LEARNING A LOW-COHERENCE DICTIONARY TO ADDRESS SPECTRAL VARIABILITY FOR HYPERSPECTRAL UNMIXING

Danfeng Hong[⋆†], *Naoto Yokoya*[⋆†§], *Jocelyn Chanussot*[‡], *Xiao Xiang Zhu*[⋆†]

[⋆]Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Munich, Germany
[†]Signal Processing in Earth Observation (SiPEO), Technical University of Munich, Munich, Germany
[§]Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan
[‡] GIPSA-Lab, CNRS, University of Grenoble Alpes, Grenoble, France

ABSTRACT

This paper presents a novel spectral mixture model to address spectral variability in inverse problems of hyperspectral unmixing. Based on the linear mixture model (LMM), our model introduces a spectral variability dictionary to account for any residuals that cannot be explained by the LMM. Atoms in the dictionary are assumed to be low-coherent with spectral signatures of endmembers. A dictionary learning technique is proposed to learn the spectral variability dictionary while solving unmixing problems simultaneously. Experimental results on synthetic and real datasets demonstrate that the performance of the proposed method is superior to state-of-the-art methods.

Index Terms— Remote sensing, spectral unmixing, spectral variability, low-coherent dictionary learning, alternating direction method of multipliers.

1. INTRODUCTION

Hyperspectral unmixing is a procedure that decomposes the measured pixel spectrum of hyperspectral data into a collection of constituent spectral signatures (or *endmembers*) and a set of corresponding fractional abundances. Hyperspectral unmixing techniques have been widely used for a variety of applications [1], such as mineral mapping, vegetation mapping, and land-cover change detection.

Spectral variability is intra-class variability due to various factors, including illumination conditions and topography, atmospheric conditions, and intrinsic variability of materials. Variations of spectral signatures for a material can result in significant errors in hyperspectral unmixing.

This work was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No [ERC-2016-StG-714087]) and from Helmholtz Association under the framework of the Young Investigators Group "SiPEO" (VH-NG-1018, www.sipeo.bgu.tum.de), as well as partially supported by ANR ASTRID (project APHYPIS) under grant ANR-16-ASTR-0027-01. The work of N. Yokoya was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI 15K20955 and Alexander von Humboldt Fellowship for postdoctoral researchers.

Quite recently, considerable attention has been paid to dealing with spectral variability in hyperspectral unmixing [2–4]. In the literature, several theories have been proposed to model the spectral variability. In [5] and [6], the normal compositional model and beta compositional model were designed by assuming that the spectral variability is fitting a given probability distribution. Obviously, the spectral variability in a certain scene can hardly be modeled by giving an explicit distribution in reality. Thouvenin *et al.* [7] indicated that the spectral variability can be represented using a perturbed linear mixing model (PLMM), where the variability is explained by an additive perturbation term for each endmember. One drawback of this model is the lack of physical meaning. For instance, as a principal spectral variability, scaling factors should be coherent with endmember spectral signatures, while other variabilities are often incoherent with endmember spectral signatures. Intuitively, such attributed spectral variability is not able to be represented by an additional term. Another interesting approach, namely the extended linear mixing mode (ELMM), has been proposed in [8]. This work mainly focuses on modeling the scaling factors on the endmembers, yet other spectral variabilities are not be considered.

To overcome the limitations of PLMM and ELMM, in this work we model the scaling factors and other spectral variability simultaneously according to their distinctive property. More specifically, our contributions can be summarized as follows: 1) a novel spectral mixture model is proposed, namely, scaling factors are modeled on the endmember dictionary and an additional dictionary is introduced to model the rest of spectral variabilities; 2) a data-driven dictionary learning method is explored in the proposed framework by giving a reasonable property assumption, namely spectral variability except for scaling factors should be low-coherent with endmember spectral signatures; 3) an alternating iterative optimization strategy is designed to solve the proposed model.

The remainder of this paper is organized as follows. Section II describes the methodology containing the proposed spectral mixture model and the optimization algorithm. Sec-

tion III presents the experimental results and corresponding analysis. Finally, Section IV concludes with a summary.

2. METHODOLOGY

In this section, we design a new framework for spectral unmixing in order to represent the spectral variability by a learned low-coherence dictionary.

2.1. Proposed Spectral Mixture Modeling

We first introduce our spectral mixture modeling as

$$\mathbf{Y} = \mathbf{AXS} + \mathbf{EB} + \mathbf{R}, \tag{1}$$

where $\mathbf{Y} = [\mathbf{Y}_1, ..., \mathbf{Y}_k, ..., \mathbf{Y}_N] \in \mathbb{R}^{D\times N}$ stands for an observed hyperspectral image with D bands and N pixels, and $\mathbf{A} = [\mathbf{A}_1, ..., \mathbf{A}_P] \in \mathbb{R}^{D\times P}$ and $\mathbf{E} = [\mathbf{E}_1, ..., \mathbf{E}_L] \in \mathbb{R}^{D\times L}$ denote endmembers and spectral variability matrices (or dictionaries), respectively. P is the number of endmembers and L is the number of basis vectors in $\mathbf{E}$. $\mathbf{X} = [\mathbf{X}_1, ..., \mathbf{X}_k, ..., \mathbf{X}_N] \in \mathbb{R}^{P\times N}$ is the abundance map with each column vector representing the abundance vector at each pixel and $\mathbf{B} = [\mathbf{B}_1, ..., \mathbf{B}_k, ..., \mathbf{B}_N] \in \mathbb{R}^{L\times N}$ is the coefficient matrix corresponding to $\mathbf{E}$. $\mathbf{S} \in \mathbb{R}^{N\times N}$ is a diagonal matrix with the non-negativity constraint on the diagonal elements ($\mathbf{S}_k \geq 0$), assuming the same scaling factor for all endmembers at each pixel. $\mathbf{R} \in \mathbb{R}^{D\times N}$ is the residual matrix.

Different from ELMM, the spectral variabilities that can not be explained by scaling are represented by spectral variability term, i.e., $\mathbf{EB}$.

2.2. Optimization Overview

The inverse problem of the model (1) is nonconvex with many variables to be estimated even when the endmember matrix is given. Therefore, it is important to use regularization based on reasonable assumptions with an alternating optimization strategy. Our optimization procedure can be divided into two parts: dictionary learning and spectral unmixing.

Based on the model in Eq. (1), we consider the following optimization problem to learn the spectral variability dictionary while solving spectral unmixing simultaneously

$$\begin{aligned}\min_{\mathbf{X},\mathbf{B},\mathbf{S},\mathbf{E}} &\|\mathbf{Y} - \mathbf{AXS} - \mathbf{EB}\|_F^2 + \alpha\|\mathbf{X}\|_{1,1} + \beta\|\mathbf{B}\|_F^2 \\ &+ \eta\|\mathbf{A}^T\mathbf{E}\|_F^2 + \tau\|\mathbf{E}^T\mathbf{E} - \mathbf{I}\|_F^2, \\ \text{s.t.}\quad &\mathbf{X} \succeq \mathbf{0}, \quad \mathbf{S} \succeq \mathbf{0}.\end{aligned} \tag{2}$$

where $\|\mathbf{X}\|_{1,1} \equiv \sum_{k=1}^{N}\|\mathbf{X}_k\|_1$ [9]. All elements of $\mathbf{X}$ and $\mathbf{S}$ are nonnegative due to physical conditions. The two regularization terms $\|\mathbf{X}\|_{1,1}$ and $\|\mathbf{B}\|_F^2$ are applied to the abundance and coefficient matrices based on two reasonable assumptions: 1) a given spectral signature is composed of a limited number of materials (sparsely represented by endmembers); 2) the spectral variability can be various due to its complex structure (many atoms in $\mathbf{E}$ should be activated). The other two regularization terms added to $\mathbf{E}$ can be explained as follows: 1) the first term $\|\mathbf{A}^T\mathbf{E}\|_F^2$ means that the spectral variability dictionary should be incoherent with the endmember dictionary because most of spectral variabilities coherent with endmembers are explained by scaling factors; 2) the other term $\|\mathbf{E}^T\mathbf{E} - \mathbf{I}\|_F^2$ has a role to make the basis vectors of $\mathbf{E}$ as orthogonal as possible while getting rid of the trivial solution. α, β, η, and τ are all penalty parameters corresponding to the different regularization terms of Eq. (2), respectively.

Once $\mathbf{E}$ has been learned by the dictionary learning step, the pixel-wised unmixing can be performed as follows

$$\begin{aligned}\min_{\mathbf{X}_k,\mathbf{B}_k,\mathbf{S}_k} &\|\mathbf{Y}_k - \mathbf{S}_k\mathbf{AX}_k - \mathbf{EB}_k\|_2^2 + \alpha\|\mathbf{X}_k\|_1 + \beta\|\mathbf{B}_k\|_2^2, \\ \text{s.t.}\quad &\mathbf{X}_k \succeq \mathbf{0}, \quad \mathbf{S}_k \geq 0.\end{aligned} \tag{3}$$

2.3. Dictionary Learning

To solve the optimization problem (2), multiple auxiliary variables $\mathbf{G}$, $\mathbf{H}$, $\mathbf{M}$, and $\mathbf{Q}$ are introduced to replace $\mathbf{X}$, $\mathbf{X}^+$, $\mathbf{XS}$,

Algorithm 1 Spectral Variability Dictionary Learning

Input: $\mathbf{Y}$, $\mathbf{A}$.
Output: $\mathbf{X}$, $\mathbf{B}$, $\mathbf{S}$, $\mathbf{E}$.
Initialize $\mathbf{X}, \mathbf{S}, \mathbf{B}, \mathbf{E}, \mathbf{G}, \mathbf{H}, \mathbf{M}, \mathbf{Q}, \alpha, \beta, \eta\, \tau, \mu, \boldsymbol{\lambda}_i (i = 1, ..., 4)$
while not converged or $t > maxIter$ **do**
$t = t + 1$
1. fix other variables and update $\mathbf{M}$ by
$$\hat{\mathbf{M}} = (\mathbf{A}^T\mathbf{A} + \mu\mathbf{I})^{-1}(\mathbf{A}^T\mathbf{Y} - \mathbf{A}^T\mathbf{EB} + \mu\mathbf{XS} - \boldsymbol{\lambda}_3)$$
2. fix other variables and update $\mathbf{B}$ by
$$\hat{\mathbf{B}} = (\mathbf{E}^T\mathbf{E} + \beta\mathbf{I})^{-1}(\mathbf{E}^T\mathbf{Y} - \mathbf{E}^T\mathbf{AM})$$
3. fix other variables and update $\mathbf{E}$ by
$$\hat{\mathbf{E}} = \arg\min_{\mathbf{E}} \frac{1}{2}\|\mathbf{Y} - \mathbf{AM} - \mathbf{EB}\|_F^2 + \frac{\eta}{2}\|\mathbf{A}^T\mathbf{E}\|_F^2 + \frac{\tau}{2}\|\mathbf{E}^T\mathbf{E} - \mathbf{I}\|_F^2$$
4. fix other variables and update $\mathbf{X}$ by
$$\hat{\mathbf{X}} = \mu(\mu\mathbf{G} + \boldsymbol{\lambda}_1 + \mu\mathbf{H} + \boldsymbol{\lambda}_2 + \boldsymbol{\lambda}_3\mathbf{S}^T + \mu\mathbf{MS}^T)(\mathbf{SS}^T + 2\mathbf{I})^{-1}$$
$$\hat{\mathbf{X}} = \hat{\mathbf{X}}(diag(\mathbf{1}^T\hat{\mathbf{X}}))^{-1}$$
5. fix other variables and update $\mathbf{S}$ by
$$\hat{\mathbf{S}} = \mu(\mathbf{X}_T\mathbf{X} + \mathbf{I})^{-1}(\mu\mathbf{X}_T\mathbf{M} + \mathbf{X}_T\boldsymbol{\lambda}_3 + \mu\mathbf{Q} + \boldsymbol{\lambda}_4)$$
6. fix other variables and update $\mathbf{G}$ by
$$\hat{\mathbf{G}} = max\{0, |\mathbf{X} - \boldsymbol{\lambda}_1/\mu| - (\alpha/\mu)\}\, sign(\mathbf{X} - \boldsymbol{\lambda}_1/\mu)$$
7. fix other variables and update $\mathbf{H}$ by
$$\hat{\mathbf{H}} = max\{0, \mathbf{X} - \boldsymbol{\lambda}_2/\mu\}$$
8. fix other variables and update $\mathbf{Q}$ by
$$\hat{\mathbf{Q}} = max\{0, \mathbf{S} - \boldsymbol{\lambda}_4/\mu\}$$
9. update multipliers by
$$\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_1 + \mu(\mathbf{G} - \mathbf{X}), \quad \boldsymbol{\lambda}_2 = \boldsymbol{\lambda}_2 + \mu(\mathbf{H} - \mathbf{X}),$$
$$\boldsymbol{\lambda}_3 = \boldsymbol{\lambda}_3 + \mu(\mathbf{M} - \mathbf{XS}), \quad \boldsymbol{\lambda}_4 = \boldsymbol{\lambda}_4 + \mu(\mathbf{Q} - \mathbf{S}).$$
10. update μ by $\mu = min(\rho\mu, \mu_{max})$
11. check the convergence conditions:
$$\|\mathbf{G} - \mathbf{X}\|_F^2 < \varepsilon; \|\mathbf{H} - \mathbf{X}\|_F^2 < \varepsilon;$$
$$\|\mathbf{M} - \mathbf{XS}\|_F^2 < \varepsilon; \|\mathbf{Q} - \mathbf{S}\|_F^2 < \varepsilon;$$
end while

and $\mathbf{S}^+$, respectively, where $()^+$ denotes an operator that converts each component of the matrix to its absolute value. The problem (2) can be converted to the following unconstrained optimization problem by the alternating direction method of multipliers (ADMM) [10]:

$$\begin{aligned}&\min_{\mathbf{X},\mathbf{B},\mathbf{S},\mathbf{E},\mathbf{G},\mathbf{H},\mathbf{M},\mathbf{Q},\boldsymbol{\lambda}} \|\mathbf{Y}-\mathbf{AM}-\mathbf{EB}\|_F^2+\alpha\|\mathbf{G}\|_{1,1}+\beta\|\mathbf{B}\|_F^2\\&+\mu\|\mathbf{M}-\mathbf{XS}+\frac{\boldsymbol{\lambda}_3}{\mu}\|_F^2+\mu\|\mathbf{Q}-\mathbf{S}+\frac{\boldsymbol{\lambda}_4}{\mu}\|_F^2+\eta\|\mathbf{A}^T\mathbf{E}\|_F^2\\&+\mu\|\mathbf{G}-\mathbf{X}+\frac{\boldsymbol{\lambda}_1}{\mu}\|_F^2+\mu\|\mathbf{H}-\mathbf{X}+\frac{\boldsymbol{\lambda}_2}{\mu}\|_F^2+\tau\|\mathbf{E}^T\mathbf{E}-\mathbf{I}\|_F^2\end{aligned} \tag{4}$$

where $\boldsymbol{\lambda}_i$ $(i = 1, ..., 4)$ are Lagrange multipliers, and μ is the penalty parameter. The object function in Eq. (4) is not convex w.r.t. all variables simultaneously, but it is a convex problem regarding the separate variable when other variables are fixed. Therefore, we can update these variables by alternating optimization. Algorithm 1 presents more specific procedures. Note that we optimize **step** 3 by derivating $\mathbf{E}$ as $\eta\mathbf{A}\mathbf{A}^T\mathbf{E}_t+\tau\mathbf{E}_{t-1}\mathbf{E}_{t-1}^T\mathbf{E}_t+\mathbf{E}_t\mathbf{J}\mathbf{J}^T=(\mathbf{Y}-\mathbf{AM})\mathbf{J}^T+\tau\mathbf{E}_{t-1}$, which is a classic Sylvester equation.

2.4. Pixel-Wise Spectral Unmixing

Unlike the dictionary learning step that considers all samples during the optimization, spectral unmixing is performed for each pixel. Consequently, the problem (3) can be reformulated as the following unconstrained optimization problem by introducing auxiliary variables $\mathbf{G}_k, \mathbf{H}_k$

$$\begin{aligned}&\min_{\mathbf{X}_k,\mathbf{B}_k,\mathbf{S}_k,\mathbf{G}_k,\mathbf{H}_k,\boldsymbol{\phi}} \|\mathbf{Y}_k-\mathbf{S}_k\mathbf{A}\mathbf{X}_k-\mathbf{E}\mathbf{B}_k\|_2^2+\alpha\|\mathbf{G}_k\|_1+\beta\|\mathbf{B}_k\|_2^2\\&+\zeta\|\mathbf{G}_k-\mathbf{X}_k+\frac{\boldsymbol{\phi}_1}{\zeta}\|_2^2+\zeta\|\mathbf{H}_k-\mathbf{X}_k+\frac{\boldsymbol{\phi}_2}{\zeta}\|_2^2\end{aligned} \tag{5}$$

where $\boldsymbol{\phi}_1$ $(i = 1, 2)$ are Lagrange multipliers, and ζ is the penalty parameter. Algorithm 2 summarizes the solution for the problem (5). Inspired by the idea of scaled constrained least squares unmixing (SCLSU) [11], $\mathbf{X}_k$ and $\mathbf{S}_k$ can be estimated simultaneously as follows:

- we first solve the subproblem with respect to $\mathbf{X}_k$;
- and then normalize $\mathbf{X}_k$ to satisfy abundance sum-to-one constraint in each iteration, as formulated by $\mathbf{X}_k = \mathbf{X}_k/\mathbf{1}^T\mathbf{X}_k$;
- finally we solve the nonnegative least squares problem [12] for $\mathbf{S}_k$:

$$\hat{\mathbf{S}}_k = arg\min_{\mathbf{S}_k\succeq\mathbf{0}}\|\mathbf{Y}_k-\mathbf{S}_k\mathbf{A}\mathbf{X}_k-\mathbf{E}\mathbf{B}_k\|_2^2. \tag{6}$$

Algorithm 2 Pixel-Wise Spectral Unmixing

Input: $\mathbf{Y}_k$, $\mathbf{A}$, $\mathbf{E}$.
Output: $\mathbf{X}_k, \mathbf{B}_k, \mathbf{S}_k$
Initialize $\mathbf{X}_k, \mathbf{S}_k, \mathbf{B}_k, \mathbf{G}_k, \mathbf{H}_k, \alpha, \beta, \zeta, \boldsymbol{\phi}_i (i = 1, 2)$.
while not converged or $t > maxIter$ **do**
$t = t + 1$
1. fix other variables and update $\mathbf{B}_k$ by
$$\hat{\mathbf{B}}_k = (\mathbf{E}^T\mathbf{E}+\beta\mathbf{I})^{-1}(\mathbf{E}^T\mathbf{Y}-\mathbf{E}^T\mathbf{S}_k\mathbf{A}\mathbf{X}_k)$$
2. fix other variables and update $\mathbf{X}_k$ and $\mathbf{S}_k$ by
$$\hat{\mathbf{X}}_k = (\mathbf{A}^T\mathbf{A}+2\zeta\mathbf{I})^{-1}(\zeta\mathbf{G}_k+\boldsymbol{\phi}_1+\zeta\mathbf{H}_k+\boldsymbol{\phi}_2+\mathbf{A}^T\mathbf{Y}-\mathbf{A}^T\mathbf{E}\mathbf{B}_k)$$
$$\hat{\mathbf{X}}_k = \hat{\mathbf{X}}_k/\mathbf{1}^T\hat{\mathbf{X}}_k$$
$$\hat{\mathbf{S}}_k = arg\min_{\mathbf{S}_k\succeq\mathbf{0}}\|\mathbf{Y}_k-\mathbf{S}_k\mathbf{A}\mathbf{X}_k-\mathbf{E}\mathbf{B}_k\|_2^2$$
3. fix other variables and update $\mathbf{G}_k$ by
$$\hat{\mathbf{G}}_k = max\{0, |\mathbf{X}_k-\boldsymbol{\phi}_1/\zeta| - (\alpha/\zeta)\}\, sign(\mathbf{X}_k-\boldsymbol{\phi}_1/\zeta)$$
4. fix other variables and update $\mathbf{H}_k$ by
$$\hat{\mathbf{H}}_k = max\{0, \mathbf{X}_k-\boldsymbol{\phi}_2/\zeta\}$$
5. update multipliers by
$$\boldsymbol{\phi}_1 = \boldsymbol{\phi}_1+\zeta(\mathbf{G}_k-\mathbf{X}_k), \quad \boldsymbol{\phi}_2 = \boldsymbol{\phi}_2+\zeta(\mathbf{H}_k-\mathbf{X}_k)$$
6. update ζ by $\zeta = min(\rho\zeta, \zeta_{max})$
7. check the convergence conditions:
$$\|\mathbf{G}_k-\mathbf{X}_k\|_2<\varepsilon; \|\mathbf{H}_k-\mathbf{X}_k\|_2<\varepsilon;$$
end while

3. EXPERIMENTS

In this section, we evaluate the performance of the proposed method using two datasets: a synthetic dataset presented in [8] and a real dataset over the Cuprite mining district. We compare the proposed method with constrained least squares unmixing (CLSU), fully CLSU (FCLSU) [13], SCLSU, CLSU+ℓ_1-norm, SCLSU+ℓ_1-norm as well as ELMM.

3.1. Synthetic Data

3.1.1. Data Description and Experimental Setup

The first synthetic data was simulated using five reference endmembers (224 spectral bands) randomly selected from United States Geological Survey (USGS) spectral library and 200×200 abundance maps generated using Gaussian fields. The dataset was designed to comply the abundance sum-to-one constraint (ASC). It should be noted that a spectral signature of each pixel in this dataset includes spectral variability due to endmember-dependent scaling factors and Gaussian mixture noise. We refer the readers to [8] for more details.Vertex component analysis (VCA) [14] is used as the endmember extraction algorithm to obtain the endmember dictionary ($\mathbf{A}$) for all algorithms under comparison, whereas Hysime method [15] is used to estimate the number of endmembers.

3.1.2. Algorithm Performance Evaluation

We use three criteria for assessing the performance of the algorithms by referring to [8]: abundance overall root mean

Table 1: Performance comparison with the different algorithms on the synthetic data

Algorithm	FCLSU	CLSU	CLSU+ℓ_1-norm	SCLSU	SCLSU+ℓ_1-norm	ELMM	Ours
aRMSE	0.0524	0.0380	0.0376	0.0251	0.0247	0.0337	0.0206
rRMSE	0.0151	0.0123	0.0127	0.0123	0.0127	0.0088	0.000006
aSAM	1.9600	1.7713	1.7715	1.7713	1.7715	1.2998	0.0007
Running Time (s)	14	6	20	7	15	366	370

square error (aRMSE), reconstruction overall root mean square error (rRMSE), and average spectral angle mapper (aSAM).

3.1.3. Results and Analysis

Table 1 lists the quantitative assessment results based on Section 3.1.2 for all algorithms. From Table 1, we can see that FCLSU shows poor performance due to the existence of spectral variability. While the performance of CLSU is better than that of FCLSU, since the abundances can be reasonably estimated in a cone not in a simplex by dropping the ASC. However, the spectral variability is not actually eliminated by CLSU, but absorbed by the abundances. By comparison, S-CLSU performs better than CLSU, particularly being robust against scaling factors. Notably, with the sparsity term, CLSU+ℓ_1-norm and S-CLSU+ℓ_1-norm can obtain better performance than those without the sparsity term, which experimentally explains that each pixel in the studied hyperspectral scene consists of a few materials. Although the ELMM approach can model endmember-dependent spectral variability, its performance is limited without consideration of spatial information as reported in [8]. The proposed method outperforms the compared methods, which implies that our method is capable of modeling the spectral variability, improving the accuracy of the abundances estimation.

3.2. Real Data

The second dataset is the hyperspectral image acquired by the airborne visible-infrared imaging spectrometer (AVIRIS) over the Cuprite mining district. We use its $304 \times 257 \times 224$ sub-image. Different from the simulated data, the real data often include complex and various spectral variability. Therefore, it is relatively difficult to obtain the pure endmembers. We first used VCA for extracting endmembers. Next, material identification was performed using USGS spectral library and spectral feature fitting [16]. Finally, the endmember dictionary ($\mathbf{A}$) was constructed based on identified spectral signatures from the library. We use the same endmember matrix for all algorithms under comparison. Fig. 1 shows the estimated abundance maps. Reference classification maps were generated by Tetracorder [17] software. From Fig. 1, we observe that the proposed method shows the best visual resemblance compared with the results from Tetracorder. The abundance maps generated by our method are more distinct and more contrastive and the distribution of each material is regional as well, which implies that various spectral variabilities could be learned effectively.

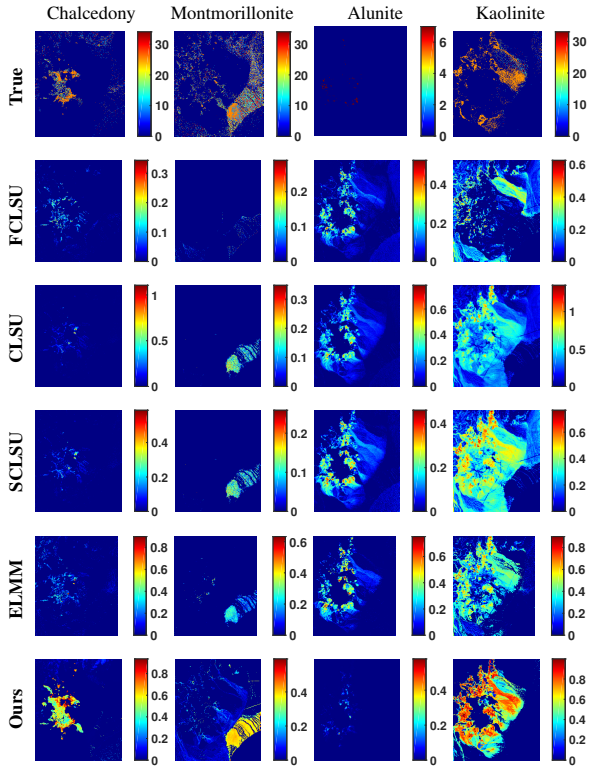

Fig. 1: The abundances estimated by five different methods (each column corresponds to one endmember extracted by VCA) and the first row shows the ground truth.

4. CONCLUSION

Based on the ELMM framework, we proposed a novel spectral mixture model considering not only the principal scaling factor but also other various spectral variabilities by introducing the spectral variability dictionary. We modeled the spectral variability as low-coherent with the endmember spectral signatures and developed an algorithm for learning the spectral variability dictionary. By analyzing experimental results, it is evident that the proposed method is prone to obtain a more accurate abundance estimation compared to other state-of-the-art algorithms.

5. REFERENCES

[1] J. M. Bioucas-Dia, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "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, Feb. 2012.

[2] A. Zare and K. C. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Process. Magazine*, vol. 31, no. 1, pp. 95–104, 2014.

[3] L. Drumetz, J. Chanussot, and C. Jutten, "Variability of the endmembers in spectral unmixing: recent advances," in *Proc. IEEE WHISPERS*, 2016, pp. 1–5.

[4] J. Bieniarz, E. Aguilera, X. Zhu, R. Mueller, and P. Reinartz, "Joint sparsity model for multilook hyperspectral image unmixing," *IEEE Geosci. and Remote Sens. Lett.*, vol. 12, no. 4, pp. 696–700, 2015.

[5] O. Eches, N. Dobigeon, C. Mailhes, and J.Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1403–1413, Jun. 2010.

[6] X. Du, A. Zare, P. Gader, and D. Dranishnikov, "Spatial and spectral unmixing using the beta compositional model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1944–2003, Jun. 2014.

[7] 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, Jun. 2016.

[8] 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.

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

[10] J. M. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. IEEE Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*. IEEE, 2010, pp. 1–4.

[11] M. A. Veganzones, L. Drumetz, G. Tochon, M. Dalla Mura, A. Plaza, J. M. Bioucas-Dias, and J. Chanussot, "A new extended linear mixing model to address spectral variability," in *Proc. IEEE Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*. IEEE, 2014, pp. 1–5.

[12] D. Kim, S. Sra, and I. Dhillon, "Trackling box-constrained convex optimization via a new projected quasi-newton approach," *SIAM J. Sci. Comput.*, vol. 32, no. 6, pp. 3548–3563, Jun. 2010.

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

[14] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 4, pp. 898–910, Apr. 205.

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

[16] R. N. Clark and T. L. Roush, "Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications," *J. Geophys. Res.*, vol. 89, no. 7, pp. 6329–6340, 1984.

[17] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the usgs tetracorder and expert systems," *Journal of Geophysical Research: Planets*, vol. 108, no. E12, 2003.