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 $\ell_1$-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
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:

\[ y_i = EB_i a_i + n_i \]  

where \( y_i \in \mathbb{R}^{L \times 1} \) is the mixed L-spectrum of the ith pixel, \( E \in \mathbb{R}^{L \times N} \) is composed of \( N \) distinct spectral signatures representing endmember bundles, \( B_i \in \mathbb{R}^{N \times K} \) gathers so-called bundling coefficients which decompose the endmember signatures according to the endmember bundles for the considered pixel, \( a_i = [a_{i1}, \ldots, a_{iK}]^T \in \mathbb{R}^{K \times 1} \) is the abundance fractions at the pixel, \( n_i \in \mathbb{R}^{L \times 1} \) represents noise in the pixel, \( L \) is the number of bands and \( K \) is the number of endmember classes. The endmember bundles \( E \) are defined as follows

\[ E = [E_1 | E_2 | \cdots | E_K] \]  

where \( E_k \in \mathbb{R}^{L \times N_k} \) represents a set of endmember spectra characterizing the \( k \)th class, \( N_k \) is the number of endmember spectra in the \( k \)th 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 \( B_i \) associated with the pixel is defined as the following block-diagonal matrix

\[ B_i = \begin{bmatrix} b_{1i} & 0_{N_1} & \cdots & 0_{N_1} \\ 0_{N_2} & b_{2i} & \cdots & 0_{N_2} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{N_K} & 0_{N_K} & \cdots & b_{Ki} \end{bmatrix} \]  

where \( b_{ki} \in \mathbb{R}^{N_k} \) is the bundling coefficients for the \( k \)th class at the pixel and \( 0_{N_k} \in \mathbb{R}^{N_k} \) is the \( N_k \)-dimensional vector whose components are zeros. Each bundling coefficient must be nonnegative and the bundling vector \( b_{ki} \) 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

\[ B_i \succeq 0 \quad \text{and} \quad \|B_i\|_0 = \sum_{k=1}^{K} \|b_{ki}\|_0 \leq s \]  

where \( \| \cdot \|_0 \) is the \( \ell_0 \)-norm that counts the number of nonzero elements and \( s \) is the maximum number of nonzero elements in \( B_i \), 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.,

\[ \forall k, \forall i, a_{ki} \geq 0, \quad \text{and} \quad \sum_{k=1}^{K} a_{ki} = 1, \quad \|a_i\|_0 \leq q \]  

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

\[ \min_{B_i, a_i} \frac{1}{2} \|EB_i a_i - y_i\|_2^2 \]  

s.t. \( \forall k, \forall i, a_{ki} \geq 0, \quad \sum_{k=1}^{K} a_{ki} = 1, \quad \|a_i\|_0 \leq q, \quad B_i \succeq 0, \quad \|B_i\|_0 \leq s \)  

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:

\[ \min_{B_i, a_i} \{ f(B_i, a_i) = f(B_i, a_i) + h(B_i) + g(a_i) \} \]  

with

\[ f(B_i, a_i) = \frac{1}{2} \|EB_i a_i - y_i\|_2^2 \]  

\[ h(B_i) = \iota_{R_+}(B_i) + \lambda_b \|B_i\|_0 \]  

\[ g(a_i) = \iota_{S}(a_i) + \lambda_a \|a_i\|_0 \]  

where \( \lambda_a \) and \( \lambda_b \) are parameters which control the balance between the data fit and the sparsity, \( \iota_C(x) \) is the indicator function on the set \( C \), i.e., \( \iota_C(x) = 0 \) when \( x \in C \) whereas \( \iota_C(x) = \infty \) when \( x \notin C \), and \( S \) 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 \( a_i \) and \( B_i \) by alternatively minimizing the objective function with respect to (w.r.t.) these parameters, i.e.,
by solving the following proximal problems
\[
B_i^{(t+1)} \in \min_{B_i} h(B_i) + \langle B_i - B_i^{(t)}, \nabla B_i f(B_i^{(t)}, a_i^{(t)}) \rangle + \frac{c_t}{2} \|B_i - B_i^{(t)}\|_2^2
\]
\[
a_i^{(t+1)} \in \min_{a_i} g(a_i) + \langle a_i - a_i^{(t)}, \nabla a_i f(B_i^{(t+1)}, a_i^{(t)}) \rangle + \frac{d_t}{2} \|a_i - a_i^{(t)}\|_2^2
\]  \hspace{1cm} (10)

**Optimization w.r.t. \(B_i\):** To optimize only w.r.t. the diagonal entries in \(B_i\), the objective function can be rewritten with the following decomposition
\[
f(b_i, a_i) = \frac{1}{2} \|U_i b_i - y_i\|_2^2
\]
\[
h(b_i) = \frac{1}{2} \|g_{b_i}(b_i)\|_2 + \lambda_b \|b_i\|_0
\]
where
\[
U_i = [E_1 \otimes a_{1i} \cdots |E_K \otimes a_{Ki}]
\]
\[
b_i = [b_{1i}^T, b_{2i}^T, \cdots, b_{Ki}^T]^T.
\]
This leads to the following updating rule
\[
\min_{b_i} h(b_i) + \frac{c_t}{2} \|b_i - (b_i^{(t)} - \frac{1}{c_t} \nabla b_i f(b_i^{(t)}, a_i^{(t)}))\|_2^2
\]
where \(\nabla b_i f(b_i^{(t)}, a_i^{(t)}) = U_i^T (U_i b_i - y_i)\). Using similar computations as in [15], this can be conducted as
\[
b_i^{(t+1)} \in \text{prox}_{\frac{c_t}{\lambda_b}} (b_i^{(t)} - \frac{1}{c_t} \nabla b_i f(b_i^{(t)}, a_i^{(t)}))
\]
where \(c_t = \gamma_m \|U_i^T U_i\|_F\) represents a step size for each iteration. The proximal operator associated with \(f\) can be computed using the approach [15]. Finally, the bundling matrix \(B_i\) can be reconstructed as \(B_i = \text{blkdiag}(b_i)\) where \(\text{blkdiag}(\cdot)\) generates the block diagonal matrix \(B_i\) from the vector \(b_i\).

**Optimization with respect to \(a_i\):** To optimize w.r.t. \(a_i\), the objective function can be rewritten using the decomposition
\[
f(B_i, a_i) = \frac{1}{2} \|S_i a_i - y_i\|_2^2
\]
\[
g(a_i) = \frac{1}{2} \|S(a_i)\|_1 + \lambda_a \|a_i\|_0
\]
where \(S_i = E B_i\). Thus, updating the abundance vector can be formulated as
\[
\min_{a_i} g(a_i) + \frac{d_t}{2} \|a_i - (a_i^{(t)} - \frac{1}{d_t} \nabla a_i f(B_i^{(t+1)}, a_i^{(t)}))\|_2^2
\]
where \(\nabla a_i f(B_i^{(t+1)}, a_i^{(t)}) = S_i^T (S_i a_i - y_i)\). Using the proximal operator, this can be written as
\[
a_i^{(t+1)} \in \text{prox}_{\frac{d_t}{\lambda_a}} (a_i^{(t)} - \frac{1}{d_t} \nabla a_i f(B_i^{(t+1)}, a_i^{(t)}))
\]  \hspace{1cm} (11)

where \(d_t = \gamma_a \|S_i^T S_i\|_F\) 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:** \(y, E\)
2. **Initialization:** \(a_i^{(0)}\) and \(B_i^{(0)}\).
3. **Main procedure:**
   4. while the stopping criterion is not satisfied do
     5. \(b_i^{(t+1)} \leftarrow \text{prox}_{\frac{c_t}{\lambda_b}} (b_i^{(t)} - \frac{1}{c_t} \nabla b_i f(b_i^{(t)}, a_i^{(t)}))\)
     6. \(B_i^{(t+1)} \leftarrow \text{blkdiag}(b_i^{(t+1)})\)
     7. \(a_i^{(t+1)} \leftarrow \text{prox}_{\frac{d_t}{\lambda_a}} (a_i^{(t)} - \frac{1}{d_t} \nabla a_i f(B_i^{(t+1)}, a_i^{(t)}))\)
   8. end while
9. **Output:** \(a_i^{(t+1)}, B_i^{(t+1)}\)

### 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]
\[
SRE \equiv \mathbb{E} \left[ \|a\|_2^2 / \mathbb{E} \|a - \hat{a}\|_2^2 \right]
\]  \hspace{1cm} (12)
where \(a\) is vectorized true abundances of all pixels, \(\hat{a}\) is vectorized estimated abundances of all pixels. To evaluate the
sparsity level (SL) induced by the methods, one monitors \([19, 20]\)

\[
\text{SL} \equiv \frac{1}{P} \sum_{i=1}^{P} \| \hat{a}_i \|_0. \tag{13}
\]

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

\[
\text{DIST} \equiv \frac{1}{P} \sum_{i=1}^{P} \max \left( |S_i|, |\hat{S}_i| \right) - |S_i \cap \hat{S}_i| \max \left( |S_i|, |\hat{S}_i| \right). \tag{14}
\]

where \(S_i\) and \(\hat{S}_i\) are the true and estimated supports of the \(i\)th pixel, i.e., the sets of index of nonzero values in \(a_i\) and \(\hat{a}_i\) of the pixel, \(|S|\) represents the total number of elements in the set \(S\) 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.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>FCLS</th>
<th>SUNSAL</th>
<th>Group lasso</th>
<th>Elitist lasso</th>
<th>MEMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRE</td>
<td>25.805</td>
<td>25.7331</td>
<td>26.2155</td>
<td>25.7959</td>
<td>26.1051</td>
</tr>
<tr>
<td>SL</td>
<td>3.45</td>
<td>3.62</td>
<td>4.22</td>
<td>4.95</td>
<td>2.35</td>
</tr>
<tr>
<td>DIST</td>
<td>0.3208</td>
<td>0.3592</td>
<td>0.4353</td>
<td>0.5033</td>
<td>0.0592</td>
</tr>
</tbody>
</table>

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 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 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
5. REFERENCES


