A PERTURBED LINEAR MIXING MODEL ACCOUNTING FOR SPECTRAL VARIABILITY

Pierre-Antoine Thouvenin, Nicolas Dobigeon and Jean-Yves Tourneret

University of Toulouse, IRIT/INP-ENSEEIHT, 2 rue Camichel, 31071 Toulouse cedex 7, France
firstname.lastname@enseeiht.fr

ABSTRACT

Hyperspectral unmixing aims at determining the reference spectral signatures composing a hyperspectral image, their abundance fractions and their number. In practice, the spectral variability of the identified signatures induces significant abundance estimation errors. To address this issue, this paper introduces a new linear mixing model explicitly accounting for this phenomenon. In this setting, the extracted endmembers are interpreted as possibly corrupted versions of the true endmembers. The parameters of this model can be estimated using an optimization algorithm based on the alternating direction method of multipliers. The performance of the proposed unmixing method is evaluated on synthetic and real data.

Index Terms— Hyperspectral imagery, linear unmixing, endmember variability, Alternating Direction Method of Multipliers (ADMM).

1. INTRODUCTION

Hyperspectral unmixing consists of identifying the spectral signatures from which the data are derived – referred to as *endmembers* – their abundances and their number according to a predefined mixing model. Assuming no microscopic interaction between the materials of the imaged scene, a linear mixing model (LMM) is known to be adapted to describe the data structure [1]. However, the spectral signatures contained in a reference pixel can vary spectrally from a pixel to another due to the varying acquisition conditions. This can result in significant estimation errors being propagated throughout the unmixing process. Various models either derived from a statistical or a deterministic point of view have been designed to address this issue [2].

Since the identified signatures can be considered as variable instances of reference endmembers, we introduce an extended version of the classical LMM to model the spectral variability. In [3], the variability is assumed to only result from scaling factors. Conversely, in this paper, inspired by a model designed in [4], each endmember is represented by a "pure" spectral signature corrupted by an additive perturbation accounting for spectral variability. The perturbation is allowed to vary from a pixel to another to represent spatial-spectral variability within the image. To the best of our knowledge, it is the first time endmember variability has been explicitly modeled as an additive perturbation. Finally, the results obtained with the ADMM in hyperspectral imagery [5] and in image deblurring [6, 7] have motivated the use of a similar framework for spectral unmixing using a perturbed LMM (PLMM).

The paper is organized as follows. The proposed PLMM is introduced in Section 2. Section 3 describes an ADMM-based algorithm to estimate the parameters of this model. Experimental results obtained on synthetic and real data are reported in Section 4 and 5 respectively. Section 6 finally concludes this work.

2. PROBLEM STATEMENT

2.1. Perturbed linear mixing model (PLMM)

Each pixel $\mathbf{y}_n$ of a hyperspectral image is represented by a linear combination of K endmembers – denoted by $\mathbf{m}_k$ – affected by a spatially varying perturbation vector $\mathbf{dm}_{n,k}$ accounting for endmember variability. The resulting PLMM can be written

$$\mathbf{y}_n = \sum_{k=1}^{K} a_{kn}\Big(\mathbf{m}_k + \mathbf{dm}_{n,k}\Big) + \mathbf{b}_n \text{ for } n = 1, \ldots, N \quad (1)$$

where L is the number of spectral bands, N is the number of pixels, $\mathbf{y}_n$ denotes the nth image pixel, $\mathbf{m}_k$ is the kth endmember, a_{kn} is the proportion of the kth endmember in the nth pixel, and $\mathbf{dm}_{n,k}$ denotes the perturbation of the kth endmember in the nth pixel. The term $\mathbf{b}_n$ represents the noise resulting from the data acquisition and the modeling errors. In matrix form, the PLMM (1) can be written as follows

$$\mathbf{Y} = \mathbf{MA} + \underbrace{\Big[\ \mathbf{dM}_1\mathbf{a}_1 \ \Big|\ \ldots \ \Big|\ \mathbf{dM}_N\mathbf{a}_N \ \Big]}_{\boldsymbol{\Delta}} + \mathbf{B} \quad (2)$$

where $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_N]$ is an $L \times N$ matrix containing the image pixels, $\mathbf{M}$ is an $L \times K$ matrix containing the endmembers $\mathbf{m}_k$, $\mathbf{A}$ is a $K \times N$ matrix composed of the abundance vectors $\mathbf{a}_n$, $\mathbf{dM}_n$ is an $L \times K$ matrix whose columns are

This work is supported by the Direction Générale de l'Armement (DGA), French Ministry of Defence, and by the Hypanema ANR Project n°ANR-12-BS03-003

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

the perturbation vectors $\mathbf{dm}_{n,k}$ associated with the nth pixel, and $\mathbf{B}$ is an $L \times N$ matrix accounting for the noise. The non-negativity and sum-to-one constraints usually considered to reflect physical considerations are

$$\begin{aligned} &\mathbf{A} \succeq \mathbf{O}_{K,N}, \quad \mathbf{A}^T\mathbf{1}_K = \mathbf{1}_N \\ &\mathbf{M} \succeq \mathbf{O}_{L,K}, \quad \mathbf{M} + \mathbf{dM}_n \succeq \mathbf{O}_{L,K}, \, \forall n = 1, \ldots, N. \end{aligned} \tag{3}$$

where $\succeq$ denotes term-wise inequality. When compared to models found in the literature to mitigate the variability impact [2], the model (1) presents the advantage to explicitly address the variability phenomenon. Besides, the variability is not assumed to only result from scaling factors as in [3].

2.2. Problem formulation

The PLMM (1) and constraints (3) can be combined to form a constrained optimization problem. We propose to define the data fitting term as the Frobenius distance between the observations and the reconstructed data. Since the problem is ill-posed, we introduce penalization functions Φ, Ψ and Υ to reflect the *a priori* knowledge on $\mathbf{M}$, $\mathbf{A}$ and $\mathbf{dM}$ respectively. As a result, the optimization problem is expressed as

$$(\mathbf{M}^*, \mathbf{dM}^*, \mathbf{A}^*) \in \underset{\mathbf{M},\mathbf{dM},\mathbf{A}}{\arg\min} \left\{ \mathcal{J}(\mathbf{M}, \mathbf{dM}, \mathbf{A}) \text{ s.t. } (3) \right\} \tag{4}$$

with

$$\begin{aligned} \mathcal{J}(\mathbf{M}, \mathbf{dM}, \mathbf{A}) = &\frac{1}{2} \left\| \mathbf{Y} - \mathbf{MA} - \mathbf{\Delta} \right\|_{\mathrm{F}}^2 + \alpha\Phi(\mathbf{A}) + \\ &\beta\Psi(\mathbf{M}) + \gamma\Upsilon(\mathbf{dM}) \end{aligned} \tag{5}$$

where the penalization parameters α, β, γ control the trade-off between the data fitting term and the penalties $\Phi(\mathbf{A})$, $\Psi(\mathbf{M})$ and $\Upsilon(\mathbf{dM})$. In addition, we assume that the penalization functions are separable, leading to

$$\begin{aligned} \Phi(\mathbf{A}) &= \sum_{n=1}^{N} \phi(\mathbf{a}_n), \quad \Psi(\mathbf{M}) = \sum_{\ell=1}^{L} \psi(\widetilde{\mathbf{m}}_\ell), \\ \Upsilon(\mathbf{dM}) &= \sum_{n=1}^{N} \upsilon(\mathbf{dM}_n) \end{aligned} \tag{6}$$

where ϕ, ψ and υ are non-negative differentiable convex functions, and $\widetilde{\mathbf{m}}_\ell$ denotes the ℓth row of $\mathbf{M}$. This assumption is used to decompose (4) into a collection of sub-problems described in Section 3. All these penalizations are described in the next paragraphs.

2.2.1. Abundance penalization

The abundance penalization Φ has been chosen to promote spatially smooth abundances as in [8]. More precisely, the abundance spatial smoothness penalization is written in matrix form as

$$\Phi(\mathbf{A}) = \frac{1}{2} \left\| \mathbf{AH} \right\|_{\mathrm{F}}^2 \tag{7}$$

where $\mathbf{H} \in \mathbb{R}^{N \times 4N}$ is a matrix computing the differences between the abundances of a given pixel and those of its 4 nearest neighbors [8]. The only terms depending on the nth abundance vector $\mathbf{a}_n$ are

$$\begin{aligned} \phi(\mathbf{a}_n) = &\underbrace{\frac{1}{2} \left(\sum_{k=0}^{3} h_{n,n+kN}^2 \right)}_{cA_n} \left\| \mathbf{a}_n \right\|_2^2 \\ &+ \underbrace{\left(\sum_{\substack{i=1 \\ i \neq n}}^{N} \sum_{k=0}^{3} h_{n,n+kN} h_{i,n+kN} \mathbf{a}_i^T \right)}_{\mathbf{c}_n^T} \mathbf{a}_n. \end{aligned} \tag{8}$$

2.2.2. Endmember penalization

As for Ψ, classical penalizations found in the literature consist of constraining the size of the simplex whose vertices are the endmember signatures. The mutual distance between each endmember introduced in [9] approximates the volume occupied by the $(K-1)$-simplex and is expressed as

$$\Psi(\mathbf{M}) = \frac{1}{2} \sum_{i \neq j} \left\| \mathbf{m}_i - \mathbf{m}_j \right\|_2^2 = \frac{1}{2} \sum_{k=1}^{K} \left\| \mathbf{MG}_k \right\|_{\mathrm{F}}^2 \tag{9}$$

where

$$\mathbf{G}_k = -\mathbf{I}_K + \mathbf{e}_k \mathbf{1}_K^T \tag{10}$$

for $k = 1, \ldots, K$ and $\mathbf{e}_k$ denotes the kth canonical basis vector of $\mathbb{R}^K$. Hence

$$\psi(\widetilde{\mathbf{m}}_\ell) = \frac{1}{2} \sum_{k=1}^{K} \left\| \widetilde{\mathbf{m}}_\ell \mathbf{G}_k \right\|_2^2 . \tag{11}$$

2.2.3. Variability penalization

The function Υ has been designed to limit the norm of the spectral variability in order to capture a reasonable endmember variability level. In this paper, we propose to consider

$$\upsilon\left(\mathbf{dM}_n\right) = \frac{1}{2} \left\| \mathbf{dM}_n \right\|_{\mathrm{F}}^2 . \tag{12}$$

3. AN ADMM-BASED ALGORITHM

Since the problem (4) is non-convex, the cost function $\mathcal{J}$ has been successively minimized with respect to each variable $\mathbf{A}$, $\mathbf{M}$ and $\mathbf{dM}$ until a stopping criterion is satisfied [6]. The assumptions made on Φ, Ψ, Υ in Section 2 allow the global optimization problem to be further decomposed into a collection of convex sub-problems exclusively involving differentiable functions. These sub-problems are finally solved using ADMM steps. The resulting algorithmic scheme is summarized in Algo 1. Some considerations about the convergence of the proposed algorithm are provided in an extended version of this paper [10].

Algorithm 1: PLMM-unmixing: global algorithm.

Data: $\mathbf{Y}, \mathbf{A}^{(0)}, \mathbf{M}^{(0)}, \mathbf{dM}^{(0)}$
begin
  $k \leftarrow 1$;
  while *stopping criterion not satisfied* **do**
(a) $\mathbf{A}^{(k)} \leftarrow \arg\min_{\mathbf{A}} \mathcal{J}\left(\mathbf{M}^{(k-1)}, \mathbf{dM}^{(k-1)}, \mathbf{A}\right)$;
(b) $\mathbf{M}^{(k)} \leftarrow \arg\min_{\mathbf{M}} \mathcal{J}\left(\mathbf{M}, \mathbf{dM}^{(k-1)}, \mathbf{A}^{(k)}\right)$;
(c) $\mathbf{dM}^{(k)} \leftarrow \arg\min_{\mathbf{M}} \mathcal{J}\left(\mathbf{M}^{(k)}, \mathbf{dM}, \mathbf{A}^{(k)}\right)$;
    $k \leftarrow k+1$;
  $\mathbf{A} \leftarrow \mathbf{A}^{(k)}$;
  $\mathbf{M} \leftarrow \mathbf{M}^{(k)}$;
  $\mathbf{dM} \leftarrow \mathbf{dM}^{(k)}$;
Result: $\mathbf{A}, \mathbf{M}, \mathbf{dM}$

3.1. Optimization with respect to A

With the assumptions made in paragraph 2.2, optimizing $\mathcal{J}$ with respect to $\mathbf{A}$ under the constraints (3) is equivalent to solving

$$\mathbf{a}_n^* = \arg\min_{\mathbf{a}_n} \left\{ \begin{array}{c} \frac{1}{2}\left\|\mathbf{y}_n - (\mathbf{M}+\mathbf{dM}_n)\mathbf{a}_n\right\|_2^2 + \alpha\phi(\mathbf{a}_n) \\ \text{s.t.} \quad \mathbf{a}_n \succeq \mathbf{0}_K, \quad \mathbf{a}_n^T\mathbf{1}_K = 1 \end{array} \right\}. \tag{13}$$

After introducing the splitting variables $\mathbf{w}_n^{(\mathbf{A})} \in \mathbb{R}^K$ for $n = 1, \ldots, N$ such that

$$\underbrace{\begin{pmatrix} \mathbf{I}_K \\ \mathbf{1}_K^T \end{pmatrix}}_{\mathbf{Q}} \mathbf{a}_n + \underbrace{\begin{pmatrix} -\mathbf{I}_K \\ \mathbf{0}_K^T \end{pmatrix}}_{\mathbf{R}} \mathbf{w}_n = \underbrace{\begin{pmatrix} \mathbf{0}_K \\ 1 \end{pmatrix}}_{\mathbf{s}} \tag{14}$$

the resulting scaled augmented Lagrangian is expressed as

$$\begin{aligned} \mathcal{L}_{\mu_n^{(\mathbf{A})}}\left(\mathbf{a}_n, \mathbf{w}_n^{(\mathbf{A})}, \boldsymbol{\lambda}_n^{(\mathbf{A})}\right) &= \frac{1}{2}\left\|\mathbf{y}_n - (\mathbf{M}+\mathbf{dM}_n)\mathbf{a}_n\right\|_2^2 \\ &+ \frac{\mu_n^{(\mathbf{A})}}{2}\left\|\mathbf{Q}\mathbf{a}_n + \mathbf{R}\mathbf{w}_n^{(\mathbf{A})} - \mathbf{s} + \boldsymbol{\lambda}_n^{(\mathbf{A})}\right\|_2^2 \\ &+ \alpha\phi(\mathbf{a}_n) + \mathcal{I}_{K,1}^+\left(\mathbf{w}_n^{(\mathbf{A})}\right) \end{aligned} \tag{15}$$

where $\mu_n^{(\mathbf{A})} > 0$ and $\mathcal{I}_{K,1}^+$ is the indicator function on $(\mathbb{R}_+)^K$. Thus, for $n = 1, \ldots, N$

$$\mathbf{a}_n^* = \left[(\mathbf{M}+\mathbf{dM}_n)^T(\mathbf{M}+\mathbf{dM}_n) + \mu_n^{(\mathbf{A})}\mathbf{Q}^T\mathbf{Q} + \alpha c A_n \mathbf{I}_K\right]^{-1} \left[(\mathbf{M}+\mathbf{dM}_n)^T\mathbf{y}_n - \alpha\mathbf{c}_n + \mu_n^{(\mathbf{A})}\mathbf{Q}^T\left(\mathbf{s} - \mathbf{R}\mathbf{w}_n^{(\mathbf{A})} - \boldsymbol{\lambda}_n^{(\mathbf{A})}\right)\right] \tag{16}$$

and

$$\mathbf{w}_n^{(\mathbf{A})^*} = \max\left(\mathbf{a}_n + \boldsymbol{\lambda}_{n,1:K}^{(\mathbf{A})}, \mathbf{0}_K\right) \tag{17}$$

where $\boldsymbol{\lambda}_{n,1:K}^{(\mathbf{A})}$ is the vector composed of the K first elements of $\boldsymbol{\lambda}_n^{(\mathbf{A})}$ and the max should be understood as a term-wise operator. In the absence of any penalization, the solution is obtained by making $\alpha = 0$ in the previous equations.

3.2. Optimization with respect to M

Similarly, optimizing $\mathcal{J}$ with respect to $\mathbf{M}$ under the constraints (3) is equivalent to solving

$$\widetilde{\mathbf{m}}_\ell^* = \arg\min_{\widetilde{\mathbf{m}}_\ell} \left\{ \begin{array}{c} \frac{1}{2}\left\|\widetilde{\mathbf{y}}_\ell - \widetilde{\mathbf{m}}_\ell\mathbf{A} - \widetilde{\boldsymbol{\delta}}_\ell\right\|_2^2 + \beta\psi(\widetilde{\mathbf{m}}_\ell) \\ \text{s.t.} \quad \text{for} \quad n = 1, \ldots, N \\ \widetilde{\mathbf{m}}_\ell \succeq \mathbf{0}_K^T, \quad \widetilde{\mathbf{m}}_\ell + \widetilde{\mathbf{dm}}_{n,\ell} \succeq \mathbf{0}_K^T \end{array} \right\} \tag{18}$$

where $\widetilde{\mathbf{y}}_\ell$, $\widetilde{\boldsymbol{\delta}}_\ell$ and $\widetilde{\mathbf{dm}}_{n,\ell}$ denote the ℓth row of $\mathbf{Y}$, $\boldsymbol{\Delta}$ and $\mathbf{dM}_n$ respectively. Introducing the splitting variables $\mathbf{W}_\ell^{(\mathbf{M})}$ in $\mathbb{R}^{(N+1)\times K}$ for $\ell = 1, \ldots, L$ such that

$$\underbrace{\begin{pmatrix} 1 \\ \mathbf{1}_N \end{pmatrix}}_{\mathbf{e}} \widetilde{\mathbf{m}}_\ell - \mathbf{W}_\ell^{(\mathbf{M})} = -\underbrace{\left[\mathbf{0}_K, \widetilde{\mathbf{dm}}_{1,\ell}^T, \ldots, \widetilde{\mathbf{dm}}_{N,\ell}^T\right]^T}_{\mathbf{F}_\ell} \tag{19}$$

the associated scaled augmented Lagrangian can be written

$$\begin{aligned} \mathcal{L}_{\mu_\ell^{(\mathbf{M})}}\left(\widetilde{\mathbf{m}}_\ell, \mathbf{W}_\ell^{(\mathbf{M})}, \boldsymbol{\Lambda}_\ell^{(\mathbf{M})}\right) &= \frac{1}{2}\left\|\widetilde{\mathbf{y}}_l - \widetilde{\mathbf{m}}_\ell\mathbf{A} - \widetilde{\boldsymbol{\delta}}_\ell\right\|_2^2 \\ &+ \frac{\mu_\ell^{(\mathbf{M})}}{2}\left\|\mathbf{e}\widetilde{\mathbf{m}}_\ell - \mathbf{W}_\ell^{(\mathbf{M})} + \mathbf{F}_\ell + \boldsymbol{\Lambda}_\ell^{(\mathbf{M})}\right\|_{\mathrm{F}}^2 \\ &+ \beta\psi(\widetilde{\mathbf{m}}_\ell) + \mathcal{I}_{N+1,K}^+\left(\mathbf{W}_\ell^{(\mathbf{M})}\right) \end{aligned} \tag{20}$$

with $\mu_\ell^{(\mathbf{M})} > 0$ and $\mathcal{I}_{N+1,K}^+$ the indicator function on $(\mathbb{R}_+)^{N+1\times K}$. Thus

$$\widetilde{\mathbf{m}}_\ell^* = \left[\left(\widetilde{\mathbf{y}}_\ell - \widetilde{\boldsymbol{\delta}}_\ell\right)\mathbf{A}^T + \mu_\ell^{(\mathbf{M})}\mathbf{e}^T\left(\mathbf{W}_\ell^{(\mathbf{M})} - \mathbf{F}_\ell - \boldsymbol{\Lambda}_\ell^{(\mathbf{M})}\right)\right] \left[\mathbf{A}\mathbf{A}^T + \beta\sum_{k=1}^K \mathbf{G}_k\mathbf{G}_k^T + \mu_\ell^{(\mathbf{M})}\left(\mathbf{e}^T\mathbf{e}\right)\mathbf{I}_K\right]^{-1} \tag{21}$$

and

$$\mathbf{W}_\ell^{(\mathbf{M})*} = \max\left(\mathbf{e}\widetilde{\mathbf{m}}_\ell + \mathbf{F}_\ell + \boldsymbol{\Lambda}_\ell^{(\mathbf{M})}, \mathbf{O}_{(N+1),K}\right). \tag{22}$$

In the absence of any endmember penalization, the solution is obtained by making $\beta = 0$ in the previous equation.

3.3. Optimization with respect to dM

Finally, optimizing $\mathcal{J}$ with respect to $\mathbf{dM}$ under the constraint (3) is equivalent to solving

$$\mathbf{dM}_n^* = \arg\min_{\mathbf{dM}_n} \left\{ \begin{array}{c} \frac{1}{2}\left\|\mathbf{y}_n - (\mathbf{M}+\mathbf{dM}_n)\mathbf{a}_n\right\|_2^2 \\ +\gamma\upsilon(\mathbf{dM}_n) \\ \text{s.t.} \quad \mathbf{M}+\mathbf{dM}_n \succeq \mathbf{O}_{L,K} \end{array} \right\} \tag{23}$$

for $n = 1, \ldots, N$. Introducing the splitting variables $\mathbf{W}_n^{(\mathbf{dM})} = \mathbf{M} + \mathbf{dM}_n$, the resulting scaled augmented Lagrangian is given by

$$\begin{aligned}\mathcal{L}_{\mu_n^{(\mathbf{dM})}}\left(\mathbf{dM}_n, \mathbf{W}_n^{(\mathbf{dM})}, \mathbf{\Lambda}_n^{(\mathbf{dM})}\right) &= \mathcal{I}_{L,K}^{+}\left(\mathbf{W}_n^{(\mathbf{dM})}\right) \\ &+ \frac{1}{2}\left\|\mathbf{y}_n - (\mathbf{M}+\mathbf{dM}_n)\mathbf{a_n}\right\|_2^2 + \gamma\upsilon(\mathbf{dM}_n) \\ &+ \frac{\mu_n^{(\mathbf{dM})}}{2}\left\|\mathbf{dM}_n + \mathbf{M} - \mathbf{W}_n^{(\mathbf{dM})} + \mathbf{\Lambda}_n^{(\mathbf{dM})}\right\|_{\mathrm{F}}^2\end{aligned} \tag{24}$$

with $\mu_n^{(\mathbf{dM})} > 0$ and $\mathcal{I}_{L,K}^{+}$ the indicator function on $(\mathbb{R}_+)^{L\times K}$. Hence

$$\begin{aligned}\mathbf{dM}_n^{*} = \Bigg[&(\mathbf{y_n} - \mathbf{M}\mathbf{a}_n)\mathbf{a}_n^T + \mu_n^{(\mathbf{dM})}\Big(\mathbf{W}_n^{(\mathbf{dM})} \\ &- \mathbf{M} - \mathbf{\Lambda}_n^{(\mathbf{dM})}\Big)\Bigg]\Big[\mathbf{a}_n\mathbf{a}_n^T + (\mu_n^{(\mathbf{dM})} + \gamma)\mathbf{I}_K\Big]^{-1}\end{aligned} \tag{25}$$

and

$$\mathbf{W}_n^{(\mathbf{dM})*} = \max\left(\mathbf{dM}_n + \mathbf{M} + \mathbf{\Lambda}_n^{(\mathbf{dM})}, \mathbf{O}_{L,K}\right). \tag{26}$$

4. EXPERIMENTS WITH SYNTHETIC DATA

The method is first evaluated on a 128×64-pixel image resulting from linear mixtures of 3 endmembers with $L = 160$ spectral bands, without any pure pixel to evaluate the method in a challenging situation. Any mixture has been corrupted by an additive white Gaussian noise to ensure a signal-to-noise ratio of 30dB. The corrupted endmembers involved in the mixture have been generated using the product of reference endmembers with randomly generated piecewise-affine functions. Different affine functions have been considered for each endmember in each pixel, which provides realistic endmembers with controlled variability [1].

4.1. State-of-the-art methods

The proposed method is compared to the VCA [11] / FCLS [12] algorithms and to the automated endmember bundles (AEB, [13]). For the proposed method, the endmembers and abundances have been initialized with VCA/FCLS and the variability matrices have been initialized with all their entries equal to eps[2]. The algorithm is stopped when the relative difference between two successive values of the objective function is less than 10^{-3}. The regularization parameters associated with the augmented Lagrangians have been initialized with the following values: $\mu_n^{(\mathbf{A})(0)} = \mu_n^{(\mathbf{dM})(0)} = 10^{-4}$, $\mu_\ell^{(\mathbf{M})(0)} = 10^{-8}$, and adjusted using the rule described in [14, p. 20] with $\tau^{\text{incr}} = \tau^{\text{decr}} = 1.1$, $\mu = 10$, $\varepsilon^{\text{abs}} = 10^{-1}$ and $\varepsilon^{\text{rel}} = 10^{-4}$.

[1] More simulations results are available in [10].

[2] MATLAB constant $eps = 2.22 \times 10^{-16}$.

Table 1. Simulation results for synthetic data (GMSE($\mathbf{A}$)$\times 10^{-2}$, GMSE($\mathbf{dM}$)$\times 10^{-4}$, RE $\times 10^{-4}$).

	VCA/FCLS	AEB	Proposed method
aSAM($\mathbf{M}$)	5.0639	5.1104	**4.1543**
GMSE($\mathbf{A}$)	2.07	2.11	**1.44**
GMSE($\mathbf{dM}$)	/	/	**4.36**
RE	2.66	2.66	**0.38**
time (s)	**1**	33	1990

The performance of the algorithm has been assessed in terms of endmember estimation using the average spectral angle mapper (aSAM). In terms of abundance and perturbation estimations, global mean square errors (GMSEs) have been computed. Finally, the reconstruction error (RE) detailed in the following lines has been considered as a measure of fit.

$$\begin{aligned}\text{aSAM}(\mathbf{M}) &= \frac{1}{K}\sum_{k=1}^{K}\frac{\langle \mathbf{m}_k | \widehat{\mathbf{m}}_k\rangle}{\|\mathbf{m}_k\|_2\|\widehat{\mathbf{m}}_k\|_2} \\ \text{GMSE}(\mathbf{dM}) &= \frac{1}{NLK}\sum_{n=1}^{N}\|\mathbf{dM}_n - \widehat{\mathbf{dM}}_n\|_{\mathrm{F}}^2 \\ \text{GMSE}(\mathbf{A}) &= \frac{1}{KN}\|\mathbf{A} - \widehat{\mathbf{A}}\|_{\mathrm{F}}^2 \\ \text{RE} &= \frac{1}{LN}\left\|\mathbf{Y} - \widehat{\mathbf{Y}}\right\|_{\mathrm{F}}^2\end{aligned}$$

where $\widehat{\mathbf{Y}}$ is the matrix formed by the pixels reconstructed with the estimated parameters $\widehat{\mathbf{A}}$, $\widehat{\mathbf{M}}$ and $\widehat{\mathbf{dM}}_n$.

4.2. Results

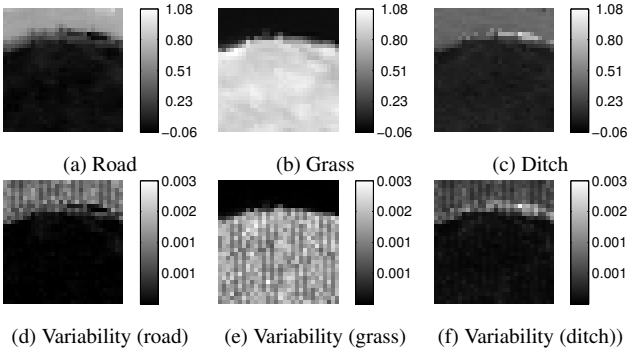
The performance measures returned by the unmixing methods are provided in Table 1. The proposed method provides competitive results when compared to other methods and exhibits lower REs, at the price of a higher computational cost.

Table 2. Experiments conducted on real data (Madonna) (RE $\times 10^{-6}$).

	VCA/FCLS	AEB	ADMM
RE	8.64	5.25	**0.43**
time (s)	**0.41**	1.77	1.88

5. EXPERIMENTS WITH REAL DATA

Keeping the same parameters as in Section 4, the proposed algorithm has been applied to a data-set (31×30) composed of 160 spectral bands extracted from an image acquired in 2010 by the Hyspex hyperspectral scanner over Villelongue, France. The image is composed of forested and urban areas.



(a) Road (b) Grass (c) Ditch

(d) Variability (road) (e) Variability (grass) (f) Variability (ditch))

Fig. 1. Abundance and variability distribution (real data).

Since there is no ground truth for this data set, we present the unmixing performance in terms of reconstruction error (see Table 2). The results depicted in Figs. 1 and 2 are consitent with those of VCA/FCLS. Moreover, regular vertical patterns almost surely due to a sensor defect or miscalibration during the data post-processing appear on the energy map of the variability terms. This observation is consistent with the remark made in [15].

6. CONCLUSION AND FUTURE WORK

This paper introduced a new linear mixing model including an additive spatially varying perturbation matrix to capture endmember variability. Hyperspectral unmixing was performed by alternating minimization of an appropriately regularized cost function, each minimization being performed by ADMM. Simulations conducted on synthetic and real data enabled the interest of the proposed solution to be appreciated. The choice of the penalization parameters α, β and γ was performed by cross validation. The development of automatic strategies to estimate these parameters remains an open problem.

REFERENCES

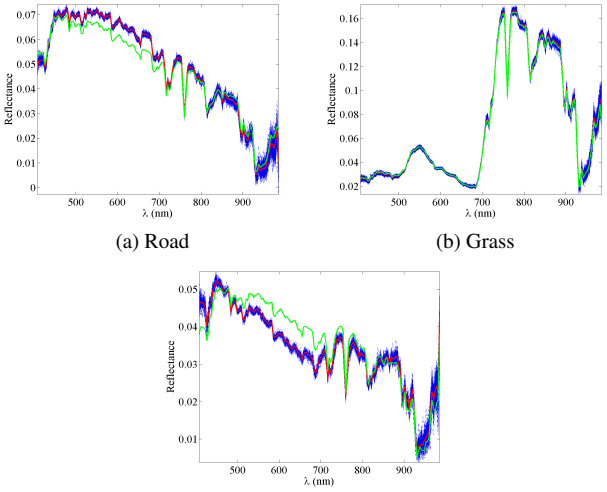
[1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observations and Remote Sens.*, vol. 5, no. 2, pp. 354–379, April 2012.

[2] A. Zare and K. C. . Ho, "Endmember variability in hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan. 2014.

[3] M. A. Veganzones, L. Drumetz, G. Tochon, M. D. Mura, A. Plaza, J. M. Bioucas-Dias, and J. Chanussot, "A new extended linear mixing model to adress spectral variability," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Process.: Evolution in Remote Sens. (WHISPERS)*, Lausanne, Switzerland, June 2014.

[4] E. C. Johnson and D. L. Jones, "Joint recovery of sparse signals and parameter perturbations with parameterized measurement models," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013, pp. 5900–5904.

[5] R. Ammanouil, A. Ferrari, C. Richard, and D. Mary, "Blind and fully constrained unmixing of hyperspectral images," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5510–5518, Dec. 2014.



(a) Road (b) Grass

(c) Ditch

Fig. 2. Endmember estimation on real data (ADMM in red lines, VCA in green lines, variability in blue dotted lines).

[6] M. S. C. Almeida and M. A. T. Figueiredo, "Blind image deblurring with unknown boundaries using the alternating direction method of multipliers," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Melbourne, Australia, 2013, pp. 586–590.

[7] A. Matakos, S. Ramani, and J. A. Fessler, "Accelerated edge-preserving image restoration with boundary artifacts," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 2019–2029, May 2013.

[8] J. Chen, C. Richard, and P. Honeine, "Nonlinear estimation of material abundance in hyperspectral images with ℓ_1-norm spatial regularization," *IEEE Trans. Geosci. and Remote Sensing*, vol. 52, no. 5, pp. 2654–2665, May 2014.

[9] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington, "ICE: A statistical approach to identifying endmembers in hyperspectral images," *IEEE Trans. Geosci. and Remote Sensing*, vol. 42, no. 10, pp. 2085–2095, Oct. 2004.

[10] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model," Feb. 2015, submitted. [Online]. Available: http://arxiv.org/abs/1502.01260/

[11] J. M. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. and Remote Sensing*, vol. 43, no. 4, pp. 898–910, April 2005.

[12] 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. and Remote Sensing*, vol. 29, no. 3, pp. 529–545, March 2001.

[13] B. Somers, M. Zortea, A. Plaza, and G. Asner, "Automated extraction of image-based endmember bundles for improved spectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observations and Remote Sens.*, vol. 5, no. 2, pp. 396–408, April 2012.

[14] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.

[15] C. Févotte and N. Dobigeon, "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," March 2014, submitted. [Online]. Available: http://arxiv.org/abs/1401.5649/