

Sparsity-Enhanced Convolutional Decomposition: A Novel Tensor-Based Paradigm for Blind Hyperspectral Unmixing

Jing Yao¹, Danfeng Hong², *Member, IEEE*, Lin Xu³, Deyu Meng⁴, *Member, IEEE*,
Jocelyn Chanussot⁵, *Fellow, IEEE*, and Zongben Xu, *Member, IEEE*

Abstract—Blind hyperspectral unmixing (HU) has long been recognized as a crucial component in analyzing the hyperspectral imagery (HSI) collected by airborne and spaceborne sensors. Due to the highly ill-posed problems of such a blind source separation scheme and the effects of spectral variability in hyperspectral imaging, the ability to accurately and effectively unmixing the complex HSI still remains limited. To this end, this article presents a novel blind HU model, called sparsity-enhanced convolutional decomposition (SeCoDe), by jointly capturing spatial-spectral information of HSI in a tensor-based fashion. SeCoDe benefits from two perspectives. On the one hand, the convolutional operation is employed in SeCoDe to locally model the spatial relation between the targeted pixel and its neighbors, which can be well explained by spectral bundles that are capable of addressing spectral variabilities effectively. It maintains, on the other hand, physically continuous spectral components by decomposing the HSI along with the spectral domain. With sparsity-enhanced regularization, an alternative optimization strategy with alternating direction method of multipliers (ADMM)-based optimization algorithm is devised for efficient model inference. Extensive experiments conducted on three different data sets demonstrate the superiority of the proposed SeCoDe compared to previous state-of-the-art methods. We will also release the code at https://github.com/danfenghong/IEEE_TGRS_SeCoDe to encourage the reproduction of the given results.

Index Terms—Blind hyperspectral unmixing (HU), convolutional sparse coding (CSC), spectral bundles, spectral variability (SV), tensor decomposition.

Manuscript received November 5, 2020; revised February 14, 2021 and March 18, 2021; accepted March 26, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant U1811461, Grant 11690011, and Grant 61721002; in part by the MIAI@Grenoble Alpes under Grant ANR-19-P3IA-0003; and in part by the AXA Research Fund. (*Corresponding author: Danfeng Hong.*)

Jing Yao, Deyu Meng, and Zongben Xu are with the School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: jasonyao92@gmail.com; dymeng@mail.xjtu.edu.cn; zbxu@mail.xjtu.edu.cn).

Danfeng Hong is with Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Weßling, Germany, and also with the University of Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France (e-mail: danfeng.hong@dlr.de).

Lin Xu is with the Institute of Artificial Intelligence, Shanghai Em-Data Technology Company, Ltd., Shanghai 200000, China (e-mail: lin.xu5470@gmail.com).

Jocelyn Chanussot is with the GIPSA-lab Grenoble Alpes, INRIA, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France, and also with Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China (e-mail: jocelyn@hi.is).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TGRS.2021.3069845>.

Digital Object Identifier 10.1109/TGRS.2021.3069845

I. INTRODUCTION

AIRBORNE or spaceborne imaging spectroscopy, also known as hyperspectral imaging, plays a fundamental role in earth observation and remote sensing. Characterized by abundant spectral information [1], hyperspectral imagery (HSI) has been favored by researches upon various applications, such as dimensionality reduction [2]–[4], land-cover and land-use classification, object detection [5]–[10], spectral unmixing [11]–[13], image segmentation [14], and multisource data fusion [15]–[17]. Nevertheless, due to the limitations of spatial resolution in hyperspectral imaging, the pixels in HSIs are usually in the form of mixtures, inevitably degrading the performance of the subsequent high-level analysis. Such a challenge motivates the constant development of spectral unmixing, which aims at decomposing the mixed spectrum into a collection of spectral signatures of pure materials (or say, *endmembers*) and their fractional abundances (or say, *abundance maps*).

A wealth of hyperspectral unmixing (HU) methods has emerged apace with the advances in statistically modeling and machine learning within the past few decades. The focus of this context starts from a relatively simplistic yet practical family of the model, i.e., the linear mixing model (LMM) [18]. Assuming that microscopic interaction between distinct materials is negligible, LMM basically seeks a linear decomposition to approximate the mixing process of the collected hyperspectral data. This kind of HU method largely depends on the endmember extraction algorithms. Except for the popular vertex component analysis (VCA) [19], Zhang *et al.* [20] for the first time proposed to use the ant colony optimization (ACO) to intelligently extract the endmembers of HSIs, which has been recently recognized as a pioneering work in the HU field and achieved the state-of-the-art unmixing results in many real hyperspectral scenes. The majority of recently proposed LMM-based methods rely on nonnegative matrix factorization (NMF) due to its good interpretability and efficient solution [21]. Besides, several kinds of regularization terms that are capable of encoding certain *prior* knowledge have been successfully employed from structure sparsity to spatial smoothness [22], [23]. Despite so, the effectiveness of these linearized blind HU methods remains limited, particularly in complex real scenarios where there exist various *spectral variabilities*.

The spectral variability (SV) refers to the phenomenon that the spectral reflectance of a given endmember can vary spatially or spectrally, due to multiple intrinsic and exterior factors (e.g., atmospheric effects, illumination conditions, geometry, and topography) [24]. To eliminate the negative effects on HU performance caused by such unpredictable variations, numerous efforts have been made to develop either stochastic or deterministic models (see [25], [26] for a recent review). Still, in the case of LMM, a representative category of studies resorts to *spectral bundles*, by representing each material within a known or estimated set of endmember “candidates,” to account for the variability [27]–[29]. More recently, physics-inspired models that extend LMM in a more constrained way to explicitly characterize SVs have received considerable attention [30], [31]. Albeit verified to be effective to some extent, these methods potentially suffer from the mismodeling effects by their vector-/matrix-based implementations.

Considering the highly ill-posed problem of blind HU, the loss of spatial correlation information brought by roughly stacking the spectrum of each mixed pixel is not tolerable. Fortunately, tensorized formulation has shown great potential in providing a faithful representation for multidimensional data. Recently proposed tensor factorization-based methods for blind HU task, by modeling the observed HSI data cube as a third-order tensor, are capable of preserving the inherent spatial–spectral information and achieving a satisfying performance [32], [33]. However, two issues are worthy of our consideration. The first issue brings failure to characterize the SV. Second, there is still room in making full use of structural regularity to enhance unmixing performance.

To overcome the aforementioned issues, this article introduces the sparsity-enhanced convolutional decomposition (SeCoDe), providing a novel tensor-based solution for solving the blind HU task. To be specific, our contributions can be summarized as follows.

- 1) We propose a novel SeCoDe decomposition strategy for tensor-based blind HU, in which a two-layered sparsity regularization is designed to further exploit the structural information of abundance maps underlying HSIs.
- 2) To the best of our knowledge, this is the first time to investigate the blind HU issue in a convolutional representation manner. Different from the traditional sparse unmixing models that focus HU at the pixel level, SeCoDe can progressively decompose the HSI from local attention to global aggregation, which enables the full utilization of spatial contextual information by a convolution operation.
- 3) Beyond the classic bandwise convolutional sparse coding (CSC), a multichannel version (MC-CSC) is developed in our model to learn a continuous spectral representation, yielding a spectrally meaningful component decomposition. The endmember products obtained by the MC-CSC can be explained by spectral bundles, which is robustly against various SVs to a great extent.
- 4) Accordingly, we provide an effective solution based on an alternative optimization strategy (AOS) with

alternating direction method of multipliers (ADMM) for the proposed SeCoDe. Extensive experiments implemented on a series of synthetic and real-world hyperspectral data sets demonstrate the superiority of the proposed method, both visually and quantitatively.

The rest of this article is organized as follows. We first introduce some necessary notions and preliminaries in Section II. Then, representative related works are investigated in Section III. Section IV presents our methodology for the blind HU task. Extensive experimental results are reported and discussed in Section V. Finally, we give concluding remarks and future prospects in Section VI.

II. NOTATIONS AND PRELIMINARIES

The notations in this article are standard in signal processing. Throughout this article, scalars, vectors, and matrices are denoted as the nonbold case, bold lower case, and bold upper case letters, respectively. In addition, the symbol \circ denotes outer product of two vectors, i.e., $\mathbf{p} \circ \mathbf{q} = \mathbf{p}\mathbf{q}^\top$; \odot means Schur–Hadamard products; \mathbb{R} marks the real coordinate space with specific dimensions; \succeq represents elementwise inequality; $\mathbf{1}$ is all-one vector/matrix with appropriate size; $\widehat{(\cdot)}$ represents variable in the discrete Fourier transform (DFT) domain; and $\|\cdot\|_{1,1}$, $\|\cdot\|_F$, and $\|\cdot\|_2$ are the matrix $\ell_{1,1}$ -norm, the Frobenius norm, and the vector ℓ_2 -norm, respectively.

We denote an N -order tensor as $\mathcal{X} = (x_{i_1, \dots, i_n, \dots, i_N}) \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, where $i_n = 1, 2, \dots, I_n$. The mode- n vectors of \mathcal{X} are a series of I_n dimensional vectors extracted from \mathcal{X} by varying other indices. The mode- n unfolding matrix $\mathbf{X}_{(n)} \in \mathbb{R}^{I_n \times (I_1 \dots I_{n-1} I_{n+1} \dots I_N)}$ is obtained by stacking the mode- n vectors as columns. The mode- n product between \mathcal{X} and a matrix $\mathbf{P} \in \mathbb{R}^{J \times I_n}$ is denoted by $\mathcal{X} \times_n \mathbf{P}$, resulting in an N -order tensor $\mathcal{Z} \in \mathbb{R}^{I_1 \times \dots \times J \times \dots \times I_N}$ with its elements $z_{i_1, \dots, j, \dots, i_N} = \sum_{i_n} x_{i_1, \dots, i_n, \dots, i_N} P_{j i_n}$. A tensor is called rank-1 if it can be written as the outer product of N vectors as $\mathcal{X} = \mathbf{m}_{i_1} \circ \mathbf{m}_{i_2} \circ \dots \circ \mathbf{m}_{i_N}$, which is also called Kronecker basis. The ℓ_1 -norm and the Frobenius norm of \mathcal{X} are defined as $\|\mathcal{X}\|_1 = \sum_{i_1, \dots, i_N} |x_{i_1, \dots, i_N}|$ and $\|\mathcal{X}\|_F = (\sum_{i_1, \dots, i_N} |x_{i_1, \dots, i_N}|^2)^{1/2}$, respectively.

III. RELATED WORK

In this section, we first briefly introduce the background of the widely studied LMM. Next, we review several representative variants of LMM against SVs and existed tensor factorization-based methods for blind HU, respectively.

A. Conventional LMM

Let $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]^\top \in \mathbb{R}^{N \times L}$ be an L -spectrum measured HSI with $N = H \times W$ pixels, $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_K]^\top \in \mathbb{R}^{K \times L}$ be the mixing matrix of K endmembers, and the abundance maps $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N]^\top \in \mathbb{R}^{N \times K}$ associated with each endmember. The LMM assumes that, for the n th pixel, the spectral responses of distinct endmembers bear no interference with each other, which can be written as

$$\mathbf{y}_n = \mathbf{s}_n^\top \mathbf{A} + \mathbf{e}_n \quad (1)$$

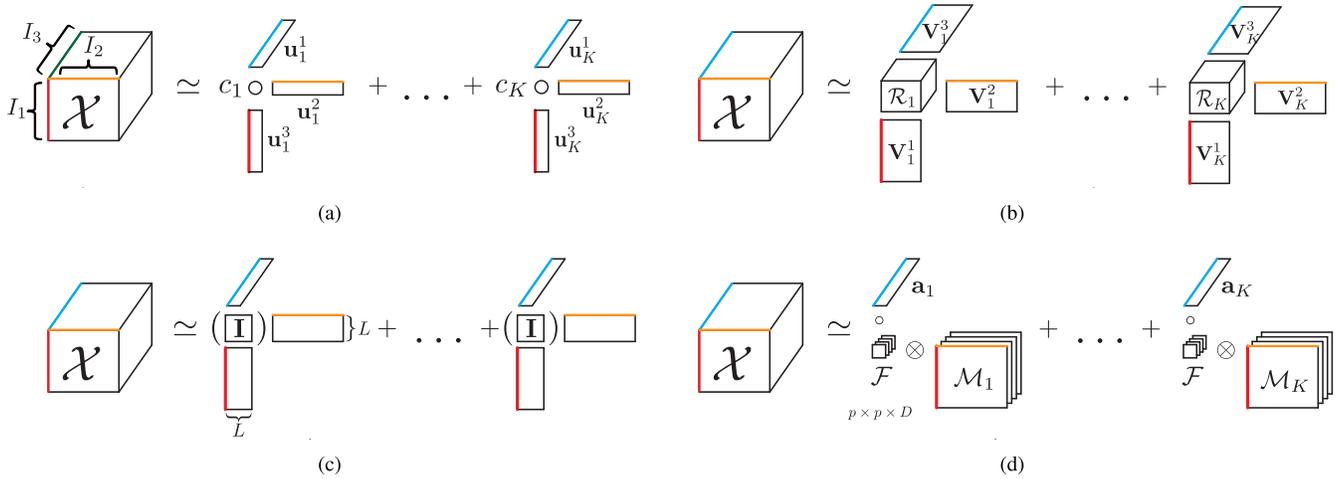


Fig. 1. Illustration to clarify the differences of four tensor-based decomposition strategies. (a)–(c) Three kinds of existed tensor decomposition [(a) CPD, (b) BTD, and (c) MVNTF]. (d) Proposed SeCoDe for HU.

where $\mathbf{e}_n \in \mathbb{R}^{1 \times L}$ represents the additive noise. To account for the physical conditions in reality, three reasonable constraints are commonly imposed, i.e., the sum-to-one constraint of \mathbf{s}_n (ASC), the nonnegativity constraint of \mathbf{s}_n (ANC), and \mathbf{a}_k . By collecting all pixels, we can rewrite the LMM in a compact matrix form as

$$\mathbf{Y} = \mathbf{S}\mathbf{A} + \mathbf{E}, \quad \text{s.t. } \mathbf{S}, \mathbf{A} \geq \mathbf{0}, \mathbf{S}\mathbf{1}_K = \mathbf{1}_N. \quad (2)$$

As only \mathbf{Y} is available in real scenarios, the above LMM is equivalent to the linear blind source separation (BSS) model in [34]. Given the number of endmembers K , a common solution is to first apply the VCA for extracting endmember signatures, and then, fully/partial constrained least-squares unmixing (FCLSU/PCLSU) [35], [36] and sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) [37] can be adopted to estimate the corresponding abundance maps.

B. LMM Addressing Spectral Variability

The main drawback of conventional LMM lies in its failure to explain the issue of SVs, which potentially hampers an accurate unmixing. For this reason, many LMM-based variants have been proposed for modeling SVs in the process of unmixing. Intuitively, the LMM and its variants can be well generalized into a unified framework as follows:

$$\mathbf{Y} = f(\mathbf{S}, \mathbf{A}) + \mathbf{E} \quad (3)$$

where f is a linear function with respect to \mathbf{S} and \mathbf{A} . According to the assumption that SVs are dominated by scaling factors [18], scaled PCLSU (SPCLSU) [38] and scaled SUnSAL (SSUnSAL) [39] first introduced a diagonal matrix $\mathbf{P}_1 \in \mathbb{R}^{N \times N}$ containing shared scaling factors across endmembers, resulting in $f(\mathbf{S}, \mathbf{A}) = \mathbf{P}_1\mathbf{S}\mathbf{A}$, to allow a pixelwise variation. Instead, extended LMM (ELMM) [40] suggested to adjust endmembers by multiplying a scaling factors to locally absorb contrast and illumination effects, i.e., $f(\mathbf{S}, \mathbf{A}) = \mathbf{P}_2 \odot \mathbf{S}\mathbf{A}$, where $\mathbf{P}_2 \in \mathbb{R}^{N \times K}$.

To overcome other kinds of SVs, Thouvenin *et al.* [41] proposed a perturbed LMM (PLMM) by adding a perturbation

term for each endmember. In this case, $f(\mathbf{S}, \mathbf{A}) = \mathbf{S}\mathbf{A} + [\mathbf{s}_1^\top \Delta_1; \mathbf{s}_2^\top \Delta_2; \dots; \mathbf{s}_N^\top \Delta_N]$, where the rows in $\Delta_n \in \mathbb{R}^{K \times L}$ are composed of perturbations of each endmember on the n th pixel. Furthermore, Hong *et al.* [31] extended ELMM to an augmented LMM (ALMM) by considering $f(\mathbf{S}, \mathbf{A}) = \mathbf{P}_1\mathbf{S}\mathbf{A} + \mathbf{S}'\mathbf{A}'$, where \mathbf{S}' and \mathbf{A}' are the SV dictionary and corresponding coefficients, respectively. An alternative is the subspace unmixing with low-rank attribute embedding (SULoRA) [30], as formulated by $f(\mathbf{S}, \mathbf{A}) = \mathbf{S}\mathbf{A}\mathbf{Q}$, where $\mathbf{Q} \in \mathbb{R}^{L \times L}$ is the low-rank subspace projection.

C. LMM Considering Tensor Factorization

Although enormous efforts have been made by the above-mentioned LMMs to address SVs in the unmixing process, these models inevitably suffer from the loss of structural information caused by their improper 2-D formulations. Recently, some tentative methods have been proposed to decompose the HSI from a tensor perspective in order to provide a more compact representation of hyperspectral data in both spatial and spectral domains. More specifically, instead of unfolding 3-D HSI into an observed matrix, tensor-based approaches perform unmixing by means of the tensor factorization technique directly acting on the original 3-D hyperspectral cube.

Zhang *et al.* [32] first introduced canonical polyadic decomposition (CPD) and applied it in identifying materials of a space object utilizing HSIs. Given a third-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, CPD seeks to find a linear combination of K rank-1 tensors

$$\mathcal{X} = \sum_{k=1}^K c_k \mathcal{U}_k = \sum_{k=1}^K c_k \mathbf{u}_k^1 \circ \mathbf{u}_k^2 \circ \mathbf{u}_k^3 \quad (4)$$

where $\mathbf{u}_k^n \in \mathbb{R}^{I_n \times 1}$ ($1 \leq n \leq 3$) denotes the factor vectors of so-called Kronecker basis \mathcal{U}_k and imposed coefficient c_k [42].

By following this, Qian *et al.* [33] designed the matrix-vector nonnegative tensor factorization (MVNTF) for HU by linking the LMM with tensor notation, which was further improved by pursuing local smoothness and nonlocal low-rankness priors [43], [44]. MVNTF can be regarded as

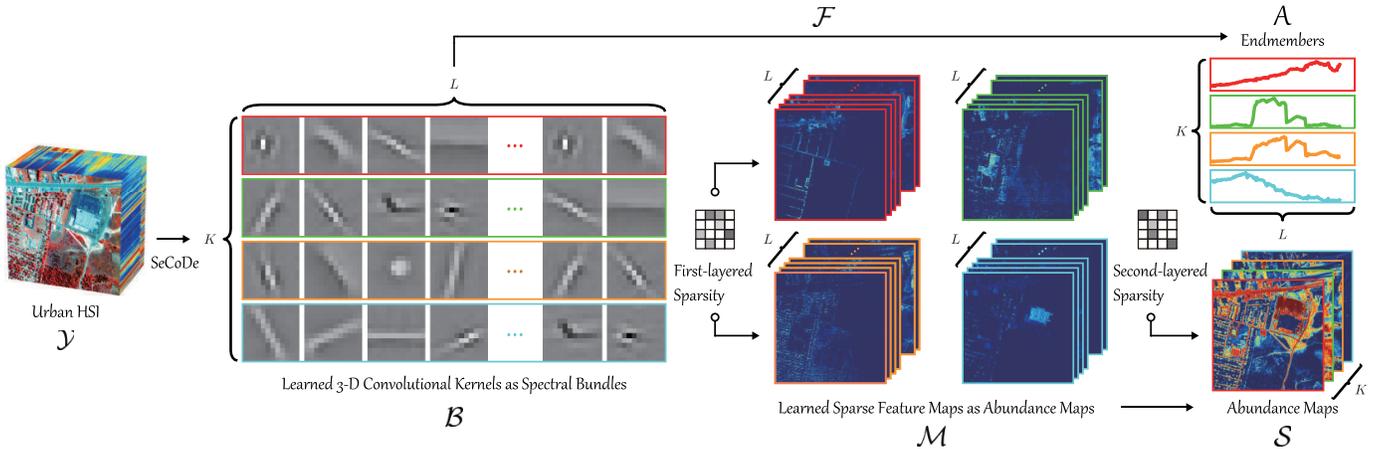


Fig. 2. Illustration to clarify the workflow of our proposed SeCoDe. Given the HSI \mathcal{Y} , the learned 3-D convolutional kernels \mathcal{B} can be regarded as spectral bundles against SVs, which can be further decomposed into K unique endmembers \mathbf{A} and a set of 2-D filters \mathcal{F} . Sparse feature maps \mathcal{M} corresponding to \mathcal{B} are generated by means of the first-layered sparsity, thereby integrating the final abundance maps \mathcal{S} with the second-layered sparsity.

a special case of block term decomposition (BTD) [45]. The BTD can be written as follows:

$$\mathcal{X} = \sum_{k=1}^K \mathcal{R}_k \times_1 \mathbf{V}_k^1 \times_2 \mathbf{V}_k^2 \times_3 \mathbf{V}_k^3 \quad (5)$$

where $\mathcal{R}_k \in \mathbb{R}^{R_1 \times R_2 \times R_3}$ is defined as the core tensor, and $\mathbf{V}_k^n \in \mathbb{R}^{I_n \times R_n}$ ($R_n \leq I_n, 1 \leq k \leq K, 1 \leq n \leq 3$) gathers R_n orthogonal bases along the n th mode of \mathcal{X} (or called *terms*). By appropriately setting $R_1 = R_2 = L, R_3 = 1, \mathcal{R}_k = \mathbf{I} \in \mathbb{R}^{L \times L \times 1}$ to be an identity matrix, and $\mathbf{v}_k^3 \in \mathbb{R}^{L \times 1}$ ($1 \leq k \leq K$), we then have the formulation of MVNTF

$$\mathcal{X} = \sum_{k=1}^K \mathbf{V}_k^1 \cdot (\mathbf{V}_k^2)^\top \circ \mathbf{v}_k^3 = \sum_{k=1}^K \mathbf{W}_k \circ \mathbf{v}_k^3 \quad (6)$$

where $\mathbf{W}_k \in \mathbb{R}^{L_1 \times L_2}$ denotes the product of first two factor matrices. It should be noted that both CPD and MVNTF can be concluded into the unified BTD framework with different ranks as rank-(1, 1, 1) and rank-($L, L, 1$) terms, respectively.

IV. PROPOSED SECoDE MODEL AND SOLVING ALGORITHM

A. Method Overview

To effectively address SVs and simultaneously preserve spatial-spectral structure information in HU, we seek to directly decompose the 3-D hyperspectral tensor into a series of convolutional kernels and corresponding feature maps in a deconvolutional manner¹ [see the example in Fig. 1(d)], where these learned 3-D convolutions can be well explained by spectral bundles and feature maps can be accordingly seen as abundance maps. Moreover, the 3-D convolutional kernels can be further decomposed into to-be-extracted pure endmembers that preserve the approximately continuous spectral properties and 2-D convolutional kernels that are capable of effectively modeling spatial relation robustly against SVs. In addition,

¹Unlike existing deconvolution-based unmixing methods that mostly only use a single filter to characterize the blur degradation [46], we here propose to adaptively learn overcomplete convolutional kernels to capture the more detailed spatial structure and directly represent the HSIs.

our SeCoDe model is further improved in unmixing performance by adaptively updating the convolutional dictionary and designing a two-layered sparsity-enhanced regularization. Fig. 2 illustrates the workflow of the proposed method.

B. Model Formulation

Before stepping into the proposed SeCoDe, we start with a brief introduction of MVNTF because the MVNTF model is not only the cornerstone of our method but also holding clearly mathematical interpretation, as well as easy implementation. Given the observed HSI $\mathcal{Y} \in \mathbb{R}^{H \times W \times L}$ and its abundances collected in tensor form $\mathcal{S} \in \mathbb{R}^{H \times W \times K}$, we can rewrite the degradation model from (2) under MVNTF framework as

$$\mathcal{Y} = \sum_{k=1}^K \mathbf{S}^{(k)} \circ \mathbf{a}_k + \mathcal{E}, \quad \text{s.t. } \mathcal{S}, \mathbf{A} \geq \mathbf{0}, \sum_{k=1}^K \mathbf{S}^{(k)} = \mathbf{1} \quad (7)$$

where $\mathbf{S}^{(k)} \in \mathbb{R}^{H \times W}$, i.e., the frontal slice of \mathcal{S} , denotes the abundance map corresponding to the k th endmember. A straightforward solution is turning (7) into a matrix-based minimization problem of mean square error by mode-3 unfolding. Alternatively, an alternating least-squares (ALS) algorithm was given in [33] by iteratively solving matrix factorization problems to each mode unfolded matrix of \mathcal{Y} . However, the factors along spatial modes with enforced low-rankness and nonnegativity properties still lack intuitive interpretations.

MVNTF is a good tool to unmix the 3-D hyperspectral cube globally, yet, it, to some extent, fails to consider an important prior assumption, that is, the spectral signature in a target pixel should share the same or similar abundances with its locally spatial neighbors. This might further lead to vulnerability to SVs. Different from the direct matrix factorization along with each mode in optimizing MVNTF, CSC is capable of capturing the information from locally to globally in a progressive way [47] to alleviate the effects of SVs by learning spectral bundles-like convolutional kernels. However, due to its bandwise reconstruction strategy, the effectiveness of directly applying the traditional CSC on the HSI remains limited, yielding the violation of continuous spectral structure.

Therefore, the well-designed MC-CSC has been applied to model spectral information in our SeCoDe continuously, i.e.,

$$\mathcal{S} = \sum_{d=1}^D \mathbf{F}_d \otimes \mathcal{M}_d \quad (8)$$

where D is the size of filter bank $\mathcal{F} = \{\mathbf{F}_d\}_{d=1}^D$, and the third-order tensor $\mathcal{M}_d \in \mathbb{R}^{H \times W \times K}$ is constructed by stacking K feature maps $\{\mathbf{M}_d^{(k)}\}_{k=1}^K$ corresponding to the d th filter $\mathbf{F}_d \in \mathbb{R}^{P \times P}$ along the third dimension. By plugging (8) into (7), the 3-D HSI \mathcal{Y} can be reconstructed in a tensor-based fashion, as illustrated in Fig. 1(d).

Note that, for the abundance map of each endmember, we have $\mathbf{S}^{(k)} = \sum_{d=1}^D \mathbf{F}_d \otimes \mathbf{M}_d^{(k)}$. Let the linear operator $\tilde{\mathbf{M}}_d^{(k)} \in \mathbb{R}^{H \times W \times P^2}$ satisfy $\tilde{\mathbf{M}}_d^{(k)} \text{vec}(\mathbf{F}_d) = \text{vec}(\mathbf{F}_d \otimes \mathbf{M}_d^{(k)})$, where $\text{vec}(\cdot)$ denotes the common vectorization. Therefore, the mode-3 unfolding matrix of the hyperspectral cube \mathcal{Y} , i.e., $\mathbf{Y}_{(3)} \in \mathbb{R}^{L \times N}$, can be approximated by

$$\begin{aligned} \mathbf{Y}_{(3)}^\top &\approx \sum_{k=1}^K \left(\sum_{d=1}^D \tilde{\mathbf{M}}_d^{(k)} \text{vec}(\mathbf{F}_d) \right) \mathbf{a}_k^\top \\ &= \sum_{k=1}^K \sum_{d=1}^D \tilde{\mathbf{M}}_d^{(k)} \mathbf{B}_d^{(k)} \end{aligned} \quad (9)$$

where \mathcal{B} represents 3-D convolutional kernels that consist of $\mathbf{B}_d^{(k)} = \text{vec}(\mathbf{F}_d) \mathbf{a}_k^\top \in \mathbb{R}^{P^2 \times L}$, $k = 1, \dots, K$, and P^2 stands for the number of pixels contained in each filter \mathbf{F}_d . With this form of representation, $\mathbf{B}_d^{(k)}$ and $\tilde{\mathbf{M}}_d^{(k)}$ clearly exhibit same physical meaning as spectral bundles and corresponding fractional maps. The complex SVs are, therefore, expected to be absorbed by such a linear combination. Furthermore, we have the following tensor-based decomposition model with the given filter bank \mathcal{F} :

$$\min_{\mathbf{A}, \mathcal{M}} \frac{1}{2} \|\mathbf{Y}_{(3)}^\top - \sum_{k=1}^K \left(\sum_{d=1}^D \tilde{\mathbf{M}}_d^{(k)} \text{vec}(\mathbf{F}_d) \right) \mathbf{a}_k^\top\|_F^2. \quad (10)$$

To integrate spectral bundle into unique endmember for each material and reduce the model complexity, we relax the model in (10) into two terms by

$$\min_{\Theta} \frac{1}{2} \left\| \mathcal{Y} - \sum_{k=1}^K \mathbf{S}^{(k)} \circ \mathbf{a}_k \right\|_F^2 + \frac{\alpha}{2} \left\| \mathcal{S} - \sum_{d=1}^D \mathbf{F}_d \otimes \mathcal{M}_d \right\|_F^2 \quad (11)$$

where $\Theta = \{\mathbf{A}, \mathcal{M}, \mathcal{S}\}$ denotes the collection of to-be-estimated variables. With commonly used priors, e.g., sparsity-promoting term in the form of ℓ_1 -norm, and necessary constraints, our baseline in (11) can be further written as

$$\begin{aligned} \min_{\Theta} \frac{1}{2} \left\| \mathcal{Y} - \sum_{k=1}^K \mathbf{S}^{(k)} \circ \mathbf{a}_k \right\|_F^2 + \frac{\alpha}{2} \left\| \mathcal{S} - \sum_{d=1}^D \mathbf{F}_d \otimes \mathcal{M}_d \right\|_F^2 \\ + \beta \sum_{d=1}^D \|\mathcal{M}_d\|_1 \\ \text{s.t. } \mathcal{S}, \mathbf{A} \geq \mathbf{0}, \sum_{k=1}^K \mathbf{S}^{(k)} = \mathbf{1} \end{aligned} \quad (12)$$

where α and β are positive weighting parameters.

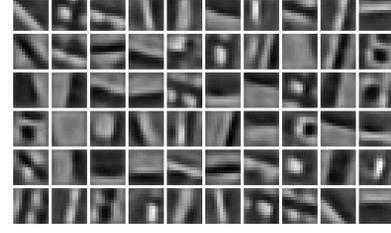


Fig. 3. Some examples of filter bank learned from real hyperspectral data.

C. Performance Amelioration

Since the sparse coding problems for distinct abundance maps are decoupled, we readily make the following extensions to further embrace the amelioration of model performance.

1) *Dictionary Update*: To enhance the generalization ability of convolutional operations in SeCoDe, we adaptively update the convolutional filters (or dictionary), namely, \mathcal{F} , according to the currently investigated hyperspectral scene. Fig. 3 visualizes some examples of convolutional filters with various learned patterns.

2) *Two-Layered Sparsity Regularization*: A reasonable assumption in the case of HU is that the observed spectrum in each pixel is probably contributed by a small set of endmembers [37]. Motivated by the recent success of sparsity-promoting enhancement in different domains [48], [49], we propose to double sparsity in the form of a two-layered sparsity regularization to further highlight the structural information in to-be-estimated maps and reduce the redundancy simultaneously. Note that other sparsity-promoting regularization terms with the forms of $\ell_{1/2}$ quasi-norm [50] and nonconvex log-sum function [23] are also optional.

Let $\Theta' = \{\mathbf{A}, \mathcal{M}, \mathcal{S}, \mathcal{F}\}$ collect all the unknown parameters, and the proposed SeCoDe aims to solve

$$\begin{aligned} \min_{\Theta'} \frac{1}{2} \left\| \mathcal{Y} - \sum_{k=1}^K \mathbf{S}^{(k)} \circ \mathbf{a}_k \right\|_F^2 + \frac{\alpha}{2} \left\| \mathcal{S} - \sum_{d=1}^D \mathbf{F}_d \otimes \mathcal{M}_d \right\|_F^2 \\ + \beta \sum_{d=1}^D \|\mathcal{M}_d\|_1 + \gamma \sum_{k=1}^K \|\mathcal{S}\|_1 \\ \text{s.t. } \Theta' \in \Omega \end{aligned} \quad (13)$$

where

$$\Omega := \begin{cases} \mathbf{M}_d^{(k)} \geq \mathbf{0} \\ \mathbf{a}_k \geq \mathbf{0} \\ \mathbf{S}^{(k)} \geq \mathbf{0} \quad \forall d, k \\ \sum_{k=1}^K \mathbf{S}^{(k)} = \mathbf{1} \\ \|\mathbf{F}_d\|_F^2 = 1. \end{cases} \quad (14)$$

D. Model Optimization

The objective function in (13) is a typical nonconvex problem with respect to the unknown variable set Θ' ; thus, it is difficult to directly obtain the analytical solution in one step. A common solution for our case is to divide the original optimization problem into two easily solved

Algorithm 1 SeCoDe for Blind HU: Global Algorithm

Input: Observed HSI \mathcal{Y} , and parameters $\alpha, \beta, \gamma, \maxIter$.

Output: Endmembers \mathbf{A} , abundance maps cube \mathcal{S} .

- 1 **Initialization:** $t = 1$, \mathbf{A} , \mathcal{S} , filter bank \mathcal{F} , $\zeta = 1e - 4$.
- 2 **while** not converged or $t > \maxIter$ **do**
- 3 Fix other variables to update \mathcal{F} and \mathcal{M} by **Algorithm 1**;
- 4 Fix other variables to update \mathcal{S} and \mathbf{A} by **Algorithm 2**;
- 5 Compute the objective function value E^{t+1} and check the convergence condition:
- 6 **if** $|\frac{E^{t+1}-E^t}{E^t}| < \zeta$ **then**
- 7 | Stop iteration;
- 8 **end**
- 9 $t \leftarrow t + 1$;
- 10 **end**

subproblems relying on an AOS that alternatively optimizes each subproblem when the other is fixed, while each subproblem can be efficiently deduced via the ADMM solver. More specifically, the two subproblems can be unfolded in the following. A comprehensive implementation for SeCoDe is given in Algorithm 1.

1) *MC-CSC*: The first subproblem can be featured as a special CSC with multichannel attentions, jointly updating convolutional filters and multichannel maps, i.e.,

$$\begin{aligned} \min_{\mathcal{F}, \mathcal{M}} \frac{\alpha}{2} \|\mathcal{S} - \sum_{d=1}^D \mathbf{F}_d \otimes \mathcal{M}_d\|_F^2 + \beta \sum_{d=1}^D \|\mathcal{M}_d\|_1 \\ \text{s.t. } \mathcal{M}_d \geq 0, \quad \|\mathbf{F}_d\|_F^2 = 1 \quad \forall d \end{aligned} \quad (15)$$

which can be efficiently solved using ADMM in an inexact fashion [47], [51]–[53].

With the filters \mathcal{F} fixed, we first update the sparse coefficient map bank $\mathcal{M} = \{\mathcal{M}_d\}_{d=1}^D$ by solving a set of K independent subproblems in parallel

$$\min_{\mathbf{M}_d^{(k)} \geq 0} \frac{1}{2} \|\mathbf{S}^{(k)} - \sum_{d=1}^D \mathbf{F}_d \otimes \mathbf{M}_d^{(k)}\|_F^2 + \frac{\alpha}{\beta} \sum_{d=1}^D \|\mathbf{M}_d^{(k)}\|_{1,1} \quad \forall k. \quad (16)$$

For each k th subproblem, we omit the superscript for simplicity, by defining

$$\tilde{\mathbf{F}} = [\tilde{\mathbf{F}}_1, \dots, \tilde{\mathbf{F}}_D], \quad \mathbf{M} = [\mathbf{M}_1^\top, \dots, \mathbf{M}_D^\top]^\top \quad (17)$$

where $\tilde{\mathbf{F}}_d$ is a linear operator such that $\tilde{\mathbf{F}}_d \mathbf{M}_d = \mathbf{F}_d \otimes \mathbf{M}_d$; we can rewrite each (16) into a compact form

$$\min_{\mathbf{M} \geq 0} \frac{1}{2} \|\mathbf{S} - \tilde{\mathbf{F}} \mathbf{M}\|_F^2 + \frac{\alpha}{\beta} \|\mathbf{M}\|_{1,1}. \quad (18)$$

We introduce auxiliary variable $\mathbf{O} = \mathbf{M}$, obtaining the corresponding augmented Lagrangian function in the scaled version

$$\begin{aligned} \mathcal{L}(\mathbf{M}, \mathbf{O}, \Gamma_1) = \frac{1}{2} \|\mathbf{S} - \tilde{\mathbf{F}} \mathbf{M}\|_F^2 + \frac{\nu_1}{2} \|\mathbf{M} - \mathbf{O} + \Gamma_1\|_F^2 \\ + \frac{\alpha}{\beta} \|\mathbf{O}\|_{1,1} + l_R^+(\mathbf{O}) \end{aligned} \quad (19)$$

Algorithm 2 MC-CSC

Input: Abundance maps cube \mathcal{S} , filter bank \mathcal{F} , and parameters α, β, \maxIter .

Output: \mathcal{F} , sparse coefficient map bank \mathcal{M} .

- 1 **Initialization:** $t = 1$, auxiliary variables $\mathbf{O} = \mathbf{M}$, $\mathbf{t}_d = \mathbf{f}_d$, and Lagrangian multipliers $\Gamma_1 = \mathbf{0}$, $\Gamma_d = \mathbf{0}$, $\zeta = 1e - 4$.
- 2 **while** not converged or $t > \maxIter$ **do**
- 3 Fix other variables to update \mathbf{M} by solving Eq. (20);
- 4 Fix other variables to update \mathbf{O} by Eq. (21);
- 5 Fix other variables to update Γ_1 by Eq. (22);
- 6 Fix other variables to update \mathbf{F} by solving Eq. (27);
- 7 Fix other variables to update \mathbf{t}_d by Eq. (28);
- 8 Fix other variables to update Γ_d by Eq. (29);
- 9 Compute the objective function value E^{t+1} and check the convergence condition:
- 10 **if** $|\frac{E^{t+1}-E^t}{E^t}| < \zeta$ **then**
- 11 | Stop iteration;
- 12 **end**
- 13 $t \leftarrow t + 1$;
- 14 **end**

where Γ_1 is the Lagrange multiplier, ν_1 is a positive penalty parameter, and $l_R^+(\cdot)$ acts as a positive-enforcement operator. Then, under the ADMM framework, we can solve the problem by iteratively updating the three variables.

For the variable \mathbf{M} , Wohlberg [47] proposed to exploit the fast Fourier transform (FFT) in equivalently solving independent linear systems

$$[(\hat{\mathbf{F}})^H \hat{\mathbf{F}} + \nu_1 \mathbf{I}] \hat{\mathbf{M}} = (\hat{\mathbf{F}})^H \mathbf{S} + \nu_1 (\hat{\mathbf{O}} - \hat{\Gamma}_1) \quad (20)$$

where $(\cdot)^H$ denotes the conjugate transpose of a complex matrix. The solution can be efficiently computed by applying the Sherman–Morrison formula.

The updates of \mathbf{O} and Γ are readily solved by

$$\mathbf{O} = \max\{0, \mathcal{D}_{\alpha/\beta\nu_1}(\mathbf{M} + \Gamma_1)\} \quad (21)$$

$$\Gamma_1 = \Gamma_1 + \mathbf{M} - \mathbf{O} \quad (22)$$

where $\mathcal{D}_\rho(\cdot)$ is the soft thresholding operator, i.e., $\mathcal{D}_\rho(x) = \text{sign}(x) \max(0, |x| - \rho)$.

After updating the sparse coefficient maps, we then fix maps \mathcal{M} and update filters \mathcal{F} by solving the following subproblem:

$$\min_{\mathbf{F}_d} \frac{1}{2} \sum_{k=1}^K \left\| \mathbf{S}^{(k)} - \sum_{d=1}^D \mathbf{F}_d \otimes \mathbf{M}_d^{(k)} \right\|_F^2, \quad \text{s.t. } \|\mathbf{F}_d\|_F^2 = 1 \quad \forall d. \quad (23)$$

Li *et al.* [54] utilized a proximal gradient descent method to solve it. Instead, we adopt ADMM again, for its prospect of interleaving the ADMM iterations for two subproblems as a whole, which ensures high efficiency and promising practicality [47].

We can rewrite (23) by vectorizing variables and defining \mathbf{P} as zero-padding operator

$$\min_{\mathbf{f}_d} \frac{1}{2} \sum_{k=1}^K \left\| \mathbf{s}^{(k)} - \sum_{d=1}^D \mathbf{f}_d \otimes \mathbf{m}_d^{(k)} \right\|_2^2, \quad \text{s.t. } \mathbf{f}_d \in C_{\text{pn}} \quad \forall d \quad (24)$$

where $C_{\text{pn}} = \{\mathbf{f} \in \mathbb{R}^N | (\mathbf{I} - \mathbf{P}\mathbf{P}^\top)\mathbf{f} = 0, \|\mathbf{f}\|_2^2 = 1\}$ is the constraint set. Similarly, we can rewrite its scaled Lagrangian function with auxiliary variable $\mathbf{t}_d = \mathbf{f}_d$

$$\mathcal{L}(\mathbf{f}_d, \mathbf{t}_d) = \frac{1}{2} \sum_{k=1}^K \left\| \mathbf{s}^{(k)} - \sum_{d=1}^D \mathbf{f}_d \otimes \mathbf{m}_d^{(k)} \right\|_2^2 + \sum_{d=1}^D \delta_{C_{\text{pn}}}(\mathbf{t}_d) + \frac{\nu_2}{2} \sum_{d=1}^D \|\mathbf{f}_d - \mathbf{t}_d + \Gamma_d\|_2^2 \quad (25)$$

where $\delta_{C_{\text{pn}}}$ is the indicator function that penalizes the out-of-set solution.

The padded filter \mathbf{f}_d is updated in the DFT domain by

$$\min_{\mathbf{F}} \frac{1}{2} \sum_{k=1}^K \|\widehat{\mathbf{s}}^{(k)} - \widehat{\mathbf{M}}^{(k)} \widehat{\mathbf{F}}\|_2^2 + \frac{\nu_2}{2} \|\widehat{\mathbf{F}} - \widehat{\mathbf{T}} + \widehat{\Gamma}_2\|_2^2 \quad (26)$$

where $\widehat{\mathbf{M}}^{(k)} = [\widehat{\mathbf{M}}_1^{(k)}, \dots, \widehat{\mathbf{M}}_D^{(k)}]$ with $\widehat{\mathbf{M}}_d^{(k)} = \text{diag}(\widehat{\mathbf{m}}_d^{(k)}) \in \mathbb{R}^{N \times N}$, and $\widehat{\mathbf{F}} = [\widehat{\mathbf{f}}_1^\top, \dots, \widehat{\mathbf{f}}_D^\top]^\top$, similar for $\widehat{\mathbf{T}}$ and $\widehat{\Gamma}_2$. By following (20), we can get the solution by solving:

$$\left[\sum_{k=1}^K (\widehat{\mathbf{M}}^{(k)})^\top \widehat{\mathbf{M}}^{(k)} + \nu_2 \mathbf{I} \right] \widehat{\mathbf{F}} = \sum_{k=1}^K (\widehat{\mathbf{M}}^{(k)})^\top \widehat{\mathbf{s}}^{(k)} + \nu_2 (\widehat{\mathbf{T}} - \widehat{\Gamma}_2) \quad (27)$$

with the iterated Sherman–Morrison formula.

The auxiliary variable \mathbf{t}_d and dual variables Γ_d are updated in forms of

$$\mathbf{t}_d = \frac{\mathbf{P}\mathbf{P}^\top(\mathbf{f}_d + \Gamma_d)}{\|\mathbf{P}\mathbf{P}^\top(\mathbf{f}_d + \Gamma_d)\|_2^2} \quad (28)$$

$$\Gamma_d = \Gamma_d + \mathbf{f}_d - \mathbf{t}_d. \quad (29)$$

More details in solving the MC-CSC are then summarized in Algorithm 2.

2) *Coupled Spectral Unmixing (CSU)*: Once the convolutional kernels and sparse feature maps are obtained by solving the MC-CSC, the second part can further integrates them into a set of unique endmembers and corresponding abundance maps by solving the following CSU:

$$\begin{aligned} \min_{\mathcal{S}, \mathbf{A}} \quad & \frac{1}{2} \left\| \mathcal{Y} - \sum_{k=1}^K \mathbf{S}^{(k)} \circ \mathbf{a}_k \right\|_F^2 + \frac{\alpha}{2} \|\mathcal{S} - \mathcal{Z}\|_F^2 + \gamma \|\mathcal{S}\|_1 \\ \text{s.t.} \quad & \mathcal{S}, \mathbf{A} \geq \mathbf{0}, \sum_{k=1}^K \mathbf{S}^{(k)} = \mathbf{1} \end{aligned} \quad (30)$$

where $\mathcal{Z} = \sum_{d=1}^D \mathbf{F}_d \otimes \mathcal{M}_d$.

We first introduce multiple auxiliary variables \mathbf{G} , \mathbf{H} , and \mathbf{J} , obtaining the scaled augmented Lagrangian function as follows:

$$\begin{aligned} \mathcal{L} = & \frac{1}{2} \|\mathbf{Y}_{(3)}^\top - \mathbf{S}_{(3)}^\top \mathbf{A}\|_F^2 + \frac{\alpha}{2} \|\mathbf{S}_{(3)}^\top - \mathbf{Z}_{(3)}^\top\|_F^2 + \gamma \|\mathbf{J}\|_{1,1} \\ & + \frac{\mu}{2} \|\mathbf{S}_{(3)}^\top - \mathbf{H} + \Lambda_2\|_F^2 + \frac{\mu}{2} \|\mathbf{S}_{(3)}^\top - \mathbf{J} + \Lambda_3\|_F^2 \\ & + \frac{\mu}{2} \|\mathbf{A} - \mathbf{G} + \Lambda_1\|_F^2 + l_R^+(\mathbf{G}) + l_R^+(\mathbf{H}) \end{aligned} \quad (31)$$

where $\{\Lambda_i\}_{i=1}^3$ are the Lagrange multipliers, and μ is the positive penalty parameter. By setting the derivatives of (31)

Algorithm 3 CSU

Input: Observed HSI \mathcal{Y} , endmembers \mathbf{A} , filter bank \mathcal{F} , co-efficient map bank \mathcal{M} , and parameters $\alpha, \gamma, \text{maxIter}$.

Output: \mathbf{A} , abundance maps cube \mathcal{S} .

```

1 Initialization:  $t = 1$ , auxiliary variables  $\mathbf{G} = \mathbf{A}$ ,
 $\mathbf{H} = \mathbf{J} = \mathbf{S}$ , and Lagrangian multipliers  $\Lambda_1 = \mathbf{0}$ ,
 $\Lambda_2 = \Lambda_3 = \mathbf{0}$ ,  $\zeta = 1e - 4$ .
2 while not converged or  $t > \text{maxIter}$  do
3   Fix other variables to update  $\mathbf{S}$  by Eq. (32);
4   Fix other variables to update  $\mathbf{A}$  by Eq. (33);
5   Fix other variables to update  $\mathbf{G}$ ,  $\mathbf{H}$ , and  $\mathbf{J}$  by Eq. (34);
6   Fix other variables to update  $\Lambda_1$ ,  $\Lambda_2$ , and  $\Lambda_3$  by
   Eq. (35);
7   Compute the objective function value  $E^{t+1}$  and check
   the convergence condition:
8   if  $|\frac{E^{t+1} - E^t}{E^t}| < \zeta$  then
9     | Stop iteration;
10  end
11   $t \leftarrow t + 1$ ;
12 end

```

with respect to each unknown variable to zero, we can readily obtain the updating rules as

$$\begin{cases} \mathbf{S}_{(3)}^\top = [\mathbf{Y}_{(3)}^\top \mathbf{A}^\top + \alpha \mathbf{Z}_{(3)}^\top + \mu(\mathbf{H} + \mathbf{J} - \Lambda_2 - \Lambda_3)] \\ \quad / (\mathbf{A}\mathbf{A}^\top + (\alpha + 2\mu)\mathbf{I}) & (32) \\ \mathbf{A} = (\mathbf{S}_{(3)} \mathbf{S}_{(3)}^\top + \mu \mathbf{I}) \setminus [\mathbf{S}_{(3)} \mathbf{Y}_{(3)}^\top + \mu(\mathbf{G} - \Lambda_1)] & (33) \\ \mathbf{G} = \max(0, \mathbf{A} + \Lambda_1), \quad \mathbf{H} = \max(0, \mathbf{S}_{(3)}^\top + \Lambda_2) & (34) \\ \mathbf{J} = \mathcal{D}_{\gamma/\mu}(\mathbf{S}_{(3)}^\top + \Lambda_3) & (34) \\ \Lambda_1 = \Lambda_1 + \mathbf{A} - \mathbf{G}, \Lambda_2 = \Lambda_2 + \mathbf{S}_{(3)}^\top - \mathbf{H} \\ \Lambda_3 = \Lambda_3 + \mathbf{S}_{(3)}^\top - \mathbf{J} & (35) \end{cases}$$

where \setminus denotes matrix left division. Algorithm 3 details the specific steps for solving the CSU's subproblem.

Finally, we repeat these optimization procedures until a presetting stopping criterion is satisfied.

3) *Computational Complexity*: The computational cost of the proposed SeCoDe is mainly dominated by the above two subproblems. The ADMM solver of the latter CSU only involves matrix multiplication, addition, and inversion, yielding $\mathcal{O}(KLN)$. As for the former MC-CSC subproblem, it is the cost of the FFTs and linear systems solvers for (20) and (27) that dominate the computational complexity, whereas the cost can be generally rated as $\mathcal{O}(DN \log N)$ by virtue of applying the (iterated) Sherman–Morrison formula.

V. EXPERIMENTS AND DISCUSSION

In this section, we conduct elaborated experiments on one synthetic and two real data sets, as shown in Fig. 4, to exhibit the effectiveness of proposed SeCoDe for blind HU. We compare the proposed method with seven representative and state-of-the-art unmixing methods, including FCLSU [35], PCLSU [36], SPCLSU [38], SSUnSAL [39], ELMM [40],

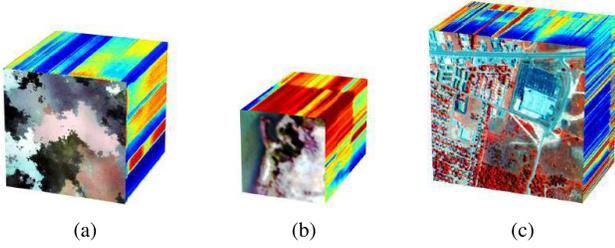


Fig. 4. Examples of three used HSIs, including one simulated scene and two real scenes. (a) Synthetic data. (b) Samson data. (c) Urban data.

MVNTF [33],² SULoRA [30],³ and ALMM [31].³ Some of these methods, such as SULoRA and ALMM, are extended to a blind unmixing version, making them able to simultaneously estimate abundance maps and extract endmembers.

A. General Experimental Setup

For a fair comparison, we optimally tune the parameters involved in competitive methods and report the average results out of ten runs. All experiments are implemented with MATLAB 2017b on a Windows 10 Operation System and conducted on Intel Core i7-8700K 3.70-GHz desktop with 64-GB memory.

1) *Implementation Details*: Since the details of implementing CSC are not the focus of this article, for space brevity, we leave out the discussion on the number of filters D , the filter size p , and different initialization. We empirically set them as 36 and 12, after weighing the performance gain and extra computational cost. The filters are simply randomly initialized with the central $1/9$ part. In addition, several important regularization parameters (α , β , and γ) in our SeCoDe will be emphatically discussed and analyzed in Section V-B2.

In our case, the number of endmembers K can be determined by using *hyperspectral signal identification by minimum error* (HySime) [55], and the endmembers are further extracted by VCA. Due to the highly ill-posed problem and strong nonconvexity of the model, a bad initialization may render the algorithm converge into a local minimum. Hence, we select the output (abundance maps) of SPCLSU as a good starting point. Finally, the materials can be identified by similarity measurement (spectral angle) between references and estimated endmembers.

2) *Performance Metrics*: We introduce three performance metrics to quantitatively assess the experimental results. Given faithful references \mathbf{S} and \mathbf{A} , the abundance overall root mean square error (aRMSE)

$$\text{aRMSE} = \frac{1}{N} \sum_{n=1}^N \sqrt{\frac{1}{K} \|\mathbf{s}_n - \hat{\mathbf{s}}_n\|_2^2} \quad (36)$$

and the spectral angle distance (SAD)

$$\text{SAD} = \frac{1}{K} \sum_{k=1}^K \arccos \left(\frac{\mathbf{a}_k^\top \hat{\mathbf{a}}_k}{\|\mathbf{a}_k\|_2 \|\hat{\mathbf{a}}_k\|_2} \right) \quad (37)$$

²<https://github.com/bearshng/mvntf>

³<https://sites.google.com/view/danfeng-hong/data-code>

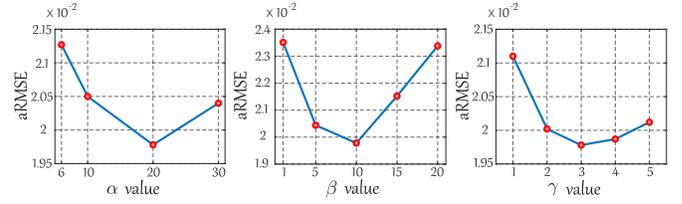


Fig. 5. Parameter sensitivity analysis in SeCoDe with respect to α , β , and γ on the synthetic data set.

are utilized to assess the fidelity of estimated abundance maps $\hat{\mathbf{S}}$ and endmembers $\hat{\mathbf{A}}$, respectively.

Due to the lack of fully reliable ground truth for abundances in the real scene, we additionally perform classification in terms of the overall accuracy (OA), to relax the requirement for references with high fidelity and give a more comprehensive assessment. The classification maps can be obtained by identifying each pixel into a certain endmember that has the maximal response.

B. Synthetic Data

1) *Data Description*: The synthetic data have been widely applied in [30], [31], and [40] to evaluate the unmixing performance quantitatively. It consists of 200×200 pixels with 224 spectral bands in the wavelength ranging from 400 to 2500 nm. Moreover, five materials selected from USGS spectral library are investigated in the studied scene. Note that the image is simulated by adding scaling factors and complex noises in order to assess the unmixing ability in the presence of SV. Please refer to [40] for more details regarding this data.

2) *Parameter Sensitivity Analysis*: Since the unmixing performance is sensitive to the setting of three regularization parameters α , β , and γ in SeCoDe to some extent, it is, therefore, indispensable to find a proper range to guide the determination of parameters. After that, we adopt a grid search strategy throughout our experiments. To be specific, we iteratively evaluate the optimal value of each parameter while fixing the others, until an optimal setting with the lowest performance metrics is reached. As shown in Fig. 5, the optimal values with respect to the three parameters, i.e., $\{2e - 2, 1e - 2, 3\}$, can be easily found with approximately convex curves.

3) *Parameter Setting*: The parameters for all the algorithms are recorded as follows. For the SUnSAL, the sparsity-promoting regularization is weighed by $2e - 3$. We set the regularization parameter in ELMM as $5e - 1$. The parameter combination for SULoRA is selected as $\{1e - 2, 1e - 2, 4e - 4\}$, and ALMMs' is $\{4e - 3, 2e - 3, 5e - 3\}$.

4) *Results and Discussion*: Table I summarizes the aRMSEs and SADs for all competitive methods. We put the abundance maps of ground truth in the left-hand column of Fig. 6. To highlight the differences of estimated abundance maps for different methods, we show the residual images between the abundance maps and GT in Fig. 6.

As listed in Table I, we can conclude that SeCoDe achieves the best results in terms of both metrics. The visual results presented in Fig. 6 exhibit a consistent tendency. Specifically, due to strictly following ASC, FCLSU inevitably absorbs SVs in abundance estimation. For that, PCLSU relaxes the

TABLE I
QUANTITATIVE PERFORMANCE COMPARISON WITH THE INVESTIGATED METHODS IN TERMS OF aRMSE AND SAD ON THE SYNTHETIC DATA SET. THE BEST ONE IS SHOWN IN BOLD

Metric	Ideal Value	Method								
		FCLSU	PCLSU	SPCLSU	SSUnSAL	ELMM	MVNTF	SULoRA	ALMM	SeCoDe
aRMSE	0	0.0679	0.0397	0.0256	0.0303	0.0356	0.0424	0.0252	0.0226	0.0202
SAD	0	0.0168	0.085	0.0085	0.0097	0.0080	0.0101	0.0081	0.0079	0.0074
Runtime (Sec.)	-	2.82	1.62	1.67	2.10	434.61	545.34	18.31	184.91	223.19

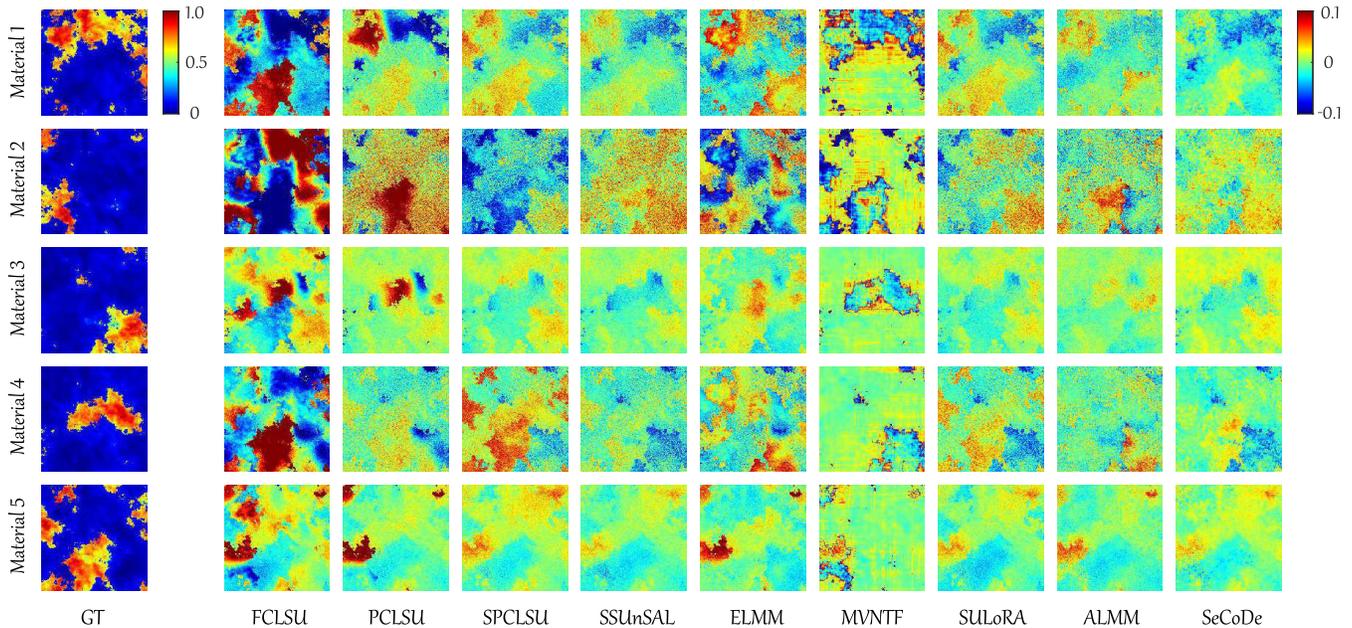


Fig. 6. Difference maps between real abundance maps and estimated abundance maps for different compared unmixing methods on the synthetic data set. The first row lists the real abundance maps for each material.

ASC in FCLSU and, hence, alleviates absorption of SVs to some extent, which brings an evident improvement. Scaled methods, such as SSUnSAL and SPCLSU, further increase the performance by modeling the main constituent of SVs, i.e., scaling factors. The results of MVNTF show evident rough patterns; the possible reason is the lack of consideration of strong SVs contained in this scene. Although the ELMM is designed with meaningfully physical consideration, the complex optimization makes the model easily trapped in local minima. As a result, it performs an unsatisfactory estimation of abundances, as also indicated by the salient region with dark colors on the residual map of material 2. Extended from ELMM, ALMM successfully avoids such a dilemma and produces the second best result. The result of SULoRA comes in third due to its robust low-rank attribute embedding. Different from the dense noises in the results of traditional methods, the SeCoDe’s results are clearer. This means that the errors are mainly located in fewer pixels, which declares a more faithful result as a whole.

C. Real Data—Samson

1) *Data Description*: The Samson data set was captured by a push broom, visible to the near-infrared sensor and

well-calibrated. The original image contains 952×952 pixels measured in 156 bands ranging from 401 to 889 nm, with a spectral resolution highly up to 3.13 nm. We choose a region of interest (ROI) with the size of 95×95 pixels. The referential abundance maps of three appeared endmembers, i.e., “#1 Rock,” “#2 Tree,” and “#3 Water,” provided from the website,⁴ are shown in Fig. 7 (left).

2) *Parameter Setting*: We record the tuned parameters for all the algorithms as follows. The sparsity-induced regularization term in SSUnSAL is weighted by $5e - 3$. The tradeoff parameter for two fidelity terms in ELMM is set as $5e - 1$. For SULoRA and ALMM, the optimal parameter combinations are set as $\{1, 1e - 2, 2e - 2\}$ and $\{2e - 4, 2e - 2, 4e - 3, 1e - 3, 20\}$, respectively. We choose α , β , and γ in SeCoDe as $1e - 1$, $1e - 2$, and $5e - 1$, respectively.

3) *Results and Analysis*: Fig. 7 visualizes the abundance maps of different methods, and the corresponding quantitative results are given in Table II. Overall, FCLSU and PCLSU yield relatively poor results as they fail to model SVs. By considering the principal SV-scaling factors, SPCLSU achieves an obvious improvement in aRMSE. ELMM assumes different scaling factors on endmembers, yielding further performance

⁴<https://rslab.ut.ac.ir/data>

TABLE II
QUANTITATIVE PERFORMANCE COMPARISON WITH THE INVESTIGATED METHODS IN TERMS OF ARMSE, SAD, AND OA ON THE SAMSON DATA SET. THE BEST ONE IS SHOWN IN BOLD

Metric	Ideal Value	Method								
		FCLSU	PCLSU	SPCLSU	SSUnSAL	ELMM	MVNTF	SULoRA	ALMM	SeCoDe
aRMSE	0	0.1653	0.1492	0.1207	0.0779	0.1158	0.1336	0.0689	0.0992	0.0517
SAD	0	0.1267	0.0844	0.0844	0.1054	0.1405	0.0732	0.0701	0.0622	0.0547
OA (%)	100	80.41	89.12	89.14	89.14	85.30	88.68	91.24	88.40	93.91
Runtime (Sec.)	-	0.42	0.56	0.57	0.54	168.33	432.65	4.80	36.97	58.14

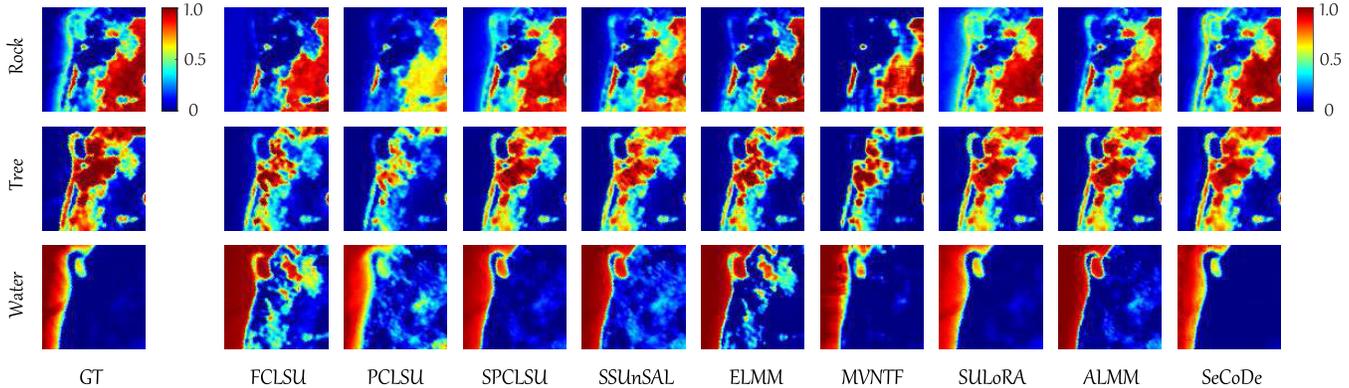


Fig. 7. Qualitative comparison of abundance maps estimated by the proposed method and state-of-the-art compared methods on the Samson data set.

TABLE III
QUANTITATIVE PERFORMANCE COMPARISON WITH THE INVESTIGATED METHODS IN TERMS OF ARMSE, SAD, AND OA ON THE HYDICE URBAN DATA. THE BEST ONE IS SHOWN IN BOLD

Metric	Ideal Value	Method								
		FCLSU	PCLSU	SPCLSU	SSUnSAL	ELMM	MVNTF	SULoRA	ALMM	SeCoDe
aRMSE	0	0.2515	0.1889	0.1259	0.1193	0.1857	0.2012	0.1179	0.1183	0.0797
SAD	0	0.2960	0.1201	0.1201	0.1130	0.0991	0.1486	0.1246	0.0919	0.1050
OA (%)	100	51.37	72.83	72.83	76.73	64.02	60.48	80.20	77.50	91.84
Runtime (Sec.)	-	2.69	1.42	1.48	1.74	1432.56	1946.68	29.26	609.06	868.47

improvement on the basis of SPCLSU. However, the strong nonconvexity hinders ELMM's performance from being further improved. The MVNTF achieves desirable results both visually and quantitatively, particularly for SAD. Beyond them, except for the principal scaling factors, ALMM models other complex SVs as well, tending to hold finer unmixing results. With the effective sparsity prior and scale modeling on abundance maps, SSUnSAL and SULoRA dramatically perform better than those previous methods for both abundance estimation and endmember extraction. As expected, the blind unmixing performance in all metrics of the proposed SeCoDe is superior to that of other competitive methods, demonstrating its superiority and effectiveness. Also, a similar trend can be visually validated in Fig. 7. It is worth noting that SeCoDe identifies the area of *Water* and *Rock* much better than SSUnSAL and SULoRA, respectively, showing a more realistic abundance estimation compared to the GT.

D. Real Data—Urban

1) *Data Description*: The second real data set was captured by the Hyperspectral Digital Image Collection Experi-

ment (HYDICE) sensor in October 1995, covering an urban area at Copperas Cove, TX, USA.⁵ The original image contains 307×307 pixels measured in 210 bands ranging from 400 to 2500 nm, at a 2-m ground sample distance (GSD) and a 10-nm spectral resolution. As a common preprocessing procedure, we manually remove severely degraded bands with the water vapor and atmospheric effects, i.e., bands with number 1–4, 76, 87, 101–111, 136–153, and 198–210. Fig. 4(c) shows the data cube with the remaining 162 bands. The widely used references for endmembers and abundance maps can be found on the website.⁶

2) *Parameter Setting*: The optimal parameters for the competitive methods on this data set are $2e - 3$ for SSUnSAL, $5e - 1$ for ELMM, $\{1, 1e - 2, 6e - 4\}$ for SULoRA, and $\{2e - 2, 2e - 2, 6e - 3\}$ for ALMM. The α , β , and γ in SeCoDe are parameterized by $6e - 1$, $1e - 2$, and 1, respectively.

3) *Results and Analysis*: The visual results are shown in Fig. 8, and the quantitative results are listed in Table III. As can be seen from the table, FCLSU and MVNTF yield relatively poor unmixing results with only around 60% OAs.

⁵<http://www.agc.army.mil/>

⁶http://www.escience.cn/people/feiyunZHU/Dataset_GT.html

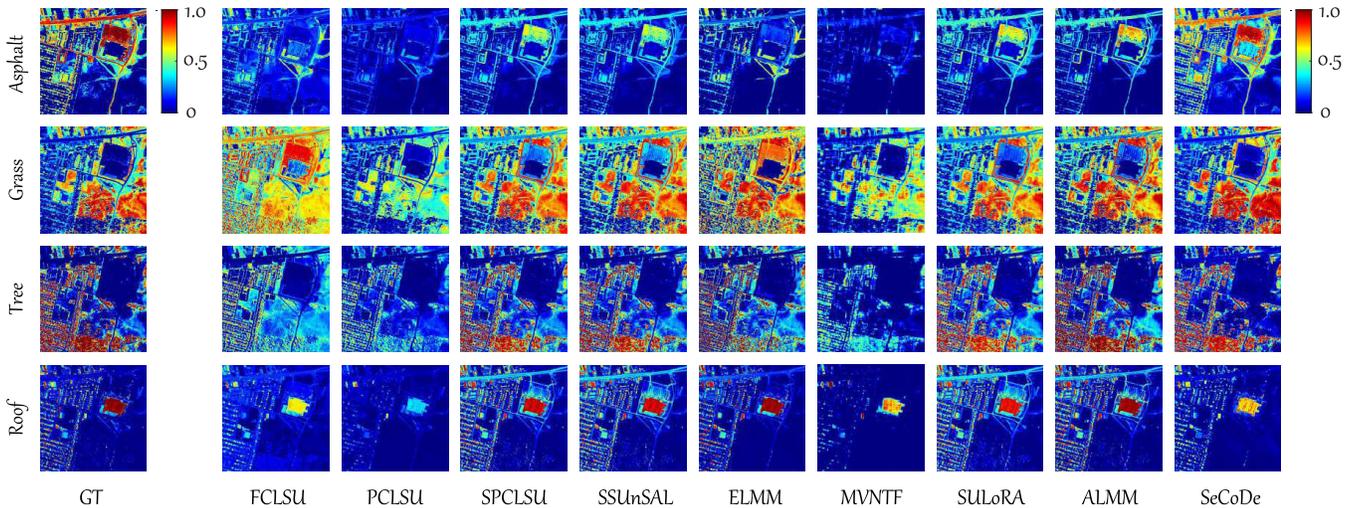


Fig. 8. Qualitative comparison of abundance maps estimated by the proposed method and state-of-the-art compared methods on the Urban data set.

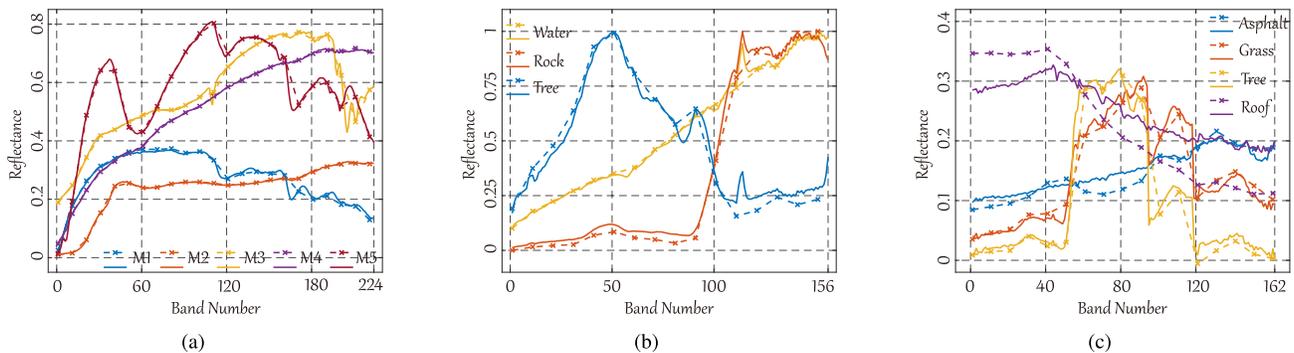


Fig. 9. Visual comparison for endmembers on the three investigated data sets, where the dashed cross line represents the estimated endmembers obtained by SeCoDe, and the solid line means the references. (a) Synthetic data. (b) Samson data. (c) Urban data.

The metrics on abundance maps of SPCLSU and SSUnSAL are much better, due to the consideration that their estimated endmembers are more accurate. ALMM is able to extract more reliable endmembers and slightly surpass SeCoDe in terms of SAD. However, great advantages of aRMSE and OA show the effectiveness of the proposed method by integrating convolutional representations, which is able to fully capture spatial information underlying hyperspectral data.

More observations can be found in Fig. 8. Although SeCoDe performs a relatively poor visualization on the material of *Roof* from a numerical perspective, it can provide a more realistic spatial distribution for the abundances of each material. Taking the long horizontal road in the top half of the scene as an example, both MVNTF and PCLSU fail to capture it, whereas FCLSU and ELMM improperly unmix it as an ingredient of *Grass*. On the other hand, it is mistakenly classified into *Roof* by the other four competitive methods, which may share similar endmembers since they are both man-made materials. Another big area of *asphalt* besides the *Roof* is also difficult to be recognized, whereas it is accurately estimated only by our SeCoDe. These pieces of evidence objectively validate the superiority of the proposed method compared to not only traditional NMF-based methods but also complex methods that addressing SVs.

E. Endmember Visualization

Finally, we analyze the endmembers estimated by our method and references on the three data sets, as visualized in Fig. 9. In the synthetic data, SeCoDe is capable of extracting high-quality endmembers, yielding a good matching with references, especially for material 2 (M2) and material 4 (M4), while, in the real scenarios, the gap between the estimations and references is relatively obvious. A reasonable explanation might be twofold. On the one hand, the complex imaging environment in real cases tends to generate more challenging SVS than in synthetic ones, which possibly brings negative effects in estimating endmembers. On the other hand, the acquisition and reliability for the references of endmembers remain limited; thus, it is difficult to accurately measure the similarities between these estimated endmembers and so-called ground truth. Despite so, the proposed SeCoDe performs relatively realistic estimations for endmembers, particularly in spectral shapes and the positions of band absorption.

VI. CONCLUSION

Motivated by the recent success of tensor analysis in the remote sensing community, we propose a tensor-based convolutional decomposition model, called SeCoDe, with a sparsity-enhanced constraint for the blind HU tasks. SeCoDe

learns 3-D convolutional kernels and regards them as spectral bundles, thereby effectively eliminating the effects of SVs. Moreover, the resulting MC-CSC is able to not only model the spatial contextual information utilizing convolutional operations but also preserve the spectral continuity. To further ameliorate the unmixing performance, we merge an adaptive filter update and a two-layered sparsity regularization into a general framework. With these well-designed strategies, the proposed method achieves the best unmixing performance among experiments on both synthetic and real data sets compared to previous methods. In future work, we will pose our efforts in developing more tensor-based models with deeper consideration of intrinsic properties, such as nonlocal information, to better understand the issue of HU and robustly address various SVs.

REFERENCES

- [1] D. Hong *et al.*, "Interpretable hyperspectral artificial intelligence: When non-convex modeling meets hyperspectral remote sensing," *IEEE Geosci. Remote Sens. Mag.*, early access, doi: [10.1109/MGRS.2021.3064051](https://doi.org/10.1109/MGRS.2021.3064051).
- [2] Y. Dong, B. Du, L. Zhang, and L. Zhang, "Dimensionality reduction and classification of hyperspectral images using ensemble discriminative local metric learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2509–2524, May 2017.
- [3] D. Hong, N. Yokoya, J. Chanussot, J. Xu, and X. X. Zhu, "Learning to propagate labels on graphs: An iterative multitask regression framework for semi-supervised hyperspectral dimensionality reduction," *ISPRS J. Photogramm. Remote Sens.*, vol. 158, pp. 35–49, Dec. 2019.
- [4] D. Hong, N. Yokoya, J. Chanussot, J. Xu, and X. X. Zhu, "Joint and progressive subspace analysis (JPSA) with spatial-spectral manifold alignment for semisupervised hyperspectral dimensionality reduction," *IEEE Trans. Cybern.*, early access, Nov. 11, 2020, doi: [10.1109/TCYB.2020.3028931](https://doi.org/10.1109/TCYB.2020.3028931).
- [5] X. Wu, D. Hong, J. Tian, J. Chanussot, W. Li, and R. Tao, "ORSIm detector: A novel object detection framework in optical remote sensing imagery using spatial-frequency channel features," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5146–5158, Jul. 2019.
- [6] X. Wu, D. Hong, J. Chanussot, Y. Xu, R. Tao, and Y. Wang, "Fourier-based rotation-invariant feature boosting: An efficient framework for geospatial object detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 302–306, Feb. 2020.
- [7] X. Cao, F. Zhou, L. Xu, D. Meng, Z. Xu, and J. Paisley, "Hyperspectral image classification with Markov random fields and a convolutional neural network," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2354–2367, May 2018.
- [8] R. Hang, Q. Liu, D. Hong, and P. Ghamisi, "Cascaded recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5384–5394, Aug. 2019.
- [9] D. Hong, X. Wu, P. Ghamisi, J. Chanussot, N. Yokoya, and X. X. Zhu, "Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 3791–3808, Jun. 2020.
- [10] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, early access, Aug. 18, 2020, doi: [10.1109/TGRS.2020.3015157](https://doi.org/10.1109/TGRS.2020.3015157).
- [11] Y. Zhou, A. Rangarajan, and P. D. Gader, "A spatial compositional model for linear unmixing and endmember uncertainty estimation," *IEEE Trans. Image Process.*, vol. 25, no. 12, pp. 5987–6002, Dec. 2016.
- [12] T. Uezato, R. J. Murphy, A. Melkumyan, and A. Chlingaryan, "Incorporating spatial information and endmember variability into unmixing analyses to improve abundance estimates," *IEEE Trans. Image Process.*, vol. 25, no. 12, pp. 5563–5575, Dec. 2016.
- [13] M. Tang, L. Gao, A. Marinoni, P. Gamba, and B. Zhang, "Integrating spatial information in the normalized P-linear algorithm for nonlinear hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1179–1190, Apr. 2018.
- [14] D. Hong, J. Yao, D. Meng, Z. Xu, and J. Chanussot, "Multimodal GANs: Toward crossmodal hyperspectral-multispectral image segmentation," *IEEE Trans. Geosci. Remote Sens.*, early access, Sep. 16, 2020, doi: [10.1109/TGRS.2020.3020823](https://doi.org/10.1109/TGRS.2020.3020823).
- [15] L. Gao, D. Hong, J. Yao, B. Zhang, P. Gamba, and J. Chanussot, "Spectral superresolution of multispectral imagery with joint sparse and low-rank learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 2269–2280, Mar. 2021.
- [16] R. Hang, Z. Li, P. Ghamisi, D. Hong, G. Xia, and Q. Liu, "Classification of hyperspectral and LiDAR data using coupled CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4939–4950, Jul. 2020.
- [17] D. Hong *et al.*, "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, early access, Aug. 24, 2020, doi: [10.1109/TGRS.2020.3016820](https://doi.org/10.1109/TGRS.2020.3016820).
- [18] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [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] B. Zhang, X. Sun, L. Gao, and L. Yang, "Endmember extraction of hyperspectral remote sensing images based on the ant colony optimization (ACO) algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 7, pp. 2635–2646, Jul. 2011.
- [21] V. P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Linear Algebra Appl.*, vol. 416, no. 1, pp. 29–47, Jul. 2006.
- [22] S. Zhang, J. Li, Z. Wu, and A. Plaza, "Spatial discontinuity-weighted sparse unmixing of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5767–5779, Oct. 2018.
- [23] J. Yao, D. Meng, Q. Zhao, W. Cao, and Z. Xu, "Nonconvex-sparsity and Nonlocal-Smoothness-Based blind hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2991–3006, Jun. 2019.
- [24] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sens. Environ.*, vol. 115, no. 7, pp. 1603–1616, Jul. 2011.
- [25] A. Zare and K. C. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan. 2014.
- [26] L. Drumetz, J. Chanussot, and C. Jutten, "Variability of the endmembers in spectral unmixing: Recent advances," in *Proc. 8th IEEE Workshop Hyperspectral Image Signal Process., Evol. Remote Sens.*, Aug. 2016, pp. 1–5.
- [27] D. A. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. O. Green, "Mapping chaparral in the santa monica mountains using multiple endmember spectral mixture models," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 267–279, Sep. 1998.
- [28] B. Somers, M. Zortea, A. Plaza, and G. P. Asner, "Automated extraction of image-based endmember bundles for improved spectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 396–408, Apr. 2012.
- [29] L. Drumetz, T. R. Meyer, J. Chanussot, A. L. Bertozzi, and C. Jutten, "Hyperspectral image unmixing with endmember bundles and group sparsity inducing mixed norms," *IEEE Trans. Image Process.*, vol. 28, no. 7, pp. 3435–3450, Jul. 2019.
- [30] D. Hong and X. X. Zhu, "SULOra: Subspace unmixing with low-rank attribute embedding for hyperspectral data analysis," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 6, pp. 1351–1363, Dec. 2018.
- [31] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.
- [32] Q. Zhang, H. Wang, R. J. Plemmons, and V. P. Pauca, "Tensor methods for hyperspectral data analysis: A space object material identification study," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 25, no. 12, pp. 3001–3012, 2008.
- [33] Y. Qian, F. Xiong, S. Zeng, J. Zhou, and Y. Y. Tang, "Matrix-vector nonnegative tensor factorization for blind unmixing of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1776–1792, Mar. 2017.
- [34] A. Cichocki and S.-I. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. New York, NY, USA: Wiley, 2002.

- [35] 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.
- [36] R. Heylen, D. Burazerovic, and P. Scheunders, "Fully constrained least squares spectral unmixing by simplex projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4112–4122, Jun. 2011.
- [37] J. M. Bioucas-Dias and M. A. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. Whispers*, Jun. 2010, pp. 1–4.
- [38] M. A. Veganzones *et al.*, "A new extended linear mixing model to address spectral variability," in *Proc. 6th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Jul. 2014, pp. 1–4.
- [39] Z. Han, D. Hong, L. Gao, B. Zhang, and J. Chanussot, "Deep half-siamese networks for hyperspectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, early access, Aug. 6, 2020, doi: [10.1109/LGRS.2020.3011941](https://doi.org/10.1109/LGRS.2020.3011941).
- [40] L. Drumetz, M.-A. Veganzones, S. Henrot, R. Phlypo, J. Chanussot, and C. Jutten, "Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3890–3905, Aug. 2016.
- [41] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 525–538, Jan. 2016.
- [42] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.
- [43] F. Xiong, Y. Qian, J. Zhou, and Y. Y. Tang, "Hyperspectral unmixing via total variation regularized nonnegative tensor factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2341–2357, Apr. 2019.
- [44] L. Sun, F. Wu, T. Zhan, W. Liu, J. Wang, and B. Jeon, "Weighted nonlocal low-rank tensor decomposition method for sparse unmixing of hyperspectral images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1174–1188, 2020.
- [45] A. Cichocki *et al.*, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, Mar. 2015.
- [46] Y. Song, "Spatial resolution improvement hyperspectral images by deconvolution joint unmixing-deconvolution," Ph.D. dissertation, Dept. Biol., Signals Syst. Cancerology Neurosci., Univ. de Lorraine, Lorraine, France, 2018.
- [47] B. Wohlberg, "Efficient algorithms for convolutional sparse representations," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 301–315, Jan. 2016.
- [48] T. Uezato, M. Fauvel, and N. Dobigeon, "Hyperspectral unmixing with spectral variability using adaptive bundles and double sparsity," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3980–3992, Jun. 2019.
- [49] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, L.-J. Deng, and Y. Wang, "FastDeRain: A novel video rain streak removal method using directional gradient priors," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 2089–2102, Apr. 2019.
- [50] 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.
- [51] Z. Lin, M. Chen, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," 2010, *arXiv:1009.5055*. [Online]. Available: <http://arxiv.org/abs/1009.5055>
- [52] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "CoSpace: Common subspace learning from hyperspectral-multispectral correspondences," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4349–4359, Jul. 2019, doi: [10.1109/TGRS.2018.2890705](https://doi.org/10.1109/TGRS.2018.2890705).
- [53] D. Hong, N. Yokoya, N. Ge, J. Chanussot, and X. X. Zhu, "Learnable manifold alignment (LeMA): A semi-supervised cross-modality learning framework for land cover and land use classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 147, pp. 193–205, Jan. 2019.
- [54] M. Li *et al.*, "Video rain streak removal by multiscale convolutional sparse coding," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6644–6653.
- [55] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.



Jing Yao received the B.Sc. degree from Northwest University, Xi'an, China, in 2014. He is pursuing the Ph.D. degree with the School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an.

From 2019 to 2020, he was a Visiting Student in signal processing in Earth observation (SiPEO) with the Technical University of Munich (TUM), Munich, Germany, and the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Weßling, Germany. His research interests include low-rank modeling, hyperspectral image analysis, and deep learning-based image processing methods.

Mr. Yao received the recognition of the first outstanding paper award named the Jose Bioucas Dias Award of the Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS) in 2021.



Danfeng Hong (Member, IEEE) received the M.Sc. degree (*summa cum laude*) in computer vision from the College of Information Engineering, Qingdao University, Qingdao, China, in 2015, and the Dr.Eng. degree (*summa cum laude*) in signal processing in Earth observation (SiPEO) from the Technical University of Munich (TUM), Munich, Germany, in 2019.

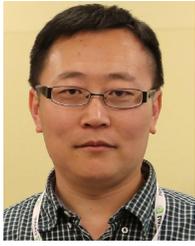
Since 2015, he has been a Research Associate with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Weßling, Germany. He is a Research Scientist and leads the Spectral Vision Working Group, IMF, DLR, and also an Adjunct Scientist with GIPSA-lab, Grenoble INP, CNRS, University of Grenoble Alpes, Grenoble, France. His research interests include signal/image processing and analysis, hyperspectral remote sensing, machine/deep learning, artificial intelligence, and their applications in Earth vision.

Dr. Hong is an Editorial Board Member of *Remote Sensing*. He received the recognition of the Best Reviewer Award of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2020 and the first outstanding paper award named the Jose Bioucas Dias award of the Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS) in 2021. He is also a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING and *Remote Sensing*. He has been serving as a Topical Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING since 2021.



Lin Xu received the Ph.D. degree from Xi'an Jiaotong University, Xi'an, China, in 2018.

Upon two years of academic experience as a Post-Doctoral Researcher in electrical and computer engineering at New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, and one-year industry experience as a Senior Researcher of the Institute of Advanced Artificial Intelligence, Horizon Robotics, Beijing, China. He joined Shanghai Em-Data Technology Company, Ltd., Shanghai, China, as the Director of the Artificial Intelligence Institute. He is also working on machine learning models/algorithms, deep learning techniques, 2-D/3-D data analysis, and high-level computer vision recognition, e.g., classification, retrieval, segmentation, detection, and their applications in the security system, traffic control, and automatic drive.



Deyu Meng (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees from Xi'an Jiaotong University, Xi'an, China, in 2001, 2004, and 2008, respectively.

From 2012 to 2014, he took his two-year sabbatical leave at Carnegie Mellon University, Pittsburgh, PA, USA. He is a Professor with School of Mathematics and Statistics, Xi'an Jiaotong University, and an Adjunct Professor with the Faculty of Information Technology, Macau University of Science and Technology, Macau. His research interests include

model-based deep learning, variational networks, and metalearning.



Jocelyn Chanussot (Fellow, IEEE) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995, and the Ph.D. degree from the Université de Savoie, Annecy, France, in 1998.

Since 1999, he has been with Grenoble INP, where he is a Professor of signal and image processing. He has been a Visiting Scholar with Stanford University, Stanford, CA, USA, the KTH Royal Institute of Technology, Stockholm, Sweden, and the National

University of Singapore (NUS), Singapore. Since 2013, he has been an Adjunct Professor with the University of Iceland, Reykjavik, Iceland. From 2015 to 2017, he was a Visiting Professor with the University of California at Los Angeles (UCLA), Los Angeles, CA, USA. He holds the AXA Chair in remote sensing and is an Adjunct Professor with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. His research interests include image analysis, hyperspectral remote sensing, data fusion, machine learning, and artificial intelligence.

Dr. Chanussot was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society from 2006 to 2008. He was a member of the Institut Universitaire de France from 2012 to 2017 and a Highly Cited Researcher (Clarivate Analytics/Thomson Reuters) from 2018 to 2019. He is also the Founding President of the IEEE Geoscience and Remote Sensing French Chapter from 2007 to 2010, which received the 2010 IEEE GRS-S Chapter Excellence Award. He has received multiple outstanding paper awards. He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Chair and the Co-Chair of the GRS Data Fusion Technical Committee from 2009 to 2011 and 2005 to 2008, respectively. He was the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing in 2009. He was the Vice-President of the IEEE Geoscience and Remote Sensing Society, in charge of meetings and symposia, from 2017 to 2019. He was the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING from 2011 to 2015. In 2014, he has served as a Guest Editor for the *IEEE Signal Processing Magazine*. He is also an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE TRANSACTIONS ON IMAGE PROCESSING, and the PROCEEDINGS OF THE IEEE.



Zongben Xu (Member, IEEE) received the Ph.D. degree in mathematics from Xi'an Jiaotong University, Xi'an, China, in 1987.

He is the Director of the Institute for Information and System Sciences, Xi'an Jiaotong University. His research interests include intelligent information processing and applied mathematics.

Dr. Xu was elected as a member of the Chinese Academy of Science in 2011. He was a recipient of the National Natural Science Award of China in 2007 and the Winner of the CSIAM Su Buchin

Applied Mathematics Prize in 2008. He delivered a speech at the International Congress of Mathematicians in 2010. He also serves as the Chief Scientist for the National Basic Research Program of China (973 Project).