Subspace Structure Regularized Nonnegative Matrix Factorization for Hyperspectral Unmixing

Lei Zhou, Xueni Zhang, Jianbo Wang, Xiao Bai, Lei Tong, Liang Zhang, Jun Zhou, and Edwin Hancock

***Abstract*—Hyperspectral unmixing is a crucial task for hyperspectral images (HSIs) processing, which estimates the proportions of constituent materials of a mixed pixel. Usually, the mixed pixels can be approximated using a linear mixing model. Since each material only occurs in a few pixels in real HSI, sparse nonnegative matrix factorization (NMF), and its extensions are widely used as solutions. Some recent works assume that materials are distributed in certain structures, which can be added as constraints to sparse NMF model. However, they only consider the spatial distribution within a local neighborhood and define the distribution structure manually, while ignoring the real distribution of materials that is diverse in different images. In this article, we propose a new unmixing method that learns a subspace structure from the original image and incorporate it into the sparse NMF framework to promote unmixing performance. Based on the self-representation property of data points lying in the same subspace, the learned subspace structure can indicate the global similar graph of pixels that represents the real distribution of materials. Then the similar graph is used as a robust global spatial prior which is expected to be maintained in the decomposed abundance matrix. The experiments conducted on both simulated and real-world HSI datasets demonstrate the superior performance of our proposed method.**

***Index Terms*—Hyperspectral unmixing (HU), linear mixing model (LMM), nonnegative matrix factorization (NMF), subspace structure, similar graph.**

I. Introduction

HYPERSPECTRAL image (HSI) analysis [1]–[5] is one of the fastest-growing technologies in recent years. However, due to low spatial resolution or specific imaging mechanism, the acquired HSIs often contain mixed pixels which span surface areas containing several types of materials. To effectively exploit hyperspectral data, hyperspectral unmixing (HU) [6]–[9] has become a basic preprocessing for effective HSI analysis.

The objective of HU is to decompose mixed pixels into components with the reference spectral signatures of each of the materials present (endmembers), and to determine their corresponding fractions (abundances). Existing unmixing algorithms mainly exploit one of two mixture models—namely, a linear model or a nonlinear model. Nonlinear mixing models [10], [11] assume that the observed pixel is mixed by a nonlinear function of the component spectral signatures of the endmembers which are weighted by the corresponding abundances. However, the process of nonlinear combination is usually difficult to model physically and to recover in real-world applications. In recent years, linear mixing model (LMM) [12] has therefore been more widely adopted in most works on HU. The reason for this is the balance between model accuracy and tractability. LMM is based on the assumption that different endmembers are mutually independent, so that the observed HSI is a linear combination of the endmembers and their corresponding abundances.

Abundant LMM unmixing algorithms have been proposed. Some of these focus on the endmember extraction from statistical and geometrical aspects, such as pixel purity index [13], N-FINDR [14], alternating projected subgradients [15], vertex component analysis [16], independent component analysis [17], and minimum-volume-based unmixing algorithms [18], etc. Other methods address the problem of abundance estimation under the assumption that the endmembers are available [19]. With the almost universal success of deep learning, there are also examples of deep neural network-based HU methods [20]–[22]. However, these methods depend on the availability of large amount of training data with groundtruth. In this article, we focus on blind unmixing which learns the endmembers as well as their abundances simultaneously. Nonnegative matrix factorization (NMF) [23], [24] is the most commonly used method for blind source separation. It aims to decompose mixed data through the product of two nonnegative matrices. This is done by minimizing the reconstruction error as measured by Euclidean distance. However, the solution of NMF is usually not unique if there are no further constraints [25]. To alleviate this problem, two kinds of constraints are commonly used on the abundance matrix.

The first is the sparsity constraint for abundance matrix. This is based on the fact that the pixels of HSI are mostly mixed by a relatively small number of endmembers. Therefore, [26], [27] presented a spare coding method on the abundance matrix for

Manuscript received April 27, 2020; revised June 16, 2020 and July 10, 2020; accepted July 14, 2020. Date of publication July 22, 2020; date of current version August 7, 2020. The code for this article is available on https://github.com/zlbuaa/Subspace-Regularized-Unmixing. *(Lei Zhou, Xueni Zhang, and Jianbo Wang are co-first authors.) (Corresponding authors: Xiao Bai; Liang Zhang.)*

Lei Zhou, Xueni Zhang, Xiao Bai, and Liang Zhang are with the School of Computer Science and Engineering, Beijing Advanced Innovation Center for Big Data and Brain Computing, State Key Laboratory of Software Development Environment, Jiangxi Research Institute, Beihang University, Beijing 100191, China (e-mail: leizhou@buaa.edu.cn; zhangxuenii@126.com; baixiao@buaa.edu.cn; liang.z@buaa.edu.cn).

Jianbo Wang is with the First Clinical Medical College of Nanchang University, Nanchang 330006, China (e-mail: 793402931@qq.com).

Lei Tong is with the Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China (e-mail: lei_tong@bjut.edu.cn).

Jun Zhou is with the School of Information and Communication Technology, Griffith University, Nathan, Qld 4111, Australia (e-mail: jun.zhou@griffith.edu.au).

Edwin Hancock is with the Department of Computer Science, University of York, YO10 5DD York, U.K. (e-mail: edwin.hancock@york.ac.uk).

Digital Object Identifier 10.1109/JSTARS.2020.3011257

HU. In this article, L_p denotes the p norm. In fact, provided $L_p(0 \leq p \leq 1)$ then the regularizer has the effect of leading to a sparse solution. Moreover, the sparsity of the $L_p(\frac{1}{2} \leq p \leq 1)$ solution is negatively correlated with p, but the sparsity of the solution for $L_p(0 \leq p \leq \frac{1}{2})$ is not sensitive with the change of p. Therefore, Qian *et al.* [28] utilized the $L_{1/2}$ regularizer on the abundance matrix to constrain the sparseness. It has been proved that the $L_{1/2}$ regularizer is more efficient in computation compared to the L_1 regularizer, and the solution is also closer to the groundtruth. In addition, to avoid the influence of noise, many norm-based robust NMF methods have been proposed. The $L_{2,1}$ norm is commonly integrated into sparse NMF to achieve robustness for pixel noise and outlier rejection since it is rotationally invariant [19], [29], [30]. Additionally, the $L_{1,2}$ norm is also effective for solving band noise problems [31], [32].

The second type of constraint incorporates information concerning the spatial distribution into abundance estimation, and has proved useful in improving the unmixing results. This is due to the fact that endmembers are distributed to form coherent geometric structures, and two correlated pixels usually have similar fractional abundances for the same endmembers. Therefore, the total variation (TV) regularizer [33]–[35] was incorporated to promote piece-wise smooth transitions in the abundance matrix for neighboring pixels of the same endmember category. In [36], abundance separation and smoothness constrained NMF was proposed for HU. The abundance separation acted on the spectral domain, and the abundance smoothness constraint was used on the spatial domains to exploit the spatial information. Due to the spatial structure learning ability of manifold method, [37] incorporated manifold structures learning into the NMF model to separate similar neighboring pixels. Inspired by the denoising method [38], Lu *et al.* [39] proposed a structure constrained sparse NMF method which exploited clustering-based approach to find the potential structure information. In [40], a clustered multitask network was proposed to solve the unmixing problem, which also used the clustering method to explore the distribution. Recently, spatial group sparsity regularized NMF (SGSNMF) [41] utilized superpixels that are obtained from image segmentation as a spatial prior to the promotion of HU.

Although the above methods try to exploit the spatial distribution of pixels, all these methods explore the correlations of pixels within a local neighborhood, and most of them are defined manually. However, each material usually occurs in many different regions in the same HSI. Thus, the spatial distribution of a particular material is not limited to local structures. Moreover, it is obvious that the distributions of materials may be quite diverse in different images. According to [42], each kind of land-cover material in a remotely sensed HSI can be treated as a different subspace. They might have different spectra because of the varying illumination, topography, and other imaging conditions. Therefore, the spatial distribution information can be captured by the subspace structure [43]. This not only represents the global distribution of the materials but can also be learned from the corresponding image. Motivated by this fact, we propose a new method aimed at incorporating the subspace structure regularization into the sparse NMF-based unmixing process. In contrast to deep subspace learning method [44], [45], here we utilize a low-rank representation (LRR) method [46], [47] to learn the similar graph that represents the subspace structures for all materials and which contains the correlations of all pixel pairs. Since the LRR constraint can be incorporated into the NMF constraint, this offers the advantage that we can optimize the subspace learning and the HU simultaneously. As a result, the spatial prior is integrated through regularization into sparse NMF and can be used to perform the HU. Furthermore, based on the assumption that an abundance matrix can be seen as the denoised feature vectors of the original image, the learned abundance matrix can be used to better learn the latent subspace structure. Hence, we introduce a novel joint framework to simultaneously optimize HU and subspace structure learning in a manner which leads to mutual enhancement.

The main contributions of this article are summarized as follows.

1) We propose a new HU method which learns the subspace structure of material reflectance to capture the global correlation of all pixels. Then the global similar graph for materials is used as a robust spatial prior to improvement of the quality of the HU.
2) We design an objective function to integrate the spectral-spatial-based unmixing and subspace structure learning into a single unified framework, in which they can be jointly optimized by an iterative algorithm. The joint framework can not only enhance the unmixing performance but also provide better subspace clustering results.
3) Experiments on both simulated and real-world HSI datasets indicate the superiority of the proposed method, which achieves comparable performance to state-of-the-art methods for HU.

The remainder of this article is structured as follows. Section II describes the background of the LMM and NMF algorithms. Section III presents our proposed method and demonstrates the implementation details. The experimental results on simulated data and real-world HSI data are presented in Section IV. Finally, we conclude this article in Section V.

II. Background

A. NMF for HU

The classic LMM for HU is based on the assumption that the observed HSI is a linear mixture of several endmembers. Consider a HSI $\mathbf{Y} \in \mathbb{R}^{L \times N}$, where the number of wavelength-indexed bands is L and the number of pixels is N. Then the original data $\mathbf{Y}$ can be reconstructed by a linear combination of endmembers as follows:

$$\mathbf{Y} = \mathbf{AS} + \mathbf{E} \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{L \times P}$ denotes the endmember matrix, in which each column represents the spectral signature of the corresponding endmember and P is the number of endmembers; $\mathbf{S} \in \mathbb{R}^{P \times N}$ denotes the abundance matrix, in which each column is the fractions of all endmembers in the corresponding pixel; and $\mathbf{E}$ is an additive Gaussian white noise.

Since the goal of HU is to estimate the endmember and abundance matrices simultaneously, in this task, we only know

the matrix $\mathbf{Y}$, and matrices $\mathbf{A}$ and $\mathbf{S}$ are the unknown targets of unmixing. To avoid the large solution space, two commonly adopted constraints can be used on the matrices $\mathbf{A}$ and $\mathbf{S}$ [48]. The first is the so-called abundance sum-to-one constraint, which restrict the proportions of each endmember sum to one. Another is the nonnegativity constraint, which restricts elements in both the endmember and abundance matrices must be greater than or equal to zero.

With the nonnegativity constraint, NMF is a good way to decompose the original image into the endmember and abundance matrices simultaneously. By reconstructing the original image $\mathbf{Y}$ through endmember matrix $\mathbf{A}$ and abundance matrix $\mathbf{S}$, the target of optimization can be defined as

$$C(\mathbf{A},\mathbf{S}) = \frac{1}{2}\|\mathbf{Y}-\mathbf{AS}\|_F^2 \quad \text{s.t. } \mathbf{A}>0, \mathbf{S}>0 \tag{2}$$

where $\|\cdot\|_F$ represents the Frobenius norm.

The multiplied iterative algorithm is commonly used to solve this objective function. When applied to (2), the multiplicative rule leads to the following two interleaved equations:

$$\mathbf{A} = \mathbf{A}.*\mathbf{YS}^T./\mathbf{ASS}^T \tag{3}$$

$$\mathbf{S} = \mathbf{S}.*\mathbf{A}^T\mathbf{Y}./\mathbf{A}^T\mathbf{AS} \tag{4}$$

where $(\cdot)^T$ denotes the matrix transposition, $.*$ denotes element-wise multiplication, and $./$ denotes element-wise division.

B. NMF With Sparsity Constraints

There are several drawbacks of the traditional NMF model (2). First, it is nonconvex, which means it is hard to get the globally optimal solution. Second, the solution of this objective function is not unique, this is because $\mathbf{AS}$ can be replaced by $(\mathbf{AD})(\mathbf{D}^{-1}\mathbf{S})$ for any nonnegative invertible matrix $\mathbf{D}$. Therefore, the classical NMF model will make the unmixing process unstable. To solve this problem, more computationally tractable constraints are incorporated into NMF.

Due to the fact that each endmember does not occur over the entire image, in most cases the abundance map is sparse. Consider NMF subject to a sparsity constraint, the objective function consists of the reconstruction error and sparsity constraint can be defined as follows:

$$C(\mathbf{A},\mathbf{S}) = \frac{1}{2}\|\mathbf{Y}-\mathbf{AS}\|_F^2 + \lambda f(\mathbf{S}) \tag{5}$$

where λ is a regularization term.

Many varieties of regularizer $f(\cdot)$ exist such that sparsity is encouraged. In this article, we choose to use the $L_{1/2}$ regularizer, which is an alternative to the L_1 regularizer. It has been proved in [28] that the $L_{1/2}$ regularizer is more efficient in computation compared to the L_1 regularizer, and the solution is closer to the groundtruth. The $L_{1/2}$ regularized NMF model is defined as

$$C(\mathbf{A},\mathbf{S}) = \frac{1}{2}\|\mathbf{Y}-\mathbf{AS}\|_F^2 + \lambda\|\mathbf{S}\|_{1/2}. \tag{6}$$

III. Approach

In this section, we propose a new method that utilizes both sparsity constraint and spatial information. First, we describe the spatial information used, and which is obtained by learning subspace structure from the original image. Then a joint framework is proposed to simultaneously perform HU and subspace structure learning.

A. Proposed Method

The traditional spectral-based NMF methods for HU usually independently processes the HSI pixels, while ignoring the spatial correlation of pixels. However, as mentioned in Section I, spatial autocorrelation is important prior knowledge for boosting the performance of HU. In previous works, several spatial regularization terms have been introduced. They are based on the assumption that pixels distributed in a local group are more likely to have the same mixed pattern in the abundance matrix. By taking benefit from the spatial structure constraints, the performance of HU has been greatly improved. However, these methods only utilize the local similarity of image pixels to achieve good performance while ignoring the global similarity over the entire image. In most cases, specific materials are distributed in different regions in the HSI. Hence, the global structure similarity shall be considered in the unmixing task.

Fig. 1 is an illustration of HU models that take different spatial regularization into consideration. By rescaling the original 3-D HSI cube into a 2-D matrix where each column denotes the spectral signature of a pixel, the observed image is expected to be approximated by two matrices—the endmember matrix and the abundance matrix. Since the endmembers are distributed in certain structure in the original images, such structure information are expected to be kept in the abundance matrix. Several spatial structures used in recent works are compared in the right of Fig. 1. In this figure, the pixels that consist of the same set of endmembers are represented in one color. We can see that there are three materials in the observed HSI, which are represented as "blue," "yellow," and "green," respectively, and they occur in different regions in the whole image. Consider the blue pixel marked with a black box in the original image, different methods capture different spatial information with different spatial structures. Fig. 1(a) shows the spatial information used by TV regularizer. It only correlates four neighbors of a pixel to promote piece-wise smooth. Instead of using Euclidean distance to measure the spatial structure, manifold regularizer in Fig. 1(b) tries to exploit the latent manifold structure of the data using heat kernel. As for the spatial group sparsity regularizer showed in Fig. 1(c), superpixels that obtained by segmentation are used to represent the spatial neighborhood. However, as mentioned before, our proposed subspace structure regularizer considers the correlation of the pixels over the entire image. It aims to explore the global structure of data to enhance the HU process, as shown in Fig. 1(d).

Subspace structure learning methods are based on the self-representation property that data points lying in the same subspace can be approximated as a linear combination of the data points from the same subspace. Therefore, the subspace structure of HSI can capture the global correlation of similar pixels which can be used as a robust spatial prior for unmixing. In our research, we make the assumption that each type of endmember forms a

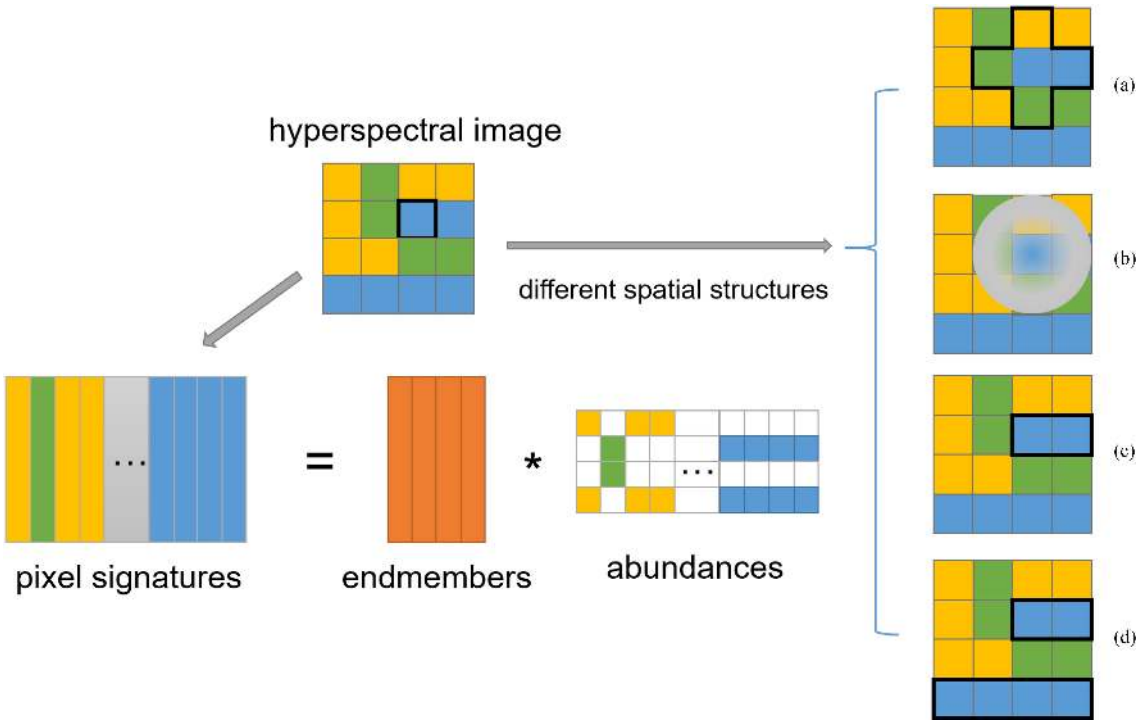

Fig. 1. Illustration of the concept of our method and several alternative methods. The original images are decomposed into two matrices—the endmembers matrix and the abundance matrix. When maintaining the spatial structure for each pixel in the abundance matrix, different methods utilized different strategies. Take the blue pixel marked with black box in the original image as an example. (a) TV regularizer considers its four neighborhood as the local structure. (b) Manifold regularizer uses heat kernel to capture the local structure. (c) Segmentation-based regularizer learns a local neighborhood. (d) Our proposed method learns a subspace structure that represents the global distribution of each material.

subspace, and all variations of endmember in the same type form the data points in the subspace.

To exploit the expected global subspace structure, we first introduce LRR, which is a classic subspace learning method. Consider the dataset $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_N]$ in $\mathbb{R}^L$, according to the self-representation property, each data points can be self-represented by themselves

$$\mathbf{Y} = \mathbf{YZ}$$

where $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_N]$ is the self-representation matrix, each $\mathbf{z}_i$ is the representation coefficient of $\mathbf{y}_i$. By looking for a LRR of $\mathbf{Z}$, the global structure of data $\mathbf{Y}$ can be obtained

$$\min_{\mathbf{Z}} rank(\mathbf{Z}) \quad \text{s.t.} \quad \mathbf{Y} = \mathbf{YZ} \tag{7}$$

whose optimal solutions $\mathbf{Z}^*$ is called the lowest-rank representations of data $\mathbf{Y}$. However, it is difficult to solve this optimization problem, since the rank function is discrete. As the nuclear norm is a good convex approximation of matrix rank, the optimization problem can be transformed as follows:

$$\min_{\mathbf{Z}} \|\mathbf{Z}\|_* \quad \text{s.t.} \quad \mathbf{Y} = \mathbf{YZ}. \tag{8}$$

Here, $\|\mathbf{Z}\|_*$ is the nuclear norm which is the sum of the singular values of the matrix.

Since the self-representation matrix $\mathbf{Z}$ contains the correlation of all pixels, it is natural to preserve this similarity in abundance matrix. In other words, the pixels in the same subspace in the original image should exist in the same subspace in the abundance matrix.

Based on the fact that there are many mixed pixels in HSI, HU is widely used as a crucial preprocessing step for HSI analysis [49] since the obtained abundance can be seen as a denoised feature representation. Therefore, it is better to preserve the latent subspace structure from the unmixed abundance map instead of the original images. By incorporating the subspace regularizer into the sparse NMF model, the optimization problem can be formulated as

$$J(\mathbf{A}, \mathbf{S}, \mathbf{Z}) = \min_{\mathbf{A},\mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathbf{AS}\|_F^2 + \lambda \|\mathbf{S}\|_{1/2} + \mu \|\mathbf{S} - \mathbf{SZ}\|_F^2 \quad \text{s.t.} \quad \mathbf{A} \geq 0, \mathbf{S} \geq 0, \mathbf{1}_K^T \mathbf{S} = \mathbf{1}_N^T \tag{9}$$

where the first two terms are reconstruction error and sparsity constraint, and the third term constrains the subspace structure of the abundance matrix.

Note that we would also want to simultaneously learn and optimize the subspace structure. Therefore, a joint framework on HU and subspace learning can be represented as follows:

$$J(\mathbf{A}, \mathbf{S}, \mathbf{Z}) = \min_{\mathbf{A},\mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathbf{AS}\|_F^2 + \lambda \|\mathbf{S}\|_{1/2} + \mu \|\mathbf{S} - \mathbf{SZ}\|_F^2 + \tau \|\mathbf{Z}\|_* \quad \text{s.t.} \quad \mathbf{A} \geq 0, \mathbf{S} \geq 0, \mathbf{1}_K^T \mathbf{S} = \mathbf{1}_N^T \tag{10}$$

where the first three terms are the objective of spectral-spatial HU and the last two terms learn the latent subspace structures of the materials.

B. Optimization

Obviously, the presented optimization problem is nonconvex. To iteratively solve this problem, we first define an auxiliary variable $\mathbf{L}$, then the optimization problem (10) can be transformed to the following problem:

$$\begin{aligned} J(\mathbf{A},\mathbf{S},\mathbf{Z}) = \min_{\mathbf{A},\mathbf{S}} \frac{1}{2}\left\|\mathbf{Y}-\mathbf{AS}\right\|_F^2 + \lambda\left\|\mathbf{S}\right\|_{1/2} \\ + \mu\left\|\mathbf{S}-\mathbf{SZ}\right\|_F^2 + \tau\left\|\mathbf{L}\right\|_* \\ \text{s.t.}\quad \mathbf{A}\geq 0, \mathbf{S}\geq 0, \mathbf{L}=\mathbf{Z}, \mathbf{1}_K^T\mathbf{S}=\mathbf{1}_N^T. \end{aligned} \tag{11}$$

Here, we consider the auxiliary variable $\mathbf{L}$ as a denoising version of $\mathbf{Z}$, then we can add the $\mathbf{L}=\mathbf{Z}$ constraint to the objective function, and the objective problem can be relaxed as

$$\begin{aligned} J(\mathbf{A},\mathbf{S},\mathbf{Z}) = \min_{\mathbf{A},\mathbf{S}} \frac{1}{2}\left\|\mathbf{Y}-\mathbf{AS}\right\|_F^2 + \lambda\left\|\mathbf{S}\right\|_{1/2} \\ + \mu\left\|\mathbf{S}-\mathbf{SZ}\right\|_F^2 + \frac{1}{2}\left\|\mathbf{L}-\mathbf{Z}\right\|_F^2 + \tau\left\|\mathbf{L}\right\|_* \\ \text{s.t.}\quad \mathbf{A}\geq 0, \mathbf{S}\geq 0, \mathbf{1}_K^T\mathbf{S}=\mathbf{1}_N^T. \end{aligned} \tag{12}$$

Subsequently, we utilize the multiplicative iterative method [24] to solve the above problem (12). Four steps are iteratively updated with other variables fixed: 1) endmember matrix estimation; 2) abundance matrix estimation; 3) reconstruction; and 4) low-rank self-representation learning. The details of each step are as follows.

1) Endmember Estimation: In this step, we use the Lagrange multiplier method to estimate the endmember matrix with other variables fixed. Then the objective function is reformulated as

$$\begin{aligned} J(\mathbf{A}) = \min_{\mathbf{A}} \frac{1}{2}\left\|\mathbf{Y}-\mathbf{AS}\right\|_F^2 + Tr(\mathbf{\Psi A}) \\ \text{s.t.}\quad \mathbf{A}\geq 0 \end{aligned} \tag{13}$$

where $\mathbf{\Psi}$ is the Lagrange multiplier. To solve this problem (13), a common method is to separate this equation and set the last term to 0. We can obtain the following equations with the Karush–Kuhn–Tucker (K–K–T) conditions

$$\nabla_{\mathbf{A}} J(\mathbf{A}) = \mathbf{ASS}^T - \mathbf{YS}^T + \mathbf{\Psi} = \mathbf{0} \tag{14}$$

$$\mathbf{A}.*\mathbf{\Psi} = \mathbf{0}. \tag{15}$$

By simultaneously multiplying both sides by $\mathbf{A}$ on (14), and then substituting (15) into (14), the endmember matrix $\mathbf{A}$ can be updated as

$$\mathbf{A} \longleftarrow \mathbf{A}.*\mathbf{YS}^T./\mathbf{ASS}^T. \tag{16}$$

2) Abundance Estimation: When the endmember matrix is updated, we fix matrix $\mathbf{A}$. Then the objective function for abundance matrix estimation can be written as

$$\begin{aligned} J(\mathbf{S}) = \min_{\mathbf{S}} \frac{1}{2}\left\|\mathbf{Y}-\mathbf{AS}\right\|_F^2 + \lambda\left\|\mathbf{S}\right\|_{1/2} \\ + \mu\left\|\mathbf{S}-\mathbf{SZ}\right\|_F^2 + Tr(\mathbf{\Gamma A}) \\ \text{s.t.}\quad \mathbf{S}\geq 0, \mathbf{1}_K^T\mathbf{S}=\mathbf{1}_N^T. \end{aligned} \tag{17}$$

The same with endmember estimation, the Lagrange multiplier method is adopted to solve problem (17). Where $\mathbf{\Gamma}$ is the Lagrange multiplier with size $K\times N$. In the same manner, the following is obtained by the K–K–T conditions:

$$\begin{aligned} \nabla_{\mathbf{S}} J(\mathbf{S}) = \mathbf{A}^T\mathbf{AS} - \mathbf{A}^T\mathbf{Y} + \frac{\lambda}{2}\mathbf{S}^{-1/2} \\ + 2\mu\mathbf{S}(\mathbf{I}-\mathbf{Z})(\mathbf{I}-\mathbf{Z})^T + \mathbf{\Gamma} = 0 \end{aligned} \tag{18}$$

$$S.*\mathbf{\Gamma} = 0. \tag{19}$$

Similarly, we multiply both sides by $\mathbf{S}$ on (18) and substitute (19) into (18), the abundance matrix $\mathbf{S}$ can be updated as

$$\begin{aligned} \mathbf{S} \longleftarrow \mathbf{S}.*\mathbf{A}^T\mathbf{Y}./\Big(\mathbf{A}^T\mathbf{AS} \\ + \frac{\lambda}{2}\mathbf{S}^{-1/2} + 2\mu\mathbf{S}(\mathbf{I}-\mathbf{Z})(\mathbf{I}-\mathbf{Z})^T\Big). \end{aligned} \tag{20}$$

3) Reconstruction: In this step, we solve the reconstruction problem with endmember matrix $\mathbf{A}$ and abundance matrix $\mathbf{S}$ fixed. The objective function is as follows:

$$J(\mathbf{Z}) = \min \mu\left\|\mathbf{S}-\mathbf{SZ}\right\|_F^2 + \frac{1}{2}\left\|\mathbf{L}-\mathbf{Z}\right\|_F^2. \tag{21}$$

By solving the above, we can get the following updating rule:

$$\mathbf{Z} \longleftarrow \mathbf{Z}.*\left(\mathbf{S}^T\mathbf{S}+\frac{2}{\mu}\mathbf{L}\right)./\left(\mathbf{S}^T\mathbf{SZ}+\frac{2}{\mu}\mathbf{Z}\right). \tag{22}$$

4) Low-Rank Self-Representation Learning: In the fourth step, the low-rank self-representation matrix is optimized by the following objective function:

$$J(\mathbf{L}) = \tau\left\|\mathbf{L}\right\|_* + \frac{1}{2}\left\|\mathbf{L}-\mathbf{Z}\right\|_F^2. \tag{23}$$

This problem has a closed-form solution and can be solved via the singular value thresholding operator [50].

Then, we solve the objective function (12) with a multiplicative iterative method. The entire process is summarized in Algorithm 1. Finally, we analyze the computational complexity of the proposed method. Compared with standard NMF, there are two more steps to compute the self-representation matrix $\mathbf{Z}$ and auxiliary variable $\mathbf{L}$. Since the dimension of $\mathbf{Z}$ and $\mathbf{L}$ is $N\times N$, the additional computational cost for $\mathbf{Z}$ and $\mathbf{L}$ is $O(PN^2)$ caused by the Singular Value Decomposition (SVD) operator. The computational complexity of standard NMF is known as $O(LPN)$. Therefore, the overall computational complexity of our method is $O(LPN+PN^2)$ which is similar with the standard NMF and is faster than $L_{1/2}$-NMF with $O(LPN+P^2N^2)$ computational complexity [28].

C. Convergence Analysis

In this section, we analyze the convergence of the proposed updating algorithm. Since we solve the optimization problem by an iterative strategy, to guarantee the convergence of the update

Algorithm 1: Subspace Structure Regularized Sparse NMF.

Input: A hyperspectral image $\mathbf{Y}$.
Output: Endmember matrix $\mathbf{A}$, abundance map $\mathbf{S}$, and self-representation matrix $\mathbf{Z}$.
Initialize $\mathbf{A}$, $\mathbf{S}$, and $\mathbf{Z}$. Let $\mathbf{L}$=$\mathbf{Z}$;
while *the stopping criteria is not reached* **do**
1) fix the others and update $\mathbf{A}$ by Equation (16);
2) fix the others and update $\mathbf{S}$ by Equation (20);
3) fix the others and update $\mathbf{Z}$ by Equation (22);
4) fix the others and update $\mathbf{L}$ by solving problem (23)

rule, we need to prove the nonincreasing property of the objective function in each update step. To formulate this problem, we use $\mathbf{A}^k$, $\mathbf{S}^k$, $\mathbf{Z}^k$, $\mathbf{L}^k$ to denote the values of the kth iteration and $\mathbf{A}^{k+1}$, $\mathbf{S}^{k+1}$, $\mathbf{Z}^{k+1}$, $\mathbf{L}^{k+1}$ to denote the values of the $(k+1)$th iteration. Then, the proof problem can be written as

$$J(\mathbf{A}^{k+1},\mathbf{S}^k,\mathbf{Z}^k,\mathbf{L}^k)\le J(\mathbf{A}^k,\mathbf{S}^k,\mathbf{Z}^k,\mathbf{L}^k) \tag{24}$$

$$J(\mathbf{A}^{k+1},\mathbf{S}^{k+1},\mathbf{Z}^k,\mathbf{L}^k)\le J(\mathbf{A}^{k+1},\mathbf{S}^k,\mathbf{Z}^k,\mathbf{L}^k) \tag{25}$$

$$J(\mathbf{A}^{k+1},\mathbf{S}^{k+1},\mathbf{Z}^{k+1},\mathbf{L}^k)\le J(\mathbf{A}^{k+1},\mathbf{S}^{k+1},\mathbf{Z}^k,\mathbf{L}^k) \tag{26}$$

$$J(\mathbf{A}^{k+1},\mathbf{S}^{k+1},\mathbf{Z}^{k+1},\mathbf{L}^{k+1})\le J(\mathbf{A}^{k+1},\mathbf{S}^{k+1},\mathbf{Z}^{k+1},\mathbf{L}^k). \tag{27}$$

Since the same problems (24), (25), (27) have been proved in [28] and [34], here we only give the proof for problem (26). Similar to [34], we consider each column of Z independently to prove this problem due to the column separability of the objective function (21). Let $\mathbf{z}$, $\mathbf{l}$ denote the same column of $\mathbf{Z}$, $\mathbf{L}$, respectively. Then the objective function becomes

$$J(\mathbf{z})=\min \mu\left\|\mathbf{S}-\mathbf{S}\mathbf{z}\right\|_F^2+\frac{1}{2}\left\|\mathbf{l}-\mathbf{z}\right\|_F^2. \tag{28}$$

To prove the nonincreasing property of the objective function, we first introduce an auxiliary function $G(\mathbf{z},\mathbf{z}^k)$ which meet the conditions $G(\mathbf{z},\mathbf{z})=J(\mathbf{z})$ and $G(\mathbf{z},\mathbf{z}^k)\ge J(\mathbf{z})$. Then $J(\mathbf{z})$ is nonincreasing when use the following updating rule:

$$\mathbf{z}^{k+1}=arg\min_{\mathbf{z}} G(\mathbf{z},\mathbf{z}^k) \tag{29}$$

since

$$J(\mathbf{z}^{k+1})\le G(\mathbf{z}^{k+1},\mathbf{z}^k)\le G(\mathbf{z}^k,\mathbf{z}^k)=J(\mathbf{z}^k). \tag{30}$$

Following [28], G can be defined as

$$G(\mathbf{z},\mathbf{z}^k)=J(\mathbf{z}^k)+(\mathbf{z}-\mathbf{z}^k)(\nabla J(\mathbf{z}^k))^T+\frac{1}{2}(\mathbf{z}-\mathbf{z}^k)K(\mathbf{z}^k)(\mathbf{z}-\mathbf{z}^k)^T \tag{31}$$

where $K(\mathbf{z}^k)$ is a diagonal matrix which is defined as

$$K(\mathbf{z}^k)=diag\left(\left(\mathbf{S}^T\mathbf{S}\mathbf{z}^k+\frac{2}{\mu}\right)./\mathbf{z}^k\right). \tag{32}$$

Since $G(\mathbf{z},\mathbf{z})=J(\mathbf{z})$, the Taylor expansion of $J(\mathbf{z})$ is

$$J(\mathbf{z})=J(\mathbf{z}^k)+(\mathbf{z}-\mathbf{z}^k)(\nabla J(\mathbf{z}^k))^T+\frac{1}{2}(\mathbf{z}-\mathbf{z}^k)\left(\mathbf{S}^T\mathbf{S}+\frac{2}{\mu}\mathbf{I}\right)(\mathbf{z}-\mathbf{z}^k)^T+O(\mathbf{z}) \tag{33}$$

where $O(\mathbf{z})$ denotes the higher-order terms of the Taylor expansion. Then the condition $G(\mathbf{z},\mathbf{z}^k)\ge J(\mathbf{z})$ is satisfied if

$$\frac{1}{2}(\mathbf{z}-\mathbf{z}^k)\left(K(\mathbf{z}^k)-\mathbf{S}^T\mathbf{S}-\frac{2}{\mu}\mathbf{I}\right)(\mathbf{z}-\mathbf{z}^k)^T\ge 0. \tag{34}$$

According to [27], $K(\mathbf{z}^k)-\mathbf{S}^T\mathbf{S}-\frac{2}{\mu}\mathbf{I}$ is a positive semidefinite matrix with the nonnegative $\mathbf{z}$. As aforementioned, next we only need to prove that the update rule (22) is coincident with selecting the minimum of $G(\mathbf{z},\mathbf{z}^k)$. This can be solved by making the gradient to be 0

$$\nabla_{\mathbf{z}}G(\mathbf{z},\mathbf{z}^k)=\mathbf{S}^T(\mathbf{S}\mathbf{z}^k-\mathbf{S})+\frac{2}{\mu}(\mathbf{z}^k-\mathbf{l})+K(\mathbf{z}^k)(\mathbf{z}-\mathbf{z}^k)=0 \tag{35}$$

then, it can be calculated

$$\begin{aligned}\mathbf{z}&=\mathbf{z}^k-K^{-1}\left(\mathbf{S}^T(\mathbf{S}\mathbf{z}^k-\mathbf{S})+\frac{2}{\mu}(\mathbf{z}^k-\mathbf{l})\right)\\&=\mathbf{z}^k-\mathbf{z}^k./\left(\mathbf{S}^T\mathbf{S}\mathbf{z}^k+\frac{2}{\mu}\mathbf{z}^k\right).*\left(\mathbf{S}^T(\mathbf{S}\mathbf{z}^k-\mathbf{S})+\frac{2}{\mu}\left(\mathbf{z}^k-\mathbf{l}\right)\right)\\&=\mathbf{z}^k-\mathbf{z}^k./\left(\mathbf{S}^T\mathbf{S}\mathbf{z}^k+\frac{2}{\mu}\mathbf{z}^k\right).*\left(\mathbf{S}^T\mathbf{S}\mathbf{z}^k+\frac{2}{\mu}\mathbf{z}^k-\mathbf{S}^T\mathbf{S}-\frac{2}{\mu}\mathbf{l}\right)\\&=\mathbf{z}^k./\left(\mathbf{S}^T\mathbf{S}\mathbf{z}^k+\frac{2}{\mu}\mathbf{z}^k\right).*\left(\mathbf{S}^T\mathbf{S}+\frac{2}{\mu}\mathbf{l}\right)\end{aligned} \tag{36}$$

which is coincident with the update rule of (22). That is to say, the proposed update algorithm can make the objective function decrease monotonically at each iteration until convergence has been reached.

D. Implementation Issues

Then, we discuss several issues during the algorithm implementation. As aforementioned issue, the optimization problem is not convex with both $\mathbf{A}$ and $\mathbf{S}$, and an iterative optimization strategy with the above updating rules is proposed to solve it. Therefore, the initialization of the matrix is crucial. Two initialization methods are frequently used—random initialization and vertex component analysis-fully constrained least squares (VCA-FCLS) initialization. Compared to random initialization that setting elements to random values between $[0,1]$, the latter that using VCA [16] to recognize endmembers as the input of $\mathbf{A}$ and then utilizing FCLS [51] to obtain the initial $\mathbf{S}$, is more effective. In this article, we use VCA-FCLS initialization in all the experiments. For self-representation matrix $\mathbf{Z}$, we initialize it using LRR on the original image $\mathbf{Y}$.

Another important issue is how to meet the basic full nonnegativity constraint and additivity constraint. Since the updating rules maintain the sign of matrix values, the former constraint

can be satisfied as long as the initial matrix is nonnegative. In terms of full additivity constraint, we exploit a similar method as [28]. We augment the original data matrix $\mathbf{Y}$ and the endmember matrix $\mathbf{A}$ by a row of constants

$$\mathbf{Y}_f = [\mathbf{Y}; \delta \mathbf{1}_N^T]$$
$$\mathbf{A}_f = [\mathbf{A}; \delta \mathbf{1}_K^T] \tag{37}$$

where δ is a weight parameter that determine the impact of the additivity constraint. When the larger δ, the more accurate the result. However, the convergence will be nonuniform. In practice, $\delta = 15$ is a good choice.

Two stopping criteria are adopted for our iterative optimization. One is to set the maximum number of iterations. We set this to 3000, in common with most alternative iterative NMF methods. The second stopping criterion is the difference in the gradients of the objective function between successive iterations

$$\|\nabla C(\mathbf{A}_i, \mathbf{S}_i)\|_2^2 \leq \epsilon \|\nabla C(\mathbf{A}_1, \mathbf{S}_1)\|_2^2 \tag{38}$$

where ϵ is set to 10^{-3}. If the gradient difference is small enough, the optimal solution is obtained.

IV. Experimental Results and Discussion

To verify the effectiveness of our proposed method, we conducted experiments on both simulated and real-world dataset. The compared HU methods include baseline methods VCA-FCLS [16] and NMF [23], sparsity-based methods $L_{1/2}$-NMF [28] and graph-regularized $L_{1/2}$-NMF (GLNMF) [37], spatial information-based methods SGSNMF [41], TV-RSNMF [34], multilayer NMF method MLNMF [52], and sparsity-constrained deep NMF with TV (SDNMF-TV) [35]. The results were evaluated with two commonly used measures to assess the quantitative unmixing performance—spectral angle distance (SAD) and root-mean-square error (RMSE). The SAD compares the similarity of the estimated signature $\hat{\mathbf{A}}_k$ and the groundtruth endmember $\mathbf{A}_k$, and is defined as

$$SAD_k = \arg\cos\left(\frac{\mathbf{A}_k^T \hat{\mathbf{A}}_k}{\|\mathbf{A}_k^T\| \left\|\hat{\mathbf{A}}_k\right\|}\right). \tag{39}$$

The RMSE is defined as

$$RMSE_k = \left(\frac{1}{N}\left\|\mathbf{S}_k - \hat{\mathbf{S}}_k\right\|^2\right)^{1/2} \tag{40}$$

where $\hat{\mathbf{S}}_k$ is the groundtruth abundance matrix for the kth endmember. As stated above, in general, a smaller SAD or RMSE corresponds to a better result.

A. Experiments on Simulated Data

1) Simulated Data: The simulated dataset in this experiment was generated according to the hyperspectral imagery synthesis (EIAs) toolbox [53]. It is a free software for users to generate simulated HSIs flexibly by controlling several parameters, such as a certain number of groundtruth endmembers, the size of the abundance map, spatial distribution of materials, and different kinds of noises. We randomly selected the endmembers from the U.S. Geological Survey (USGS)[1] mineral spectra library, and generated the corresponding abundance maps according to the Gaussian field. To demonstrate the effectiveness of utilizing the global spatial information, we designed the abundance map by mosaicing four smaller abundance matrix together so that each material occurs in different regions of the entire HSIs. Fig. 2 shows nine selected endmembers and Fig. 3 shows the groundtruth abundance maps built from the nine endmembers. Here, the simulated dataset has a size of 100×100 pixels and 224 spectral bands.

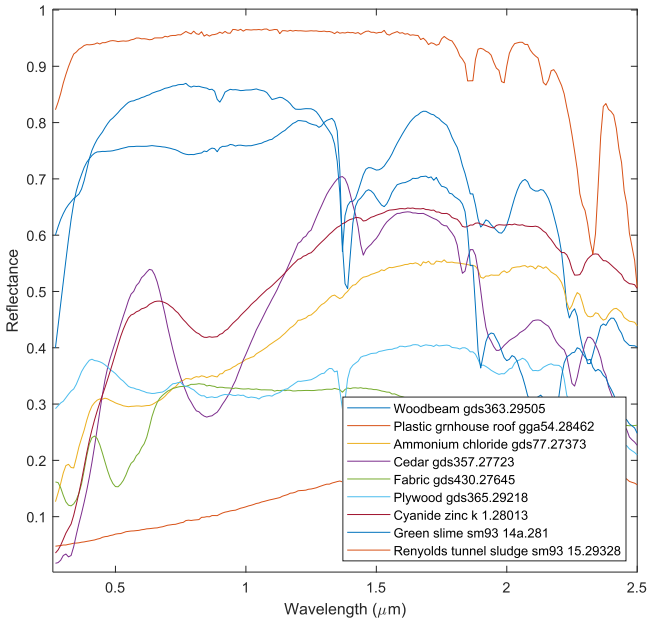

Fig. 2. Spectral curve of the nine endmembers selected from the USGS mineral spectra library on the simulated datasets.

2) Parameter Analysis: There are two key parameters λ and μ in our proposed method, where λ measure the sparsity constraints and μ is for subspace structure regularization. First, we discuss the influence of these two parameters on the simulated dataset at the circumstance of SNR=20 dB. In this experiment, we changed λ at the interval $\{0.0005, 0.001, 0.003, 0.01, 0.05, 0.1, 0.2, 0.3\}$ and μ at the interval $\{0.0001, 0.001, 0.01, 0.1\}$ to test our proposed method. We set parameter τ as 0.001 the same with [34]. The performance of our method for different parameter λ and μ is shown in Fig. 4, where (a) displays the SAD results and (b) displays the RMSE results. In general, SAD and RMSE results with respect to λ and μ reveal the same trend. When λ and μ both are near zero, the results are stable. It should be noted that when λ and μ both are zeros, the results correspond to classic NMF. As λ increases it gradually converges to local minima. When λ is too large, the results will be worse than NMF. Similar trend can be seen in parameter μ. This indicates the effectiveness of the sparsity constraint as well as the subspace structure constraint. With the proper choice of parameter values, the SAD and RMSE can be

[1][Online]. Available: https://www.usgs.gov/labs/spec-lab

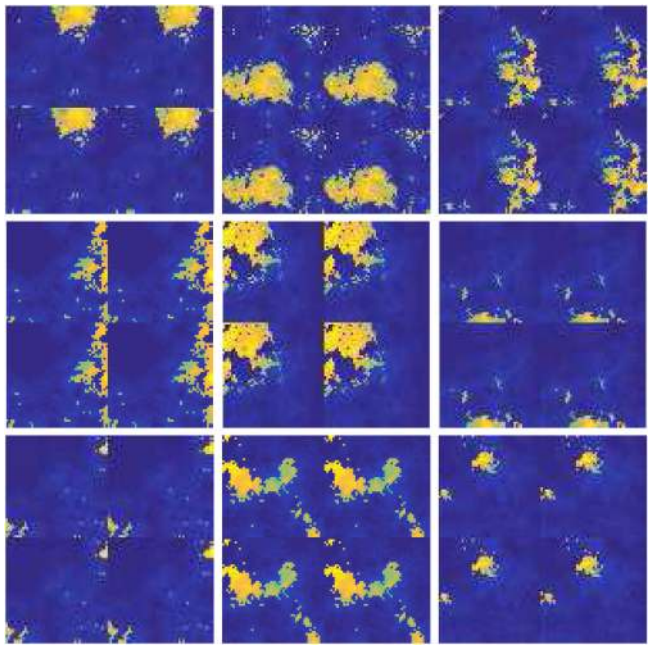

Fig. 3. Abundance maps for the simulated dataset. To demonstrate the effectiveness of our proposed subspace structure constraints, each abundance map consists of four smaller maps built from endmembers in the same subspace.

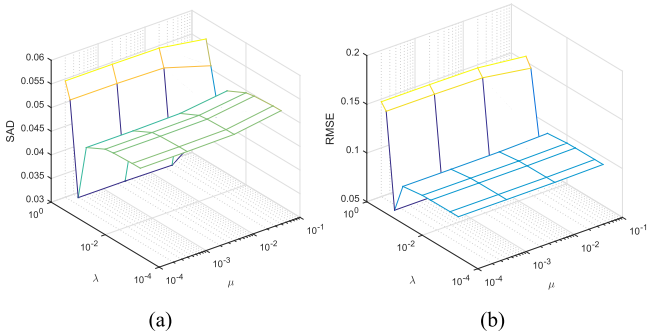

Fig. 4. Performance of our proposed method with respect to different λ and μ when SNR $= 20$ dB. (a) SAD results. (b) RMSE results.

significantly decreased. Relatively, parameter μ is more robust than λ, which can be seen more obviously in RMSE results.

Since the parameter μ is not sensitive to the results, its value is set as 0.01 for our experiments. As for the sparse regularization parameter λ, we utilize the same strategy in [28] to determine its value

$$\lambda_e = \frac{1}{\sqrt{L}} \sum_{l=1}^{N} \frac{\sqrt{N} - \|Y_l\|_1 / \|Y_l\|_2}{\sqrt{N-1}}. \tag{41}$$

Usually, the optimal parameter λ is smaller than λ_e. Therefore we can search it at the interval $[\lambda_e/10, \lambda_e]$. We set the value of λ as 0.1 in the simulated date experiments.

3) Performance Comparisons: Inevitably real-world HSIs are easily corrupted by noises, which is a great challenge for unmixing. Therefore, different levels of white Gaussian noises was added to the simulated data, which exist in most HSIs. We choose the noise level as $\{10, 20, 30, 40\}$ dB. Table I presents the SAD values of different methods under different noise levels. We can see that unmixing methods that integrate spatial information like TV-RSNMF and SDNMF-TV have lower SAD values compared to methods that only use sparsity constraints in most cases. This indicates the effectiveness of the utilization of spatial relationships. It can be seen that our method outperforms all the compared methods on different noise levels except when SNR $= 30$ dB, as the VCA-FCLS is slightly better than our method. This may be caused by some specific noise which influences the subspace clustering result. In most cases, as the mixed noise increases, the unmixing problem becomes more difficult, our method has more obvious advantages. This verifies that the subspace regularizer, which captures the global spatial relationship, is useful in the HU task. Similar results can be seen in Table II, which displays the RMSE values of different methods. In order to demonstrate that the improvement introduced by our method is not at the cost of excessive computational cost, we provide the average running time of different methods. The last row of Table I shows the average running time on different noise levels, we can see that our method is more efficient than most compared methods that define the distribution structure manually. In addition, the time cost of our method is significant superior to the multilayer and deep NMF methods.

To further validate the effectiveness of our proposed subspace structure regularizer, we present the visualization of self-representation matrices of randomly chosen points in different endmembers. As shown in Fig. 5, the lighter areas indicate larger weight in the self-representation matrix. It can be seen that the

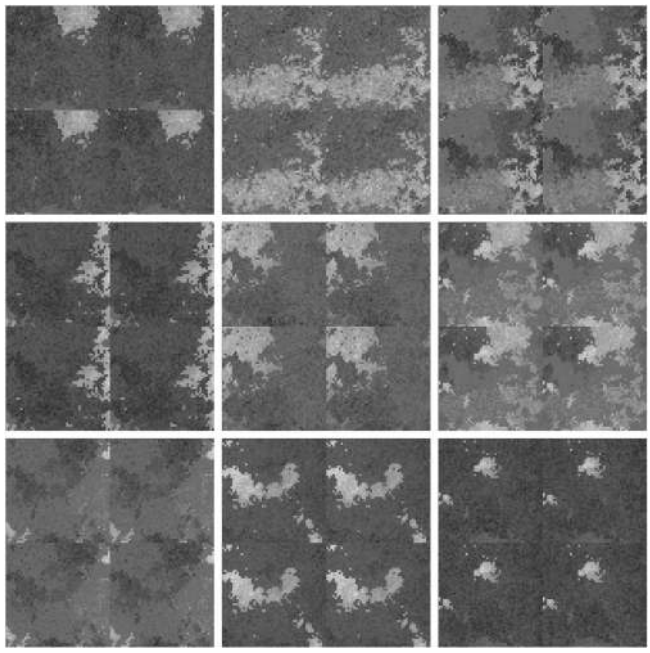

Fig. 5. Visualization of self-representation matrices for different endmembers.

TABLE I
SAD VALUES AND RUNNING TIMES OF THE DIFFERENT METHODS WITH THE SIMULATED DATA

Method	VCA-FCLS	NMF	$L_{1/2}$-NMF	GLNMF	MLNMF	SGSNMF	TV-RSNMF	SDNMF-TV	Ours
SNR=10dB	0.1315	0.1802	0.1774	0.1809	0.1325	0.1392	0.1193	0.1183	**0.1104**
SNR=20dB	0.0366	0.0485	0.0344	0.0361	0.0340	0.0348	0.0509	0.0324	**0.0297**
SNR=30dB	**0.0102**	0.0152	0.0133	0.0124	0.0201	0.0238	0.0200	0.0105	0.0107
SNR=40dB	0.0024	0.0035	0.0027	0.0026	0.0028	0.0236	0.0037	0.0023	**0.0019**
Time(s)	**72.2**	125.4	160.1	187.2	425.8	116.3	190.4	573.5	153.7

TABLE II
RMSE VALUES OF THE DIFFERENT METHODS WITH THE SIMULATED DATA

Method	VCA-FCLS	NMF	$L_{1/2}$-NMF	GLNMF	MLNMF	SGSNMF	TV-RSNMF	SDNMF-TV	Ours
SNR=10dB	0.2397	0.1974	0.1977	0.2092	0.1085	0.1052	0.2256	0.0954	**0.0873**
SNR=20dB	0.1129	0.0914	0.0585	0.0625	0.0612	0.0623	0.0942	0.0593	**0.0545**
SNR=30dB	0.0532	0.0420	0.0408	0.0415	0.0387	0.0398	0.0386	0.0352	**0.0312**
SNR=40dB	0.0179	0.0155	0.0151	0.0142	0.0253	0.0396	0.0134	0.0132	**0.0122**

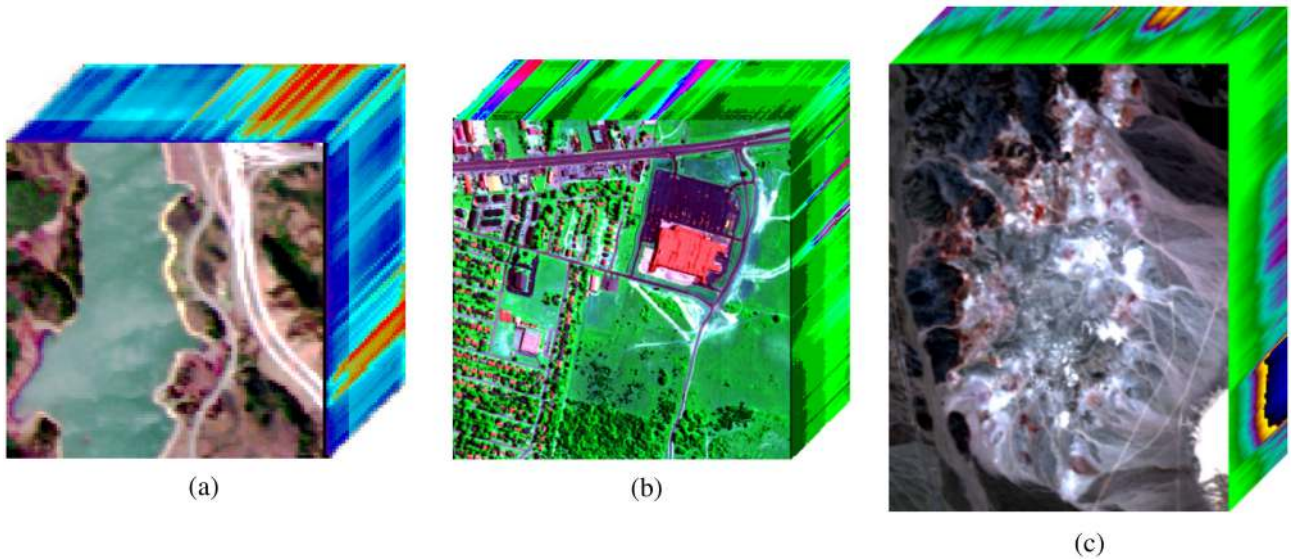

Fig. 6. Three real-world HSIs. (a) HYDICE Jasper Ridge dataset. (b) HYDICE urban dataset. (c) AVIRIS Cuprite dataset.

subspace structure is mostly approximate with the abundance map. Therefore, the learned subspace structure can be used as a robust global spatial prior for unmixing.

B. Experiments on Real Data

In this section, we validate our method on the real-world HSIs. We conducted unmixing experiments on three public hyperspectral datasets—the hyperspectral digital imagery collection experiment (HYDICE) Jasper Ridge dataset, the HYDICE Urban dataset, and the AVIRIS Cuprite dataset. Specifically, we obtain the groundtruth following [16]. For the Cuprite dataset, the reference endmember signatures were chosen from the USGS digital spectral library.

1) HYDICE Jasper Ridge Dataset: Jasper Ridge is a widely used hyperspectral data for evaluating the unmixing method, which contains 512×614 pixels. There are 224 spectral bands from 380 to 2500 nm. Since the groundtruth of this HSI is difficult to obtain, we only used a part of image with 100×100 pixels. Specifically, the first pixel of the chosen part is (105, 269). To avoid the atmospheric effects and dense water vapor problems, we removed related bands (1–3, 108–112, 154–166, 220, 224), remaining an image of 198 bands, which is the same with other HU methods. As shown in Fig 6(a), the endmembers of Jasper Ridge are "Tree," "Soil," "Water," and "Road."

Quantitative evaluation is presented in Table III which shows the mean SAD and RMSE values of different HU methods. As a representative solution, NMF balances the estimation of

TABLE III
SAD AND RMSE VALUES AND RUNNING TIMES OF DIFFERENT METHODS ON THE JASPER RIDGE DATASET

	VCA-FCLS	NMF	$L_{1/2}$-NMF	GLNMF	MLNMF	SGSNMF	TV-RSNMF	SDNMF-TV	Ours
SAD	0.3001	0.3457	0.1891	0.2903	0.1602	0.1686	0.2330	0.1523	**0.1466**
RMSE	0.2367	0.2126	0.1912	0.2185	0.1835	0.1897	0.2279	0.1735	**0.1653**
Time(s)	**31.4**	89.5	108.3	121.8	242.1	80.5	135.9	274.5	99.6

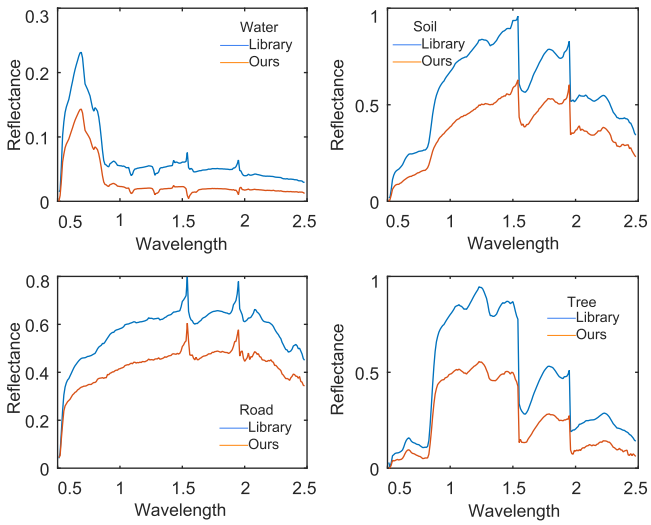

Fig. 7. Comparison of the library spectra with the endmember signatures extracted by our method on the Jasper Ridge dataset.

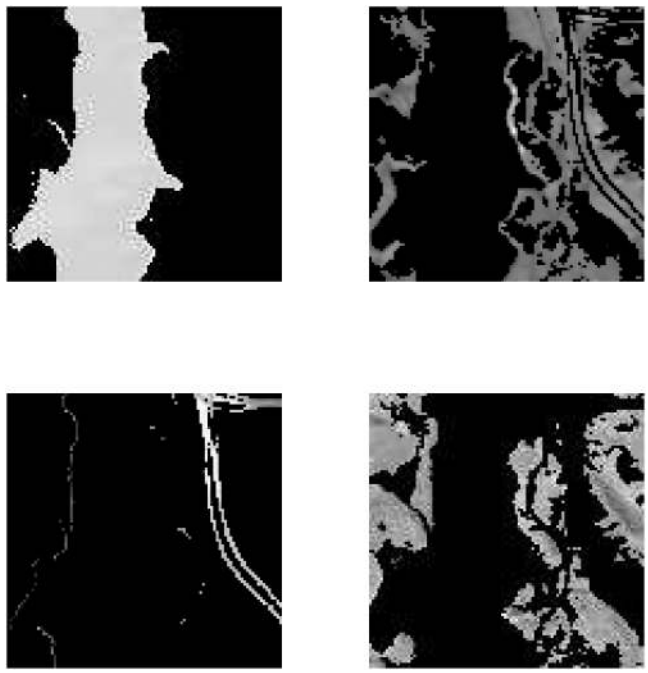

Fig. 8. Abundance maps of four different endmembers obtained using our method on the Jasper Ridge dataset. From left to right and from top to bottom are water, soil, road, and tree, respectively.

endmembers and abundance matrix compared with VCA-FCLS. They both only use nonnegative constraints. $L_{1/2}$-NMF and GLNMF add different kinds of sparsity constraints and obtain better results. This may because sparse constraints is more effective for unmixing problem, and it can detect expressive endmembers [54]. However, these methods often have poor RMSE performance since they only focus on endmembers. The utilization of spatial information solves this problem to a certain degree. Neighbor-based TV-RSNMF and deep NMF with TV SDNMF-TV both have great performance, and SDNMF-TV is slightly better than the other compared methods. It can be seen from Table III that our proposed method that learns spatial information from original images rather than design manually achieve better performance for real-world HU. In general, our proposed method achieves the lowest mean SAD values as well as the lowest mean RMSE compared with the other methods. This validates the superiority of the proposed subspace regularizer.

The qualitative unmixing results are shown in Figs. 7 and 8. From Fig. 7, we can see that the endmember signatures extracted by our method is almost coincident with the reference signatures obtained from the spectral library. Fig. 8 displays the abundance map obtained by our method. The corresponding endmember is illustrated with dark pixels. From Fig. 8, we can see the results quite agree with the four targets, "Water," "Soil," "Road," and "Tree," respectively.

Simultaneously, we obtained the clustering results. Since our method can jointly learn the subspace structure of the dataset, then the clustering result can be obtained by a standard spectral clustering algorithm. Fig. 9 shows the results when the number of clusters is set as 2, 3, and 4, respectively. It can be seen that the clustering results conform to the real image intuitively.

2) HYDICE Urban Dataset: HYDICE Urban dataset is another widely used HSI. It includes 307×307 pixels, and each pixel has 210 spectral bands ranging from 400–2500 nm. Here, noisy bands (1–4) and water-absorption (76, 87, 101–111, 136–153, and 198–210) bands were removed, resulting in an image of 162 bands. The groundtruth includes six endmembers: "Asphalt," "Tree," "Grass," "Roof#1," "Roof#2," and "Concrete road" as shown in Fig 6(b).

Similar to the previous experiment, Table IV shows the mean SAD and RMSE values. We observe that the proposed subspace learning regularized NMF method outperforms all the other methods. In this experiment, TV-RSNMF and SDNMF-TV that use spatial information obtain slightly better results than sparse-based methods. This indicates the effectiveness of spatial relationships for complex image unmixing.

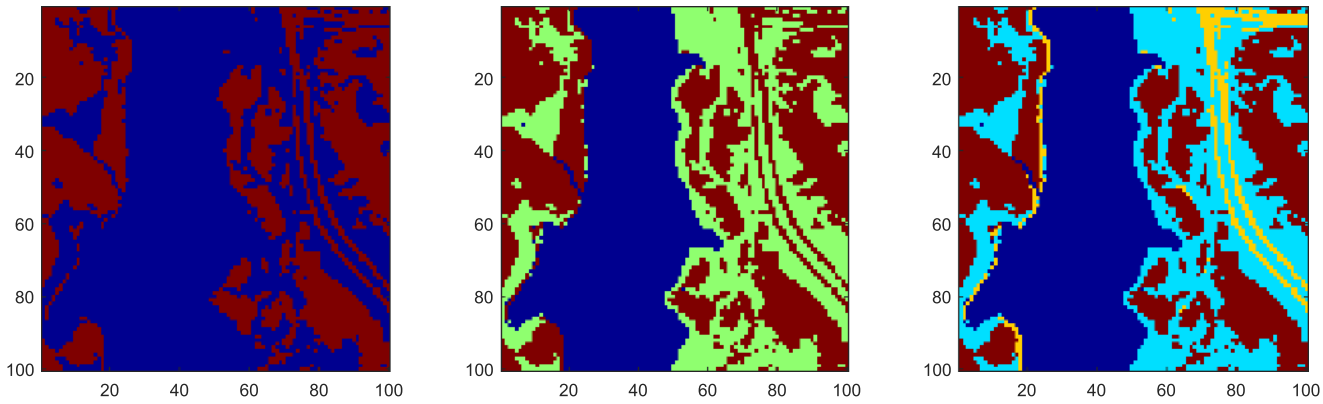

Fig. 9. Clustering results on Jasper Ridge dataset when the number of clusters is set as 2, 3, and 4, respectively.

TABLE IV
SAD AND RMSE VALUES AND RUNNING TIMES OF DIFFERENT METHODS ON THE URBAN DATASET

	VCA-FCLS	NMF	$L_{1/2}$-NMF	GLNMF	MLNMF	SGSNMF	TV-RSNMF	SDNMF-TV	Ours
SAD	0.3966	0.3721	0.2674	0.3129	0.2431	0.2410	0.2559	0.2375	**0.2307**
RMSE	0.2764	0.2417	0.1726	0.2274	0.1706	0.1744	0.1904	0.1642	**0.1593**
Time(s)	**124.3**	253.2	341.8	352.6	895.2	203.4	318.3	975.3	289.4

TABLE V
SAD VALUES AND RUNNING TIMES OF DIFFERENT METHODS ON THE CUPRITE DATASET

Method	VCA-FCLS	NMF	$L_{1/2}$-NMF	GLNMF	MLNMF	SGSNMF	TV-RSNMF	SDNMF-TV	Ours
Alunite	0.1094	0.1164	0.1245	0.1090	**0.0958**	0.1072	0.1032	0.0980	0.1099
Andradite	**0.0568**	0.0806	0.0732	0.0665	0.0743	0.1062	0.0810	0.0673	0.0711
Buddingtonite	0.1215	0.3762	0.1173	0.1130	0.1319	0.1197	0.1141	0.1141	**0.1059**
Dumortierite	**0.0759**	0.1216	0.0974	0.0798	0.0856	0.0761	0.0997	0.1015	0.0959
Kaolinite-1	0.0985	0.1072	0.1356	0.0994	0.0991	0.0778	**0.0771**	0.1075	0.0963
Kaolinite-2	0.0603	0.0901	0.0549	0.0624	0.0775	0.0844	**0.0489**	0.0753	0.0740
Muscovite	0.2130	0.2739	0.1443	0.1002	**0.0745**	0.1498	0.1427	0.0989	0.1306
Montmorillonite	0.0983	0.0922	**0.0535**	0.1030	0.0921	0.0595	0.0599	0.0800	0.0616
Nontronite	0.0733	0.4640	0.0744	0.0733	0.1177	0.1254	**0.0702**	0.1277	0.0784
Pyrope	0.1711	0.1817	0.0931	0.2455	0.1045	**0.0605**	0.0705	0.1118	0.0610
Sphene	0.0577	0.1155	0.2618	0.0552	**0.0508**	0.3159	0.2737	0.0642	0.1930
Chalcedony	0.0992	0.0992	**0.0698**	0.1450	0.1387	0.0980	0.1207	0.1477	0.0988
Mean	0.1022	0.1765	0.1083	0.1044	0.1001	0.1147	0.1055	0.0995	**0.0981**
Time(s)	**53.2**	102.1	153.5	169.3	410.2	93.5	174.5	493.5	142.8

3) AVIRIS Cuprite Dataset: Cuprite dataset contains 224 spectral bands cover the range of 400–2500 nm. A total of 188 bands remained by removing noisy bands (1–2 and 221–224) and water-vapor absorption bands (104–113 and 148–167). In this experiment, a spatial size of 250×191 was tailored, which contains 14 kinds of minerals [16]. Since there are only tiny differences between signatures of several minerals, the estimated number of endmembers was reduced to 12 for the unmixing. It is shown in Fig 6(c).

Table V compares the SAD results of different HU methods. We use bold to indicate the best and underline for the second best performance for each endmember. As shown in Table V, our method outperforms the compared methods for the mean SAD values. Different methods are good at estimating different endmembers, this might be because most of the endmembers in this dataset are tiny and fragmented, the spatial structure is not so obvious and unified. For endmembers like "Buddingtonite," "pyrope," and "Chalcedony," our proposed method has great advantages. Since the Cuprite dataset has no groundtruth, we only show the grayscale abundance maps obtained by our method in Fig. 10. Compared to the original image shown in Fig 6(c), the results can be verified intuitively.

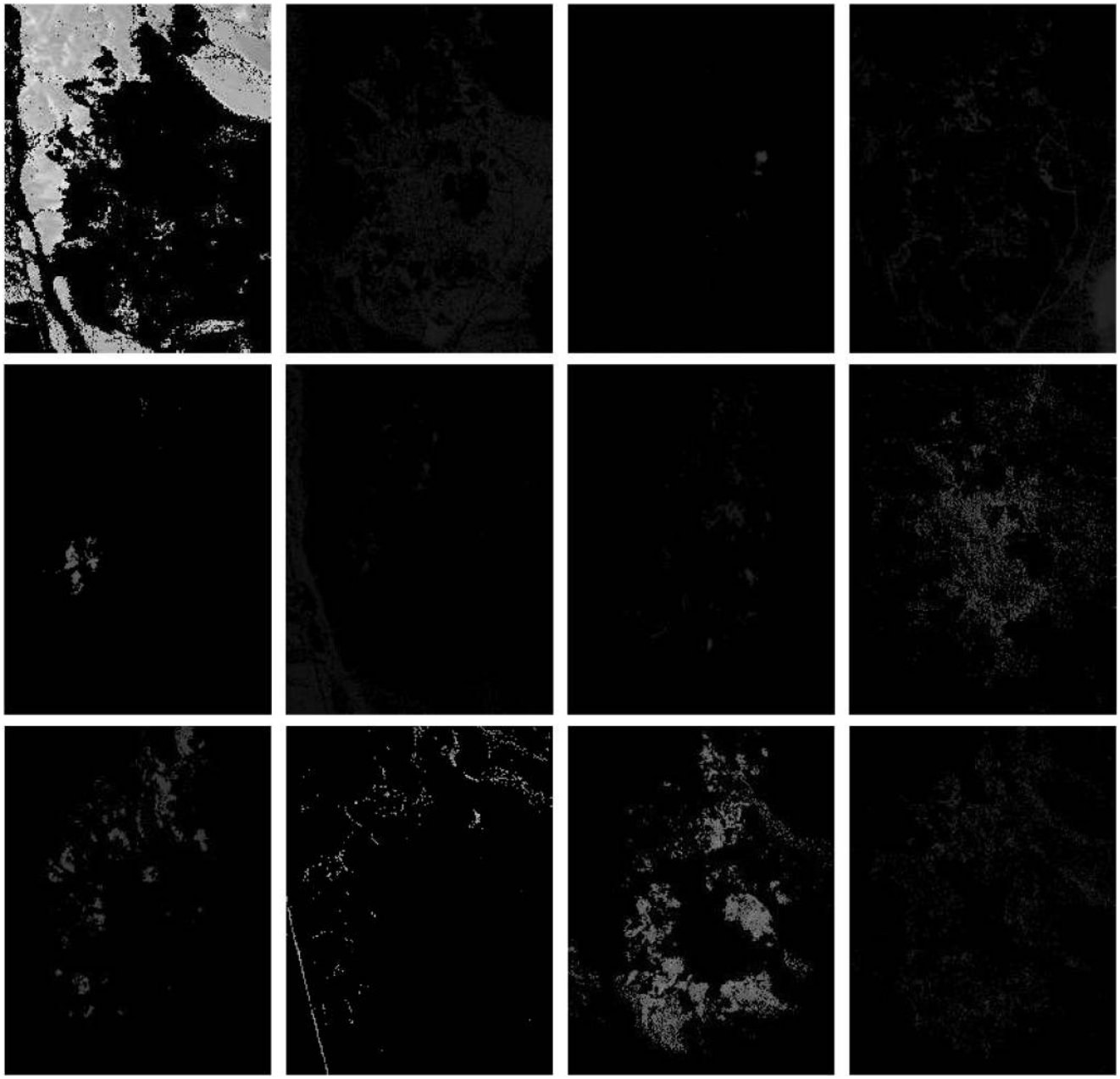

Fig. 10. Abundance maps of 12 different endmembers obtained using our method on the Cuprite dataset. From left to right and from top to bottom are Sphene, Andradite, Muscovite, Montmorillonite, Buddingtonite, Kaolinite-2, Alunite, Dumortierite, Kaolinite-1, Pyrope, Chalcedony, and Nontronite, respectively.

V. Conclusion

In this article, we have proposed a spatial information-based NMF by learning the subspace structure from the original image for blind HU. The presented model effectively exploits the subspace structure of the abundance map to constrain the NMF method. We first incorporate the subspace structure regularizer into the sparse NMF model as a spatial prior to improve the unmixing performance. The learned subspace structure can capture the global distribution of materials in different image regions. Then, we have integrated the spectral-spatial-based unmixing and subspace structure learning in a single unified framework and presented a multiplicative iterative method to optimize it. We compared our method with plenty classical and state-of-the-art NMF-based HU methods on both simulated and real-word HSI datasets. Both quantitative and qualitative results demonstrate the effectiveness of our method.

References

[1] X. Bai, H. Zhang, and J. Zhou, "VHR object detection based on structural feature extraction and query expansion," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6508–6520, Oct. 2014.

[2] J. Liang, J. Zhou, Y. Qian, L. Wen, X. Bai, and Y. Gao, "On the sampling strategy for evaluation of spectral-spatial methods in hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 862–880, Feb. 2017.

[3] S. Mei, J. Hou, J. Chen, L.-P. Chau, and Q. Du, "Simultaneous spatial and spectral low-rank representation of hyperspectral images for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2872–2886, May 2018.

[4] M. Zhang, W. Li, and Q. Du, "Diverse region-based CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2623–2634, Jun. 2018.

[5] X. Bai, F. Xu, L. Zhou, Y. Xing, L. Bai, and J. Zhou, "Nonlocal similarity based nonnegative tucker decomposition for hyperspectral image denoising," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 701–712, Mar. 2018.

[6] J. Li, Y. Li, R. Song, S. Mei, and Q. Du, "Local spectral similarity preserving regularized robust sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7756–7769, Oct. 2019.

[7] D. Wang, Z. Shi, and X. Cui, "Robust sparse unmixing for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1348–1359, Mar. 2018.

[8] Y. E. Salehani, S. Gazor, and M. Cheriet, "Sparse hyperspectral unmixing via heuristic l_p-norm approach," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1191–1202, Apr. 2018.

[9] L. Tong, J. Zhou, Y. Qian, X. Bai, and Y. Gao, "Nonnegative-matrix-factorization-based hyperspectral unmixing with partially known endmembers," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6531–6544, Nov. 2016.

[10] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1844–1868, Jun. 2014.

[11] B. Yang, B. Wang, and Z. Wu, "Nonlinear hyperspectral unmixing based on geometric characteristics of bilinear mixture models," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 1–21, Feb. 2018.

[12] F. J. Garca-Haro, M. A. Gilabert, and J. Meli, "Linear spectral mixture modelling to estimate vegetation amount from optical spectral data," *Int. J. Remote Sens.*, vol. 17, no. 17, pp. 3373–3400, 1996.

[13] J. W. Boardman, "Automating spectral unmixing of AVIRIS data using convex geometry concepts," in *Proc. Ann. JPL Airborne Geoscience Workshop* 1993, vol. 1, pp. 11–14.

[14] M. E. Winter, "N-FINDr: An algorithm for fast autonomous spectral endmember determination in hyperspecral data," *Proc. SPIE Int. Soc. Opt. Eng.*, vol. 3753, pp. 266–275, 1999.

[15] A. Zymnis, S. J. Kim, J. Skaf, M. Parente, and S. Boyd, "Hyperspectral image unmixing via alternating projected subgradients," in *Proc. Asilomar Conf.*, 2008, pp. 1164–1168.

[16] 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.

[17] J. Wang and C. I. Chang, "Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2601–2616, Sep. 2006.

[18] M. Craig, "Minimum volume transforms for remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 542–552, May 1994.

[19] J. Huang, T.-Z. Huang, L.-J. Deng, and X.-L. Zhao, "Joint-sparse-blocks and low-rank representation for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2419–2438, Apr. 2019.

[20] 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.

[21] F. Khajehrayeni and H. Ghassemian, "Hyperspectral unmixing using deep convolutional autoencoders in a supervised scenario," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 567–576, Feb. 2020.

[22] R. A. Borsoi, T. Imbiriba, and J. C. M. Bermudez, "Deep generative endmember modeling: An application to unsupervised spectral unmixing," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 374–384, 2020.

[23] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[24] D. Lee and S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 556–562.

[25] A. Cichocki, R. Zdunek, A. H. Phan, and S. I. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. Hoboken, NJ, USA: Wiley, 2009.

[26] P. Hoyer, "Non-negative matrix factorization with sparseness constraints," *J. Mach. Learn. Res.*, vol. 5, no. 1, pp. 1457–1469, 2004.

[27] P. O. Hoyer, "Non-negative sparse coding," in *Proc. 12th IEEE Workshop Neural Netw. Signal Process.*, 2002, pp. 557–565.

[28] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via $l_{1/2}$ sparsity-constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4282–4297, Nov. 2011.

[29] Y. Ma, C. Li, X. Mei, C. Liu, and J. Ma, "Robust sparse hyperspectral unmixing with $l_{2,1}$ norm," *Remote Sens.*, vol. 55, no. 3, pp. 1227–1239, Mar. 2017.

[30] D. Kong, C. Ding, and H. Huang, "Robust nonnegative matrix factorization using $l_{2,1}$-norm," in *Proc. 20th ACM Int. Conf. Inf. Knowl. Manage.*, 2011, pp. 673–682.

[31] W. He, H. Zhang, and L. Zhang, "Sparsity-regularized robust non-negative matrix factorization for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4267–4279, Sep. 2016.

[32] R. Huang, X. Li, and L. Zhao, "Spectral–spatial robust nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 8235–8254, Oct. 2019.

[33] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.

[34] W. He, H. Zhang, and L. Zhang, "Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3909–3921, Jul. 2017.

[35] X.-R. Feng, H.-C. Li, J. Li, Q. Du, A. Plaza, and W. J. Emery, "Hyperspectral unmixing using sparsity-constrained deep nonnegative matrix factorization with total variation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6245–6257, Oct. 2018.

[36] X. Liu, X. Wei, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 757–772, Feb. 2011.

[37] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li, "Manifold regularized sparse NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2815–2826, May 2013.

[38] W. Dong, X. Li, L. Zhang, and G. Shi, "Sparsity-based image denoising via dictionary learning and structural clustering," in *Proc. IEEE CVPR*, 2011, pp. 457–464.

[39] X. Lu, H. Wu, and Y. Yuan, "Double constrained NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2746–2758, May 2014.

[40] S. Khoshsokhan, R. Rajabi, and H. Zayyani, "Clustered multitask nonnegative matrix factorization for spectral unmixing of hyperspectral data," *J. Appl. Remote Sens.*, vol. 13, no. 2, 2019, Art. no. 026509.

[41] X. Wang, Y. Zhong, L. Zhang, and Y. Xu, "Spatial group sparsity regularized nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6287–6304, Nov. 2017.

[42] H. Zhang, Z. Han, L. Zhang, and P. Li, "Spectral spatial sparse subspace clustering for hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3672–3684, Jun. 2016.

[43] L. Zhou, X. Bai, X. Liu, J. Zhou, and E. R. Hancock, "Learning binary code for fast nearest subspace search," *Pattern Recognit.*, vol. 98, 2020, Art. no. 107040.

[44] P. Ji, T. Zhang, H. Li, M. Salzmann, and I. Reid, "Deep subspace clustering networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 24–33.

[45] L. Zhou *et al.*, "Latent distribution preserving deep subspace clustering," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, 2019, pp. 4440–4446.

[46] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in *Proc. 27th Int. Conf. Mach. Learn.*, 2010, pp. 663–670.

[47] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.

[48] N. Keshava, "A survey of spectral unmixing algorithms," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 55–78, 2003.

[49] F. Xiong, J. Zhou, and Y. Qian, "Material based object tracking in hyperspectral videos," *IEEE Trans. Image Process.*, vol. 29, no. 1, pp. 3719–3733, 2020.

[50] J. F. Cai, E. J. Cands, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *Siam J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2008.

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

[52] R. Rajabi and H. Ghassemian, "Spectral unmixing of hyperspectral imagery using multilayer NMF," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 38–42, Jan. 2015.

[53] S. Grupo de Inteligencia Computacional, Universidad del Pas Vasco / Euskal Herriko Unibertsitatea (UPV/EHU). "Hyperspectral Imagery Synthesis (EIAs) Toolbox." [Online]. Available: http://www.ehu.es/ccwintco/index.php/Hyperspectral Imagery Synthesis tools for MATLAB

[54] S. Z. Li, X. W. Hou, H. J. Zhang, and Q. S. Cheng, "Learning spatially localized, parts-based representation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, 2001, vol. 1, pp. I–I.

Lei Zhou received the bachelor's degree, in 2016 from the School of Mathematics and Systems Science, Beihang University, Beijing, China, where he is currently working toward the Ph.D. degree with the School of Computer Science and Engineering.

His current research interests include machine learning, computer vision, and remote sensing image processing.

Xueni Zhang received the bachelor's degree, in 2017 from the College of Computer Science and Technology, Jilin University, Jilin, China, and received the master's of engineering degree, in 2020 from the School of Computer Science and Engineering, Beihang University, Beijing, China.

Her research interests include computer vision and remote sensing image processing.

Jianbo Wang is currently an undergraduate from the First Clinical Medical College of Nanchang University, Jiangxi, China.

His research interests include machine learning and image processing.

Xiao Bai received the B.Eng. degree in computer science from Beihang University, Beijing, China, in 2001, and the Ph.D. degree in computer science from the University of York, York, U.K., in 2006.

He was a Research Officer (Fellow, Scientist) with the Computer Science Department, University of Bath, Bath, U.K. until 2008. He is currently a Full Professor with the School of Computer Science and Engineering, Beihang University. He has authored or coauthored more than 60 papers in journals and refereed conferences. His current research interests include pattern recognition, image processing, and remote sensing image analysis.

Lei Tong received the B.E. degree in measurement and control technology and instrumentation, the M.E. degree in measurement technology and automation devices from Beijing Jiaotong University, Beijing, China, in 2010 and 2012, respectively, and the Ph.D. degree in engineering from Griffith University, Brisbane, Australia, in 2016.

Currently, he is a Lecturer with the Faculty of Information Technology, Beijing University of Technology, Beijing, China. His current research interests include signal and image processing, pattern recognition, and remote sensing.

Liang Zhang received the B.Eng. degree in computer science and the Ph.D. degree in computer science from Beihang University, Beijing, China, in 2001 and 2007, respectively.

His current research interests include machine learning, computer vision, and image processing.

Jun Zhou received the B.S. degree in computer science and the B.E. degree in international business from the Nanjing University of Science and Technology, Nanjing, China, in 1996 and 1998, respectively; the M.S. degree in computer science from Concordia University, Montreal, QC, Canada, in 2002; and the Ph.D. degree in computing science from the University of Alberta, Edmonton, AB, Canada, in 2006.

He was a Research Fellow with the Research School of Computer Science, Australian National University, Canberra, ACT, Australia, and a Researcher with the Canberra Research Laboratory, NICTA, Canberra, ACT, Australia. In June 2012, he joined the School of Information and Communication Technology, Griffith University, Nathan, QLD, Australia, where he is currently an Associate Professor. His research interests include pattern recognition, computer vision, and spectral imaging with their applications in remote sensing and environmental informatics.

Edwin Hancock (Fellow, IEEE) received the B.Sc. degree in physics, in 1977, the Ph.D. degree in high-energy physics, in 1981, and the D.Sc. degree, in 2008 from the University of Durham, Durham, U.K., and a doctorate Honoris Causa from the University of Alicante, Alicante, Spain, in 2015.

He is a Professor with the Department of Computer Science, where he leads a group of some faculty, research staff, and Ph.D. students working in the areas of computer vision and pattern recognition. His main research interests are in the use of optimization and probabilistic methods for high and intermediate level vision.

He is a Fellow of the International Association for Pattern Recognition. He is currently the Editor-in-Chief of the journal *Pattern Recognition*, and was the Founding Editor-in-Chief of *IET Computer Vision* from 2006 until 2012. He has also been a member of the editorial boards of the journals IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, *Pattern Recognition*, *Computer Vision and Image Understanding*, *Image and Vision Computing*, and the *International Journal of Complex Networks*.