A MULTIPLE ENDMEMBER MIXING MODEL TO HANDLE SPECTRAL VARIABILITY IN HYPERSPECTRAL UNMIXING

Tatsumi Uezato$^{(1)}$, *Mathieu Fauvel*$^{(2)}$ *and Nicolas Dobigeon*$^{(1)}$

$^{(1)}$ University of Toulouse, IRIT/INP-ENSEEIHT, 31071 Toulouse, France
$^{(2)}$ University of Toulouse, Dynafor/INP-ENSAT, 31326 Castanet-Tolosan, France
firstname.lastname@enseeiht.fr, mathieu.fauvel@ensat.fr

ABSTRACT

This paper proposes a novel mixing model that incorporates spectral variability. The proposed approach relies on the following two ingredients: *i*) a mixed spectrum is modeled as a combination of a few endmember signatures which belong to some endmember bundles (referred to as classes), *ii*) sparsity is promoted for the selection of both endmember classes and endmember spectra within a given class. This leads to an adaptive and hierarchical description of the endmember spectra. A proximal alternating linearized minimization algorithm is derived to minimize the objective function associated with this model, providing estimates of the bundling coefficients and abundances. Results showed that the proposed method outperformed the existing methods in terms of promoting sparsity and selecting endmember classes within each pixel.

Index Terms— Hyperspectral imagery, spectral unmixing, endmember variability, sparse unmixing, double sparsity

1. INTRODUCTION

Spectral unmixing is a technique that decomposes a mixed spectrum into a collection of *pure* spectra (i.e. *endmembers*) and their corresponding proportions (i.e. *abundances*). Although many studies have developed a great number of spectral unmixing methods, there are still major problems [1, 2]. One of the major problems is caused by spectral variability of each material class present in an image [3–5]. The spectral variability of each material class is defined, in this paper, as *endmember variability*.

Many spectral unmixing methods that incorporate endmember variability have been developed [6]. To describe this endmember variability, a wide and commonly used approach resorts to multiple endmember spectra within a given class of materials (i.e. *endmember bundles*). Endmember bundles can be collected from field investigation or from the image itself using automated endmember extraction methods, e.g., [7]. Multiple endmember spectral mixture analysis (MESMA) is one of the most successful methods that use endmember bundles [8]. MESMA, however, is computationally expensive and may be greatly degraded when endmember bundles do not completely represent *true* spectral variability within each class [3].

A new class of methods that incorporate all spectra belonging to endmember bundles simultaneously and estimate their corresponding abundances has been developed [3, 4, 9, 10]. The estimated multiple abundances are summed to generate a single abundance within each class. These methods are different from MESMA, which solves a combinatorial problem, and are computationally feasible. These methods, however, tend to select a large number of endmember spectra to unmix a given mixed spectrum, leading to multiple abundances corresponding to the selected endmember spectra. Sparsity inducing regularization (e.g., derived from ℓ_1-norm) has been used to select a fewer number of endmember spectra and generate a fewer number of abundances for each pixel [4]. The regularization, however, imposes sparsity on the selection of spectra, not the selection of classes. This shows that it ignores the structure of groups in endmember bundles. A few methods are designed to impose sparsity on the selection of classes [5, 11]. Although the aforementioned methods show great potential to achieve good performance in computationally feasible time, it has several constraints: *i*) it is difficult to interpret the physical meaning of the models because it first generate *unrealistic* multiple abundances and sum the multiple abundances for each class, *ii*) it lacks from flexibility to describe endmember variability, *iii*) the methods do not explicitly generate adaptive endmember spectra used for unmixing each pixel, like the model-driven methods proposed in [12, 13].

The paper proposes a novel unmixing method that addresses the aforementioned problems. The proposed method is inspired by a double sparsity-based method [14]. Indeed, it captures the hierarchical structure of each endmember class. It owns the major advantages of providing a physically meaningful model that is composed of endmember

Part of this work has been funded by EU FP7 through the ERANETMED JC-WATER program, MapInvPlnt Project ANR-15-NMED-0002-02 and by the MUESLI IDEX ATS project, Toulouse INP.

bundles, bundling coefficients (which relate the spectra to the endmember classes) and abundances. Moreover, it generate adaptive endmember spectra incorporating hierarchical structure for unmixing each pixel. The proposed method is compared to the state-of-the-art methods using simulated and real hyperspectral data.

2. MULTIPLE ENDMEMBER UNMIXING

2.1. Multiple endmember mixing model

The proposed model relies on the definition of 3 components representing endmember bundles, bundling coefficients and abundances. According to this model, each endmember bundle is mixed to provide a suitable and adaptive endmember spectrum used to unmix a given pixel. The proposed multiple endmember mixing model (MEMM) is defined as follows:

$$\mathbf{y}_i = \mathbf{E}\mathbf{B}_i\mathbf{a}_i + \mathbf{n}_i \tag{1}$$

where $\mathbf{y}_i \in \mathbb{R}^{L\times 1}$ is the mixed L-spectrum of the ith pixel, $\mathbf{E} \in \mathbb{R}^{L\times N}$ is composed of N distinct spectral signatures representing endmember bundles, $\mathbf{B}_i \in \mathbb{R}^{N\times K}$ gathers so-called bundling coefficients which decompose the endmember signatures according to the endmember bundles for the considered pixel, $\mathbf{a}_i = [a_{1i}, \dots, a_{Ki}]^T \in \mathbb{R}^{K\times 1}$ is the abundance fractions at the pixel, $\mathbf{n}_i \in \mathbb{R}^{L\times 1}$ represents noise in the pixel, L is the number of bands and K is the number of endmember classes. The endmember bundles $\mathbf{E}$ are defined as follows

$$\mathbf{E} = \left[\, \mathbf{E}_1 \mid \mathbf{E}_2 \mid \cdots \mid \mathbf{E}_K \,\right] \tag{2}$$

where $\mathbf{E}_k \in \mathbb{R}^{L\times N_k}$ represents a set of endmember spectra characterizing the kth class, N_k is the number of endmember spectra in the kth class and N is the total number of endmember spectra of all classes with $N = \sum_{k=1}^{K} N_k$. To enforce the bundle structure, the bundling coefficients $\mathbf{B}_i$ associated with the pixel is defined as the following block-diagonal matrix

$$\mathbf{B}_i = \begin{bmatrix} \mathbf{b}_{1i} & \mathbf{0}_{N_1} & \cdots & \mathbf{0}_{N_1} \\ \mathbf{0}_{N_2} & \mathbf{b}_{2i} & \cdots & \mathbf{0}_{N_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N_K} & \mathbf{0}_{N_K} & \cdots & \mathbf{b}_{Ki} \end{bmatrix} \tag{3}$$

where $\mathbf{b}_{ki} \in \mathbb{R}^{N_k}$ is the bundling coefficients for the kth class at the pixel and $\mathbf{0}_{N_k} \in \mathbb{R}^{N_k}$ is the N_k-dimensional vector whose components are zeros. Each bundling coefficient must be nonnegative and the bundling vector $\mathbf{b}_{ki}$ is expected to be sparse. Indeed multiple endmember spectra within each class are usually redundant and only a few endmember spectra within each class should be enough to unmix a pixel. This property can be induced by considering the following bundling constraints

$$\mathbf{B}_i \succeq 0 \quad \text{and} \quad \|\mathbf{B}_i\|_0 = \sum_{k=1}^{K} \|\mathbf{b}_{ki}\|_0 \le s \tag{4}$$

where $\|\cdot\|_0$ is the ℓ_0-norm that counts the number of nonzero elements and s is the maximum number of nonzero elements in $\mathbf{B}_i$, i.e., the maximum number of endmembers to be used within each class to describe the pixel. The abundance nonnegativity constraint (ANC) and the abundance sum-to-one constraint (ASC) are usually imposed. In addition, in this work, complementary sparsity is imposed on each abundance vector, i.e.,

$$\forall k, \forall i, a_{ki} \ge 0, \quad \text{and} \quad \sum_{k=1}^{K} a_{ki} = 1, \;\; \|\mathbf{a}_i\|_0 \le q \tag{5}$$

where q is the number of endmember classes to be used to decompose the image pixel.

2.2. Algorithm

Unmixing according to the proposed MEMM can be formulated as the minimization problem

$$\begin{aligned} &\min_{\mathbf{B}_i,\mathbf{a}_i} \frac{1}{2}\|\mathbf{E}\mathbf{B}_i\mathbf{a}_i - \mathbf{y}_i\|_2^2 \\ &\text{s.t. } \forall k, \forall i,\, a_{ki} \ge 0, \quad \sum_{k=1}^{K} a_{ki} = 1, \quad \|\mathbf{a}_i\|_0 \le q, \\ &\mathbf{B}_i \succeq 0, \quad \|\mathbf{B}_i\|_0 \le s. \end{aligned} \tag{6}$$

This minimization problem is similar to the double sparsity-inducing method proposed in [14]. Using an alternative formulation, the minimization problem can be written as the following non-convex minimization problem:

$$\min_{\mathbf{B}_i,\mathbf{a}_i} \left\{ \mathcal{J}\left(\mathbf{B}_i, \mathbf{a}_i\right) = f(\mathbf{B}_i, \mathbf{a}_i) + h(\mathbf{B}_i) + g(\mathbf{a}_i) \right\}$$

with

$$f(\mathbf{B}_i, \mathbf{a}_i) = \tfrac{1}{2}\|\mathbf{E}\mathbf{B}_i\mathbf{a}_i - \mathbf{y}_i\|_2^2 \tag{7}$$

$$h(\mathbf{B}_i) = \iota_{\mathbb{R}_+}(\mathbf{B}_i) + \lambda_b\|\mathbf{B}_i\|_0 \tag{8}$$

$$g(\mathbf{a}_i) = \iota_{\mathbb{S}}(\mathbf{a}_i) + \lambda_a\|\mathbf{a}_i\|_0 \tag{9}$$

where λ_a and λ_b are parameters which control the balance between the data fit and the sparsity, $\iota_{\mathcal{C}}(\mathbf{x})$ is the indicator function on the set C, i.e., $\iota_{\mathcal{C}}(\mathbf{x}) = 0$ when $\mathbf{x} \in \mathcal{C}$ whereas $\iota_{\mathcal{C}}(\mathbf{x}) = \infty$ when $\mathbf{x} \notin \mathcal{C}$, and $\mathbb{S}$ is the set defined by the ASC and ANC. Solving this optimization problem is challenging since the the regularization functions h and g are nonconvex and nonsmooth. It can be tackled thanks to the proximal alternating linearized minimization (PALM) [15]. With guarantees to converge to a critical point, PALM iteratively updates the parameters $\mathbf{a}_i$ and $\mathbf{B}_i$ by alternatively minimizing the objective function with respect to (w.r.t.) these parameters, i.e., by solving the following proximal problems

$$
\begin{aligned}
\mathbf{B}_i^{(t+1)} &\in \min_{\mathbf{B}_i} h(\mathbf{B}_i) + \langle \mathbf{B}_i - \mathbf{B}_i^{(t)}, \nabla_{\mathbf{B}_i} f(\mathbf{B}_i^{(t)}, \mathbf{a}_i^{(t)}) \rangle \\
&\quad + \frac{c_t}{2} \|\mathbf{B}_i - \mathbf{B}_i^{(t)}\|_2^2 \\
\mathbf{a}_i^{(t+1)} &\in \min_{\mathbf{a}_i} g(\mathbf{a}_i) + \langle \mathbf{a}_i - \mathbf{a}_i^{(t)}, \nabla_{\mathbf{a}_i} f(\mathbf{B}_i^{(t+1)}, \mathbf{a}_i^{(t)}) \rangle \\
&\quad + \frac{d_t}{2} \|\mathbf{a}_i - \mathbf{a}_i^{(t)}\|_2^2
\end{aligned} \tag{10}
$$

Optimization w.r.t. $\mathbf{B}_i$: To optimize only w.r.t. the diagonal entries in $\mathbf{B}_i$, the objective function can be rewritten with the following decomposition

$$
\begin{aligned}
f(\mathbf{b}_i, \mathbf{a}_i) &= \frac{1}{2} \|\mathbf{U}_i \mathbf{b}_i - \mathbf{y}_i\|_2^2 \\
h(\mathbf{b}_i) &= \iota_{\mathbb{R}_+}(\mathbf{b}_i) + \lambda_b \|\mathbf{b}_i\|_0
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{U}_i &= [\mathbf{E}_1 \odot a_{1i} | \cdots | \mathbf{E}_K \odot a_{Ki}] \\
\mathbf{b}_i &= \left[\mathbf{b}_{1i}^T, \mathbf{b}_{2i}^T, \cdots, \mathbf{b}_{Ki}^T\right]^T.
\end{aligned}
$$

This leads to the following updating rule

$$
\min_{\mathbf{b}_i} h(\mathbf{b}_i) + \frac{c_t}{2} \|\mathbf{b}_i - (\mathbf{b}_i^{(t)} - \frac{1}{c_t} \nabla_{\mathbf{b}} f(\mathbf{b}_i^{(t)}, \mathbf{a}_i^{(t)}))\|_2^2
$$

where $\nabla_{\mathbf{b}_i} f(\mathbf{b}_i^{(t)}, \mathbf{a}_i^{(t)}) = \mathbf{U}_i^T (\mathbf{U}_i \mathbf{b}_i - \mathbf{y}_i)$. Using similar computations as in [15], this can be conducted as

$$
\mathbf{b}_i^{(t+1)} \in \mathrm{prox}_{c_t/\lambda_b}^{h} (\mathbf{b}_i^{(t)} - \frac{1}{c_t} \nabla_{\mathbf{b}_i} f(\mathbf{b}_i^{(t)}, \mathbf{a}_i^{(t)}))
$$

where $c_t = \gamma_m \|\mathbf{U}_i^T \mathbf{U}_i\|_{\mathrm{F}}$ represents a step size for each iteration. The proximal operator associated with f can be computed using the approach [15]. Finally, the bundling matrix $\mathbf{B}_i$ can be reconstructed as $\mathbf{B}_i = \mathrm{blkdiag}(\mathbf{b}_i)$ where $\mathrm{blkdiag}(\cdot)$ generates the block diagonal matrix $\mathbf{B}_i$ from the vector $\mathbf{b}_i$.

Optimization with respect to $\mathbf{a}_i$: To optimize w.r.t. $\mathbf{a}_i$, the objective function can be rewritten using the decomposition

$$
\begin{aligned}
f(\mathbf{B}_i, \mathbf{a}_i) &= \tfrac{1}{2} \|\mathbf{S}_i \mathbf{a}_i - \mathbf{y}_i\|_2^2 \\
g(\mathbf{a}_i) &= \iota_{\mathbb{S}}(\mathbf{a}_i) + \lambda_a \|\mathbf{a}_i\|_0
\end{aligned}
$$

where $\mathbf{S}_i = \mathbf{E}\mathbf{B}_i$. Thus, updating the abundance vector can be formulated as

$$
\min_{\mathbf{a}_i} g(\mathbf{a}_i) + \frac{d_t}{2} \left\| \mathbf{a}_i - \left(\mathbf{a}_i^{(t)} - \frac{1}{d_t} \nabla_{\mathbf{a}_i} f(\mathbf{B}_i^{(t+1)}, \mathbf{a}_i^{(t)}) \right) \right\|^2
$$

where $\nabla_{\mathbf{a}_i} f(\mathbf{B}_i^{(t+1)}, \mathbf{a}_i^{(t)}) = \mathbf{S}_i^T (\mathbf{S}_i \mathbf{a}_i - \mathbf{y}_i)$. Using the proximal operator, this can be written as

$$
\mathbf{a}_i^{(t+1)} \in \mathrm{prox}_{d_t/\lambda_a}^{g} \left(\mathbf{a}_i^{(t)} - \frac{1}{d_t} \nabla_{\mathbf{a}_i} f(\mathbf{B}_i^{(t+1)}, \mathbf{a}_i^{(t)}) \right) \tag{11}
$$

where $d_t = \gamma_a \|\mathbf{S}_i^T \mathbf{S}_i\|_{\mathrm{F}}$ represents a step size for each iteration. Moreover the proximal mapping associated with g can be performed using the method developed in [16]. The pseudocode for MEMM-based unmixing is shown in Algo. 1.

Algorithm 1 Algorithm for MEMM-based unmixing

1. **Input** : $\mathbf{y}_i, \mathbf{E}$
2. **Initialization**: $\mathbf{a}_i^{(0)}$ and $\mathbf{B}_i^{(0)}$.
3. **Main procedure**:
4. **while** the stopping criterion is not satisfied **do**
5. $\quad \mathbf{b}_i^{(t+1)} \leftarrow \mathrm{prox}_{c_t/\lambda_b}^{h}(\mathbf{b}_i^{(t)} - \frac{1}{c_t} \nabla_{\mathbf{b}_i} f(\mathbf{b}_i^{(t)}, \mathbf{a}_i^{(t)}))$
6. $\quad \mathbf{B}_i^{(t+1)} = \mathrm{blkdiag}(\mathbf{b}_i^{(t+1)})$
7. $\quad \mathbf{a}_i^{(t+1)} \leftarrow \mathrm{prox}_{d_t/\lambda_a}^{g}(\mathbf{a}_i^{(t)} - \frac{1}{d_t} \nabla_{\mathbf{a}_i} f(\mathbf{B}_i^{(t+1)}, \mathbf{a}_i^{(t)}))$
8. **end while**
9. **Output** : $\mathbf{a}_i^{(t+1)}, \mathbf{B}_i^{(t+1)}$

3. EXPERIMENTS

3.1. Experiments using simulated data

Simulated data: The performance of the proposed method has been evaluated thanks to simulated data generated as follows. First, $K = 5$ spectra have been selected from the USGS library to define the classes. Second, $N_k = 20$ ($k = 1, \ldots, K$) endmember spectra have been synthetically generated for each endmember class using the approach in [17]. Third, a number of endmember classes and a number of endmember spectra within each class have been randomly determined to define the mixture. Forth, a mixed spectrum has been generated by a linear combination of randomly selected endmember spectra within each selected class and randomly generated abundances. Finally, Gaussian random noise has been added to the generated mixed spectrum with a signal-to-noise ratio 40dB. This process has been repeated to generate a set of 100 mixed spectra.

Validation of methods: The proposed method has been compared with fully constrained least squares (FCLS) [18], sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL [18]) and methods based on sparse representations, namely, group lasso and elitist lasso [5]. Note that a single abundance within each class has been estimated by summing multiple abundances for each class in the methods for comparison. Abundances estimated by the methods have been validated using the following 3 different criteria. To evaluate the quality of the reconstruction, one defines the signal-to-reconstruction error (SRE) as [18]

$$
\mathrm{SRE} \equiv \mathbb{E}\left[\|\mathbf{a}\|_2^2\right] / \mathbb{E}\left[\|\mathbf{a} - \hat{\mathbf{a}}\|_2^2\right] \tag{12}
$$

where $\mathbf{a}$ is vectorized true abundances of all pixels, $\hat{\mathbf{a}}$ is vectorized estimated abundances of all pixels. To evaluate the sparsity level (SL) induced by the methods, one monitors [19, 20]

$$\mathrm{SL} \equiv \frac{1}{P}\sum_{i=1}^{P} \|\hat{\mathbf{a}}_i\|_0. \tag{13}$$

where $\hat{\mathbf{a}}_i$ is the estimated abundances of the ith pixel, P is the number of pixels. Finally, one considers the distance between the two actual and estimated supports [20, 21])

$$\mathrm{DIST} \equiv \frac{1}{P}\sum_{i=1}^{P} \frac{\max\left(|\mathcal{S}_i|, |\hat{\mathcal{S}}_i|\right) - |\mathcal{S}_i \cap \hat{\mathcal{S}}_i|}{\max\left(|\mathcal{S}_i|, |\hat{\mathcal{S}}_i|\right)}. \tag{14}$$

where $\mathcal{S}_i$ and $\hat{\mathcal{S}}_i$ are the true and estimated supports of the ith pixel, i.e., the sets of index of nonzero values in $\mathbf{a}_i$ and $\hat{\mathbf{a}}_i$ of the pixel, $|\mathcal{S}|$ represents the total number of elements in the set $\mathcal{S}$ and $\cap$ denotes the intersection operation. DIST aims at evaluating whether the methods correctly select the combination of endmember classes. Finally, parameters required for the methods have been empirically adjusted to reach the highest SRE.

Criteria	FCLS	SUNSAL	Group lasso	Elitist lasso	MEMM
SRE	25.805	25.7331	**26.2155**	25.7959	26.1051
SL	3.45	3.62	4.22	4.95	**2.35**
DIST	0.3208	0.3592	0.4353	0.5033	**0.0592**

Table 1: SRE, SL and DIST estimated by the 5 methods.

Results: SRE, SL and DIST have been calculated from abundances estimated from the 5 methods and are reported in Table 1. MEMM produces consistent SRE compared with other methods. However, MEMM significantly outperforms the state-of-the-art methods in terms of SL and DIST. This shows that MEMM could promote more sparsity than other methods while selecting more appropriate combination of endmember classes.

3.2. Experiment using real hyperspectral data

Validation of methods: A 40×40-pixel subset of the real Pavia University hyperspectral image has been considered to evaluate the relevance of the proposed MEMM. In this image, four spatially isolated materials are present and mixed pixels are located only on the boundary of these materials. Note that multiple endmember spectra are expected to coexist within each class. Endmember bundles extracted by the N-Dimensional visualizer in ENVI are shown in in Fig. 2 (first row). The image has been unmixed according to the proposed MEMM and the compared methods using these 4 pre-defined endmember bundles.

Results: Abundances estimated by MEMM and other methods are depicted in Fig. 1. MEMM estimates larger abundances in each spatially discrete region. Endmember bundles show more detailed spectral variability than the original

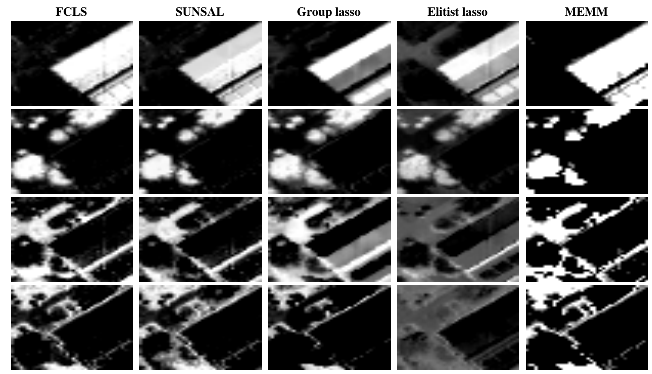

Fig. 1: Abundances estimated by the 5 methods. Brighter pixels show large abundances of a material while darker pixels show small abundances of a material

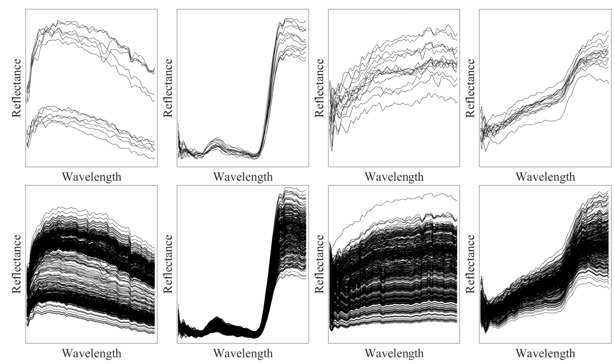

Fig. 2: First row represents endmember bundles used for unmixing. Second row represents synthetic endmember spectra generated by MEMM.

endmember bundles used for unmixing (see Fig. 2). This shows that MEMM successfully promotes sparsity to select endmember classes while generating adaptive endmember spectra within each class. This leads to more realistic abundance maps.

4. CONCLUSION

This paper proposed a novel spectral unmixing method that incorporates endmember variability. The proposed model has the following advantages compared to the existing methods: *i*) it generates physically realistic abundance fractions corresponding to each endmember class, *ii*) it generates adaptive endmember spectra for each pixel and *iii*) it captures a hierarchical structure of each endmember class. The method showed comparable results for estimating abundances while it outperformed other methods in terms of selecting a set of endmember classes within each pixel. Work is underway to develop a method to extract multiple endmembers within each class.

5. REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, D. Qian, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 5, no. 2, pp. 354–379, 2012.

[2] T. Uezato, M. Fauvel, and N. Dobigeon, "Hyperspectral image unmixing with LiDAR data-aided spatial regularization," *IEEE Trans. Geosci. Remote Sens.*, 2018.

[3] T. Uezato, R. J. Murphy, A. Melkumyan, and A. Chlingaryan, "A novel spectral unmixing method incorporating spectral variability within endmember classes," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2812–2831, 2016.

[4] M. A. Veganzones, L. Drumetz, G. Tochon, M. Dalla-Mura, A. J. Plaza, J. M. Bioucas Dias, and J. Chanussot, "A new extended linear mixing model to address spectral variability," in *Proc. IEEE GRSS Workshop Hyperspectral Image SIgnal Process.: Evolution in Remote Sens. (WHISPERS)*, 2014, pp. 1–4.

[5] T. R. Meyer, L. Drumetz, J. Chanussot, A. L. Bertozzi, and C. Jutten, "Hyperspectral unmixing with material variability using social sparsity," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2016, pp. 2187–2191.

[6] 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, 2014.

[7] T. Uezato, R. J. Murphy, A. Melkumyan, and A. Chlingaryan, "A novel endmember bundle extraction and clustering approach for capturing spectral variability within endmember classes," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6712–6731, 2016.

[8] 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. Environment*, vol. 65, no. 3, pp. 267–279, 1998.

[9] 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, 2016.

[10] M. A. Goenaga, M. C. Torres-Madronero, M. Velez-Reyes, S. J. Van Bloem, and J. D. Chinea, "Unmixing analysis of a time series of Hyperion images over the Guanica dry forest in Puerto Rico," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 6, no. 2, pp. 329–338, 2013.

[11] M. Iordache, J. Bioucas-Dias, and A. Plaza, "Hyperspectral unmixing with sparse group lasso," in *Proc. IEEE Int. Conf. Geosci. Remote Sens. (IGARSS)*, 2011, pp. 3586–3589.

[12] 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, 2016.

[13] 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, 2016.

[14] R. Rubinstein, M. Zibulevsky, and M. Elad, "Double sparsity: Learning sparse dictionaries for sparse signal approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1553–1564, 2010.

[15] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Mathematical Programming*, vol. 146, no. 1-2, pp. 459–494, 2014.

[16] K. Anastasios, B. Stephen, and C. Volkan, "Sparse projections onto the simplex," in *Proc. Int. Cong. Machine Learning*, 2013.

[17] P. A. Thouvenin, N. Dobigeon, and J. Y. Tourneret, "Online unmixing of multitemporal hyperspectral images accounting for spectral variability," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 3979–3990, 2016.

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

[19] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, 2014.

[20] F. Chen, K. Wang, and T. F. Tang, "Spectral unmixing using a sparse multiple-endmember spectral mixture model," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5846–5861, 2016.

[21] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer, 2010.