Extracted endmembers from the Cuprite dataset. (a) Alunite. (b) Andradite. (c) Buddingtonite. (d) Dumortierite. (e) Kaolinite1. (f) Kaolinite2. (g) Muscovite. (h) Montmorillonite. (i) Nontronite. (j) Pyrope. (k) Sphene. (l) Chalcedony.

TABLE XV
EFFECT OF NOISE ON SYNTHETIC DATASETS FOR ALL ALGORITHMS

Image	SNR	SVMAX	AVMAX	VCA	TRIP	PPI	ICA	AMEE	SPEE	RBSPP	SSPP	Proposed
S1	20	2.116	1.882	1.936	2.129	2.144	2.127	**1.346**	1.482	1.917	1.906	1.927
	40	2.164	2.164	2.162	2.164	2.409	2.533	**1.414**	2.214	2.012	2.025	1.523
	60	2.450	2.450	2.452	2.450	2.420	2.452	1.973	2.217	1.962	2.035	**1.950**
	80	2.451	2.451	2.451	2.451	2.532	2.434	1.974	2.218	1.964	**1.918**	1.954
	100	2.451	2.451	2.451	2.451	2.390	2.463	1.999	2.218	2.006	**1.944**	2.051
	Mean	2.326	2.280	2.290	2.329	2.379	2.402	**1.741**	2.070	1.972	1.965	1.881
S2	20	1.712	1.772	1.708	1.540	1.520	1.768	1.367	**0.823**	1.431	1.542	0.976
	40	1.848	1.827	1.827	1.848	1.898	1.848	1.827	**1.413**	1.529	1.536	**1.413**
	60	1.859	1.859	1.859	1.859	1.844	1.877	1.859	**1.451**	1.600	1.772	**1.451**
	80	1.859	1.859	1.859	1.859	1.859	1.943	1.859	1.859	**1.600**	1.622	1.859
	100	1.859	1.859	1.859	1.859	1.798	1.943	1.859	1.859	**1.600**	1.801	1.859
	Mean	1.828	1.835	1.822	1.793	1.784	1.876	1.754	**1.481**	1.552	1.654	1.512
S3	20	0.968	1.043	0.968	**0.886**	0.999	1.087	1.372	0.932	1.130	1.338	0.925
	40	1.079	1.079	1.149	1.079	1.438	0.993	1.145	1.522	1.273	1.549	**0.823**
	60	1.075	1.075	1.075	1.075	0.954	0.931	1.394	1.336	1.258	1.260	**0.824**
	80	**1.075**	**1.075**	**1.075**	**1.075**	1.375	1.375	1.394	1.335	1.258	1.380	**1.075**
	100	**1.075**	**1.075**	**1.075**	**1.075**	1.375	1.395	1.394	1.335	1.258	1.301	**1.075**
	Mean	1.055	1.070	1.069	1.038	1.228	1.156	1.340	1.292	1.236	1.365	**0.945**
S4	20	1.396	1.467	1.507	1.458	1.470	1.863	**1.310**	1.626	1.406	1.437	1.410
	40	**1.552**	**1.552**	1.601	**1.552**	1.602	1.842	1.606	1.601	1.695	1.698	**1.552**
	60	1.271	1.271	1.271	1.271	1.302	1.957	1.549	1.550	1.692	1.261	**1.258**
	80	1.271	1.271	1.271	1.271	1.271	1.371	1.549	1.269	1.691	**1.251**	1.271
	100	1.271	1.271	1.271	1.271	1.293	1.371	1.461	1.269	1.691	**1.232**	1.271
	Mean	**1.352**	1.366	1.384	1.365	1.388	1.681	1.495	1.463	1.635	1.376	**1.352**
S5	20	1.301	1.332	1.358	1.301	1.317	**1.136**	1.231	1.317	1.484	1.609	**1.136**
	40	1.411	1.411	1.411	1.411	1.199	1.253	1.240	1.515	1.283	**1.166**	1.253
	60	**1.087**	**1.087**	**1.087**	**1.087**	1.097	1.187	1.237	2.144	1.258	1.152	**1.087**
	80	**1.087**	**1.087**	**1.087**	**1.087**	1.088	1.088	1.237	2.144	1.258	1.187	**1.087**
	100	**1.087**	**1.087**	**1.087**	**1.087**	1.097	1.097	1.237	2.144	1.258	1.171	**1.087**
	Mean	1.195	1.201	1.206	1.195	1.159	1.152	1.236	1.853	1.308	1.257	**1.130**

the band number. The overall comparison of experiments with Urban and Cuprite dataset is shown in Tables XIII and XIV, respectively.

B. Experiments With Synthetic Datasets

The TSAM value obtained for the synthetic images by all the considered algorithms is shown in Table XV. Each synthetic image has Gaussian noise of SNR 20, 40, 60, 80, and 100 dB. As a result, each of the five types of synthetic images contains five types of noise (i.e., we have a total of 25 synthetic images). For each variant of a synthetic image, TSAM is calculated. It can be seen in the last column of Table XV that the proposed algorithm gives the lowest TSAM values for 13 out of the 25 generated images. The bold entries in the Table XV represents the best performing algorithm in that particular row. It can also be seen in the table that the proposed algorithm gives better performance than other algorithms for Exponential, Rational, and Spheric types of synthetic images. Moreover, the proposed algorithm gives almost similar TSAM values for Legendre and Matern synthetic images. The mean values of TSAM are also calculated

in this experiment to showcase the average effect of all noise types on each particular image. The mean value of TSAM for the proposed algorithm is lower for images S3, S4, and S5, while the lowest mean values for S1 and S2 are obtained by AMEE and SPEE, respectively. It can also be noted that the proposed algorithm is close to the best performer for S1 and S2 synthetic images.

V. CONCLUSION

This article introduces a new algorithm for endmember extraction from hyperspectral images. An innovative characteristic of the proposed method is that it combines the concept of band entropy (to model spatial information) and convex geometry (to characterize spectral information). The proposed framework for the integration of spatial and spectral information is innovative and offers the advantage that rare endmembers are not discarded by our method (as opposed to other traditional algorithms for endmember extraction, particularly those that focus on extracting endmembers from spatially homogeneous areas in the scene). The proposed algorithm is evaluated by a detailed comparison with other methods, using both synthetic and real

datasets, and comparative metrics such as SAM, SID, ED, and NXC. Our results indicate that the performance of the proposed algorithm is generally better than that of state-of-the-art algorithms, not only in terms of endmember extraction accuracy but also in terms of abundance estimation accuracy (evaluated using the RMSE). The efficacy of the proposed algorithm is also demonstrated with noise-corrupted synthetic scenes, in which ground endmembers and abundances are known. The proposed approach has the specific advantage of being able to find rare endmember signatures from the scene, which is particularly helpful in target detection and accurate material mapping.

ACKNOWLEDGMENT

REFERENCES

[1] A. J. Brown *et al.*, "Hydrothermal formation of clay-carbonate alteration assemblages in the Nili Fossae region of Mars," *Earth Planet. Sci. Lett.*, vol. 297, nos. 1/2, pp. 174–182, 2010.

[2] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sen.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[3] J. S. Bhatt, "Novel approaches for spectral unmixing of hyperspectral data," Ph.D. dissertation, Dept. Inf. Commun. Technol., Dhirubhai Ambani Inst, Inf. Commun. Technol., Gandhinagar, India, 2014.

[4] R. Li and S. Latifi, "Improving hyperspectral subpixel target detection using hybrid detection space," *J. Appl. Remote Sens.*, vol. 12, no. 1, 2018, Art. no. 015022.

[5] Y. Zhang, X. Huang, X. Hao, J. Wang, W. Wang, and T. Liang, "Fractional snow-cover mapping using an improved endmember extraction algorithm," *J. Appl. Remote Sens.*, vol. 8, no. 1, 2014, Art. no. 084691.

[6] F. Huang, Y. Yu, and T. Feng, "Hyperspectral remote sensing image change detection based on tensor and deep learning," *J. Vis. Commun. Image Representation*, vol. 58, pp. 233–244, 2019.

[7] B. Tu, N. Li, L. Fang, H. Fei, and D. He, "Classification of hyperspectral images via weighted spatial correlation representation," *J. Vis. Commun. Image Representation*, vol. 56, pp. 160–166, 2018.

[8] J.-C. Wu and G.-c. Tsuei, "Comparison of hyperspectral endmember extraction algorithms," *J. Appl. Remote Sens.*, vol. 7, no. 1, 2013, Art. no. 073525.

[9] X. Geng, L. Ji, F. Wang, Y. Zhao, and P. Gong, "Statistical volume analysis: A new endmember extraction method for multi/hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6100–6109, Oct. 2016.

[10] X. R. Wang, A. J. Brown, and B. Upcroft, "Applying incremental EM to Bayesian classifiers in the learning of hyperspectral remote sensing data," in *Proc. 7th Int. Conf. Inf. Fusion*, vol. 1, 2005, p 8.

[11] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[12] S. Ozkan, B. Kaya, and G. B. Akar, "EndNet: Sparse autoencoder network for endmember extraction and hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 482–496, Jan. 2019.

[13] Y. Su, J. Li, A. Plaza, A. Marinoni, P. Gamba, and S. Chakravortty, "DAEN: Deep autoencoder networks for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4309–4321, Jul. 2019.

[14] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, "End-member extraction for hyperspectral image analysis," *Appl. Opt.*, vol. 47, no. 28, pp. F77–F84, 2008.

[15] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proc. SPIE*, vol. 3753, pp. 266–276, 1999.

[16] W. Xiong, C. Chang, C. Wu, K. Kalpakis, and H. M. Chen, "Fast algorithms to implement N-FINDR for hyperspectral endmember extraction," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 3, pp. 545–564, Sep. 2011.

[17] C.-I. Chang, C.-C. Wu, W. Liu, and Y. Ouyang, "A new growing method for simplex-based endmember extraction algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2804–2819, Oct. 2006.

[18] T. Chan, W. Ma, A. Ambikapathi, and C. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4177–4193, Nov. 2011.

[19] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[20] A. Ambikapathi, T. Chan, C. Chi, and K. Keizer, "Hyperspectral data geometry-based estimation of number of endmembers using p-norm-based pure pixel identification algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2753–2769, May 2013.

[21] C.-I. Chang and A. Plaza, "A fast iterative algorithm for implementation of pixel purity index," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 63–67, Jan. 2006.

[22] J. D. Bayliss, J. A. Gualtieri, and R. F. Cromp, "Analyzing hyperspectral data with independent component analysis," *Proc. SPIE*, vol. 3240, pp. 133–144, 1998.

[23] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 9, pp. 2025–2041, Sep. 2002.

[24] M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2679–2693, Aug. 2009.

[25] G. Martin and A. Plaza, "Region-based spatial preprocessing for endmember extraction and spectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 745–749, Jul. 2011.

[26] G. Martin and A. Plaza, "Spatial-spectral preprocessing prior to endmember identification and unmixing of remotely sensed hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sen.*, vol. 5, no. 2, pp. 380–395, Apr. 2012.

[27] D. M. Rogge, B. Rivard, J. Zhang, A. Sanchez, J. Harris, and J. Feng, "Integration of spatial–spectral information for the improved extraction of endmembers," *Remote Sens. Environ.*, vol. 110, no. 3, pp. 287–303, 2007.

[28] J. Zhang, B. Rivard, and D. M. Rogge, "The successive projection algorithm (SPA), an algorithm with a spatial constraint for the automatic search of endmembers in hyperspectral data," *Sensors*, vol. 8, no. 2, pp. 1321–1342, 2008.

[29] S. Mei, M. He, Z. Wang, and D. Feng, "Spatial purity based endmember extraction for spectral mixture analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3434–3445, Sep. 2010.

[30] H. Li and L. Zhang, "A hybrid automatic endmember extraction algorithm based on a local window," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4223–4238, Nov. 2011.

[31] M. Xu, B. Du, and L. Zhang, "Spatial-spectral information based abundance-constrained endmember extraction methods," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2004–2015, Jun. 2014.

[32] M. C. Torres-Madronero and M. Velez-Reyes, "Integrating spatial information in unsupervised unmixing of hyperspectral imagery using multiscale representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1985–1993, Jun. 2014.

[33] M. Xu, L. Zhang, and B. Du, "An image-based endmember bundle extraction algorithm using both spatial and spectral information," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2607–2617, Jun. 2015.

[34] L. Wang, C. Shi, C. Diao, W. Ji, and D. Yin, "A survey of methods incorporating spatial information in image classification and spectral unmixing," *Int. J. Remote Sens.*, vol. 37, no. 16, pp. 3870–3910, 2016.

[35] X. Shen, W. Bao, and K. Qu, "Clustering based spatial spectral preprocessing for hyperspectral unmxing," in *Proc. 4th Int. Conf. Commun. Inf. Process.*, 2018, pp. 313–316.

[36] X. Shen and W. Bao, "A spatial energy and spectral purity based preprocessing algorithm for fast hyperspectral endmember extraction," in *Proc. 10th Workshop Hyperspectral Imag. Signal Process., Evol. Remote Sens.*, 2019, pp. 1–5.

[37] X. Shen and W. Bao, "Hyperspectral endmember extraction using spatially weighted simplex strategy," *Remote Sens.*, vol. 11, no. 18, 2019, Art. no. 2147.

[38] Y. Yan, W. Hua, X. Liu, Z. Cui, and D. Diao, "Spatial–spectral pre-processing for spectral unmixing," *Int. J. Remote Sens.*, vol. 40, no. 4, pp. 1357–1373, 2019.

[39] M. Parente, J. T. Clark, A. J. Brown, and J. L. Bishop, "End-to-end simulation and analytical model of remote-sensing systems: Application to CRISM," *IEEE Trans. Geosci. Remote Sen.*, vol. 48, no. 11, pp. 3877–3888, Nov. 2010.

[40] C. E. Shannon, "A mathematical theory of communication," *ACM SIG-MOBILE Mobile Comput. Commun. Rev.*, vol. 5, no. 1, pp. 3–55, 2001.

[41] G. P. Abousleman, M. W. Marcellin, and B. R. Hunt, "Hyperspectral image compression using entropy-constrained predictive trellis coded quantization," *IEEE Trans. Image Process.*, vol. 6, no. 4, pp. 566–573, Apr. 1997.

[42] B. Guo, S. R. Gunn, R. I. Damper, and J. D. Nelson, "Band selection for hyperspectral image classification using mutual information," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 4, pp. 522–526, Oct. 2006.

[43] J. M. Nascimento and J. M. Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan. 2005.

[44] W. Sun and Q. Du, "Hyperspectral band selection: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 118–139, Jun. 2019.

[45] R. Yang, L. Su, X. Zhao, H. Wan, and J. Sun, "Representative band selection for hyperspectral image classification," *J. Vis. Commun. Image Representation*, vol. 48, pp. 396–403, 2017.

[46] A. J. Brown, "Spectral curve fitting for automatic hyperspectral data analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1601–1608, Jun. 2006.

[47] D. Oliva, S. Hinojosa, V. Osuna-Enciso, E. Cuevas, M. Pérez-Cisneros, and G. Sanchez-Ante, "Image segmentation by minimum cross entropy using evolutionary methods," *Soft Comput.*, vol. 23, no. 2, pp. 431–450, 2019.

[48] T. O. Kvålseth, "On the measurement of randomness (uncertainty): A more informative entropy," *Entropy*, vol. 18, no. 5, 2016, Art. no. 159.

[49] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Trans. Inf. theory*, vol. 37, no. 1, pp. 145–151, Jan. 1991.

[50] L. Jost, "Entropy and diversity," *Oikos*, vol. 113, no. 2, pp. 363–375, 2006.

[51] V. Soltan, *Lectures on Convex Sets*, vol. 986. Hackensack, NJ, USA: World Scientific, 2015.

[52] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 713–718, Mar. 1992.

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

[54] F. Zhu, Y. Wang, B. Fan, S. Xiang, G. Meng, and C. Pan, "Spectral unmixing via data-guided sparsity," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5412–5427, 2014.

[55] R. N. Clark *et al.*, "USGS digital spectral library splib06a," U.S. Geological Survey, Digital Data Series 231 2007.

[56] S. Computational Intelligence Group, University of the Basque Country/Euskal Herriko Unibertsitatea (UPV/EHU), "Hyperspectral imagery synthesis (EIAS) toolbox," 2010. Accessed: Mar. 28, 2019. [Online]. Available: http://www.ehu.es/ccwintco/index.php/Hyperspectral_Imagery_Synthesis_to%ols_for_MATLAB

[57] J. M. Bioucas-Dias and J. M. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.

[58] D. C. Heinz and C. -I-Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sen.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

Dharambhai Shah (Student Member, IEEE) received the B.E. degree in electronics and communications engineering and the M.E. degree in signal processing and very large scale integration technology from Gujarat Technological University, Ahmedabad, India, in 2012 and 2014, respectively. He is currently working toward the Ph.D. degree with the Institute of Technology, Nirma University, Ahmedabad.

He was an Assistant Professor with Uka Tarsadia University, Bardoli, India, for more than three years. His research interests include signal processing, hyperspectral unmixing, and machine learning.

Tanish Zaveri (Senior Member, IEEE) received the B.E. degree in electronics engineering from the Sardar Vallabhbhai National Institute of Technology (SVNIT), Surat, India, in 1998, the M.Tech. degree in biomedical engineering from the Indian Institute of Technology Bombay, Mumbai, India, in 2005, and the Ph.D. degree in image fusion engineering from SVNIT, in 2010.

He is currently a Professor with the Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, Ahmedabad, India. He has completed various funded projects on remote sensing. His research interests include remote sensing, video processing, and medical image processing.

Yogesh N. Trivedi received the B.E. degree in electronics from Sardar Patel University, Anand, India, in 1993, the M.E. degree in industrial electronics from the Maharaja Sayajirao University of Baroda, Vadodara, India, in 1995, and the Ph.D. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 2011.

He is currently a Professor with the Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, Ahmedabad, India. From 1995 to 1996, he was a Lecturer with the Faculty of Technology and Engineering, the Maharaja Sayajirao University of Baroda. Since 1996, he has been with Nirma University, where he works as a PG Coordinator for Communication Engineering program and is the Head of the Centre for Continuing Education. He has authored or coauthored several papers in national/international conferences and international journals. He has completed two research projects sponsored by GUJCOST, Gandhinagar, India. He has been guiding students for their M.Tech and Ph.D. degrees with Nirma University. He has organized several short-term training programs for the benefit of research scholars. His current research interests include signal processing, wireless communication, and cognitive radio.

Antonio Plaza (Fellow, IEEE) received the M.Sc. and Ph.D. degrees in computer engineering from the University of Extremadura, Badajoz, Spain, in 1999 and 2002, respectively.

He is currently a Full Professor and the Head of the Hyperspectral Computing Laboratory, the Department of Technology of Computers and Communications, the University of Extremadura. He has authored more than 600 publications, including 284 JCR journal papers (more than 170 in IEEE journals), 24 book chapters, and more than 300 peer-reviewed conference proceeding papers. He has guest edited ten special issues on hyperspectral remote sensing for different journals. He has reviewed more than 500 manuscripts for more than 50 different journals. His main research interests include hyperspectral data processing and parallel computing of remote sensing data.

Prof. Plaza was a recipient of the recognition of Best Reviewers of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, in 2009 and the recognition of Best Reviewers of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, in 2010. He is also a recipient of the Best Column Award of the *IEEE Signal Processing Magazine*, in 2015, the 2013 Best Paper Award of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and the most highly cited paper (2005–2010) in the *Journal of Parallel and Distributed Computing*. He received best paper awards at the IEEE International Conference on Space Technology and the IEEE Symposium on Signal Processing and Information Technology. He was an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING from 2007 to 2012. He is an Associate Editor for the IEEE ACCESS. He was a member of the Editorial Board of the IEEE GEOSCIENCE AND REMOTE SENSING NEWSLETTER from 2011 to 2012 and the *IEEE Geoscience and Remote Sensing Magazine* in 2013. He was also a member of the steering committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. He was the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) from 2011 to 2012 and the President of the Spanish Chapter of the IEEE GRSS from 2012 to 2016. He was the Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING from 2013 to 2017. He was included in the 2018 and 2019 Highly Cited Researchers List (Clarivate Analytics).