UNMIXING MULTITEMPORAL HYPERSPECTRAL IMAGES WITH VARIABILITY: AN ONLINE ALGORITHM

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 consists in determining the reference spectral signatures composing a hyperspectral image and their relative abundance fractions in each pixel. In practice, the identified signatures may be affected by a significant spectral variability resulting for instance from the temporal evolution of the imaged scene. This phenomenon can be accounted for by using a perturbed linear mixing model. This paper studies an online estimation algorithm for the parameters of this extended linear mixing model. This algorithm is of interest for the practical applications where the size of the hyperspectral images precludes the use of batch procedures. The performance of the proposed method is evaluated on synthetic data.

Index Terms— Multi-temporal hyperspectral image, linear unmixing, endmember variability, online estimation, two-stage stochastic program.

1. INTRODUCTION

Hyperspectral unmixing aims at identifying the reference spectral signatures composing a hyperspectral image –referred to as *endmembers* – and their relative abundance fractions in each pixel. When the microscopic interactions between the materials and the scene relief are negligible, a linear mixing model (LMM) is traditionally used to describe hyperspectral data [1]. However, varying acquisition conditions such as illumination or natural evolution of the scene can significantly modify the acquired spectral signatures, thus affecting the extracted endmembers from an image to another. In this context, unmixing several images acquired over the same area at different time instants gives the opportunity to track the endmembers contained in this area and to analyze the potential temporal variability.

So far, endmember variability has essentially been considered for a single hyperspectral image [2–6]. However, some applications require to unmix a sequence of multi-temporal hyperspectral images, which can preclude the use of batch estimation procedures due to limited memory resources, contrasting with the batch approach proposed in [7]. Moreover, a sequential analysis has the advantage of processing any image individually, when it is acquired, avoiding to handle a huge amount of data and avoiding any storage. This sequential processing will be clearly interesting for the analysis of the numerous multi-temporal images that will be regularly acquired in the coming years (by the European Space agency[1] for instance). Since the identified endmembers can be considered as time-varying instances of the desired reference endmembers, we exploit the perturbed linear mixing model (PLMM) proposed in [8] to account for

[1]`http://www.esa.int/Our_Activities/Observing_the_Earth/Proba-1/Going_hyperspectral`

spectral variability. Inspired by the works presented in [9, 10], we formulate the unmixing problem associated with the PLMM as a two-stage stochastic program and finally perform an online estimation of its parameters.

The paper is organized as follows. The PLMM accounting for temporal variability is introduced in Section 2. Section 3 investigates a sequential algorithm to solve the optimization problem associated with the PLMM in an online manner. Experimental results obtained on synthetic data are reported in Section 4. Conclusions are finally reported in Section 5.

2. PROBLEM STATEMENT

2.1. Perturbed linear mixing model (PLMM)

Consider hyperspectral images sharing the same K endmembers – where K is known – whose pixels are similarly affected by spectral variability. The nth pixel of the tth hyperspectral image $\mathbf{y}_{nt}$ is modeled by a linear combination of the K endmembers – denoted as $\mathbf{m}_1, \ldots \mathbf{m}_K$ – affected by a time-varying perturbation vector $\mathbf{dm}_{kt}$ accounting for temporal endmember variability. The resulting PLMM [8] can be written as

$$\mathbf{y}_{nt} = \sum_{k=1}^{K} a_{knt}\Big(\mathbf{m}_k + \mathbf{dm}_{kt}\Big) + \mathbf{b}_{nt} \tag{1}$$

for $n = 1, \ldots, N$ and $t = 1, \ldots, T$, where $\mathbf{m}_k$ is the kth endmember, a_{knt} is the proportion of the kth endmember in the nth pixel of the image $\sharp t$, $\mathbf{dm}_{kt}$ denotes the perturbation of the kth endmember at time t and $\mathbf{b}_{nt}$ denotes the noise resulting from the data acquisition and the modeling errors. In matrix form, the model (1) can be expressed as

$$\mathbf{Y}_t = (\mathbf{M} + \mathbf{dM}_t)\mathbf{A}_t + \mathbf{B}_t \tag{2}$$

where $\mathbf{Y}_t = [\mathbf{y}_{1t}, \ldots, \mathbf{y}_{Nt}]$ is an $L \times N$ matrix containing the tth image pixels, $\mathbf{M}$ denotes an $L \times K$ matrix containing the endmembers, $\mathbf{A}_t$ is a $K \times N$ matrix composed of the abundance vectors $\mathbf{a}_{nt} = [a_{1nt}, \ldots, a_{Knt}]^T$, $\mathbf{dM}_t$ is an $L \times K$ matrix whose columns are the perturbation vectors at time t, and $\mathbf{B}_t$ is an $L \times N$ matrix containing the noise terms. The non-negativity and sum-to-one constraints usually considered to reflect physical considerations are

$$\mathbf{A}_t \succeq \mathbf{0}_{K,N}, \quad \mathbf{A}_t^T \mathbf{1}_K = \mathbf{1}_N, \ \forall t = 1, \ldots, T \tag{3}$$

$$\mathbf{M} \succeq \mathbf{0}_{L,K} \tag{4}$$

where $\succeq$ stands for term-wise inequality. We further assume the following constraints

$$\|\mathbf{dM}_t\|_F^2 \leq \sigma^2, \quad \text{for } t = 1, \ldots, T \tag{5}$$

$$\left\|\sum_{t=1}^{T} \mathbf{dM}_t\right\|_F^2 \leq \kappa^2 \tag{6}$$

where σ and κ are predefined positive constants. The constraint (5) sets a limit on the variability energy, whereas (6) constrains the energy of the average temporal variability to be small. In other words, (6) ensures that the endmember signatures (i.e., each column of $\mathbf{M}$) will be close to the mean of the corresponding perturbed endmembers. We denote as $\mathcal{A}_K$, $\mathcal{M}$ and $\mathcal{D}_t$ the spaces associated with the constraints relative to the abundances (see Eq. (3)), the endmembers (see Eq. (4)) and the variability (see Eqs. (5) and (6)). The multi-temporal unmixing procedure proposed in this paper is aimed at estimating $\mathbf{A}_t$, $\mathbf{M}$ and $\mathbf{dM}_t$ without loading all the images simultaneously into memory.

2.2. Problem formulation

The unmixing problem introduced in the previous section can be formulated as a two-stage stochastic problem associated with the following empirical risk minimization

$$\min_{\mathbf{M} \in \mathcal{M}} \frac{1}{T} \sum_{t=1}^{T} h(\mathbf{Y}_t, \mathbf{M}) + \beta \Psi(\mathbf{M}) \tag{7}$$

$$h(\mathbf{Y}_t, \mathbf{M}) = \min_{(\mathbf{A}, \mathbf{dM}) \in \mathcal{A}_K \times \mathcal{D}_t} f(\mathbf{Y}_t, \mathbf{M}, \mathbf{A}, \mathbf{dM}) \tag{8}$$

where Ψ is an endmember regularization term. f represents a regularized instantaneous discrepancy measure, and h denotes the cost of the tth optimal decision taken to update the endmembers $\mathbf{M}$ given the data available at time t. Strictly speaking, h should be a set-valued function since f is not necessarily convex with respect to $(\mathbf{A}, \mathbf{dM})$. In this work, we assume that, for fixed $\mathbf{Y}_t$ and $\mathbf{M}$, $f(\mathbf{Y}_t, \mathbf{M}, \cdot, \cdot)$ admits locally unique stationary points, and we consequently define $h(\mathbf{Y}_t, \mathbf{M})$ as the value of $f(\mathbf{Y}_t, \mathbf{M}, \cdot, \cdot)$ at the stationary point $(\mathbf{A}_t, \mathbf{dM}_t)$ that any non-convex optimization method yields from a given starting point $(\mathbf{A}_t^{(0)}, \mathbf{dM}_t^{(0)})$. Assuming the noise matrix in (2) is white and Gaussian, f can be defined as

$$f(\mathbf{Y}, \mathbf{M}, \mathbf{A}, \mathbf{dM}) = \frac{1}{2} \|\mathbf{Y} - (\mathbf{M} + \mathbf{dM})\mathbf{A}\|_F^2 \tag{9}$$
$$+ \alpha \Phi(\mathbf{A}) + \gamma \Upsilon(\mathbf{dM})$$

where Φ and Υ are appropriate penalty terms, and the penalization parameters α, β, γ control the trade-off between the data fitting term and the penalties. The expression of the penalties considered in this paper are detailed below.

The abundance and variability penalties Φ_t and Υ_t ($t = 1, \ldots, T$) are chosen to promote smooth time variations between two consecutive images [12], i.e.,

$$\Phi_t(\mathbf{A}) = \frac{1}{2} \|\mathbf{A} - \mathbf{A}_{t-1}\|_F^2 \tag{10}$$

$$\Upsilon_t(\mathbf{dM}) = \frac{1}{2} \|\mathbf{dM} - \mathbf{dM}_{t-1}\|_F^2 \tag{11}$$

since the images are assumed to be acquired at time instants close from one another (i.e., only exhibit a smooth temporal variation). Classical endmember penalizations found in the literature aim at constraining the size of the $(K-1)$-simplex whose vertices are the endmembers. In this paper, we consider the mutual distance between the endmembers introduced in [13, 14], expressed as

$$\Psi(\mathbf{M}) = \frac{1}{2} \sum_{i=1}^{K} \left(\sum_{\substack{j=1 \\ j \neq i}}^{K} \|\mathbf{m}_i - \mathbf{m}_j\|_2^2 \right). \tag{12}$$

Algorithm 1: Online unmixing algorithm.

Data: $\mathbf{M}_0, \mathbf{A}_0, \mathbf{dM}_0, \alpha > 0, \beta > 0, \gamma > 0, \xi \in]0, 1]$
begin
    $\mathbf{C}_0 \leftarrow \mathbf{0}_{K,K}$;
    $\mathbf{D}_0 \leftarrow \mathbf{0}_{L,K}$;
    $\mathbf{E}_0 \leftarrow \mathbf{0}_{L,K}$;
    for $t = 1$ **to** T **do**
    a   Random selection of an image $\mathbf{Y}_t$;

        // Abundance and variability estimation by PALM [11], cf. §3.1
    b $(\mathbf{A}_t, \mathbf{dM}_t) \in \underset{(\mathbf{A}, \mathbf{dM}) \in \mathcal{A}_K \times \mathcal{D}_t}{\arg\min} f(\mathbf{Y}_t, \mathbf{M}_t, \mathbf{A}, \mathbf{dM})$;
        $\mathbf{C}_t \leftarrow \xi \mathbf{C}_{t-1} + \mathbf{A}_t \mathbf{A}_t^T$;
        $\mathbf{D}_t \leftarrow \xi \mathbf{D}_{t-1} + (\mathbf{dM}_t \mathbf{A}_t - \mathbf{Y}_t) \mathbf{A}_t^T$;
        $\mathbf{E}_t \leftarrow \xi \mathbf{E}_{t-1} + \mathbf{dM}_t$;
        // Endmember update [10, Alg. 2], cf. §3.2
    c $\mathbf{M}_t \leftarrow \underset{\mathbf{M} \in \mathcal{M}}{\arg\min} \hat{g}_t(\mathbf{M})$;

Result: $\mathbf{M}_T, (\mathbf{A}_t)_{t=1,\ldots,T}, (\mathbf{dM}_t)_{t=1,\ldots,T}$

---

3. AN ONLINE UNMIXING ALGORITHM

Whenever an image $\mathbf{Y}_t$ is received, we propose to solve the local problem (8), i.e., to estimate the abundances and the variability, by using a proximal alternating linearized minimization (PALM) [11] which is guaranteed to converge to a critical point of $f(\mathbf{Y}_t, \mathbf{M}_t, \cdot, \cdot)$. The endmembers are then updated by proximal gradient descent steps to solve (7) in the limit, similarly to [10]. The main algorithm is summarized in Algorithm 1 and detailed in the following paragraphs.

3.1. Abundance and variability estimation

Applying PALM [11] to the abundance and the variability estimation leads to the following update rule

$$\mathbf{A}_t^{(k+1)} = \mathcal{P}_{\mathcal{A}_K} \left(\mathbf{A}_t^{(k)} - \frac{1}{\lambda_t^{(k)}} \nabla_\mathbf{A} f(\mathbf{Y}_t, \mathbf{M}_t, \mathbf{A}_t^{(k)}, \mathbf{dM}_t^{(k)}) \right)$$
$$\lambda_t^{(k)} = \gamma_1 L_{1t}^{(k)}, \ \gamma_1 > 1$$

where k denotes the current PALM iteration, $L_{1t}^{(k)}$ is the Lipschitz constant of $\nabla_\mathbf{A} f(\mathbf{Y}_t, \mathbf{M}_t, \cdot, \mathbf{dM}_t^{(k)})$ and $\mathcal{P}_{\mathcal{A}_K}$ is the projector on $\mathcal{A}_K$ [15, 16]. The variability term is similarly updated by
$$\mathbf{dM}_t^{(k+1)} =$$

$$\mathcal{P}_{\mathcal{D}_t} \left(\mathbf{dM}_t^{(k)} - \frac{1}{\nu_t^{(k)}} \nabla_\mathbf{dM} f(\mathbf{Y}_t, \mathbf{M}_t, \mathbf{A}_t^{(k+1)}, \mathbf{dM}_t^{(k)}) \right)$$

$$\nu_t^{(k)} = \gamma_2 L_{2t}^{(k)}, \ \gamma_2 > 1$$

where $\mathcal{P}_{\mathcal{D}_t}$ is the projector on $\mathcal{D}_t$ and $L_{2t}^{(k)}$ is the Lipschitz constant of $\nabla_\mathbf{dM} f(\mathbf{Y}_t, \mathbf{M}_t, \mathbf{A}_t^{(k+1)}, \cdot)$. This projection is computed by the Dykstra algorithm [17, 18]. Indeed, we can introduce an auxiliary variable $\mathbf{E}_t$ and observe that

$$\mathcal{D}_t = \mathcal{B}_F(\mathbf{0}, \sigma) \cap \{\mathbf{dM}_t : \|\mathbf{dM}_t + \mathbf{E}_{t-1}\|_F \leq \kappa\}$$

$$\mathbf{E}_{t-1} = \sum_{i=1}^{t-1} \mathbf{dM}_i.$$

where $\mathcal{B}_F(\mathbf{0}, \sigma)$ denotes the Frobenius ball of radius σ centered in $\mathbf{0}$.

3.2. Endmember estimation

A direct application of the method detailed in [10] leads to the following endmember update rule

$$\mathbf{M}_t^{(k+1)} = \mathcal{P}_+ \left(\mathbf{M}_t^{(k)} - \frac{1}{\mu_t^{(k)}} \nabla_{\mathbf{M}} \hat{g}_t(\mathbf{M}_t^{(k)}) \right)$$

$$\mu_t^{(k)} = \gamma_3 L_{3t}, \ \gamma_3 > 1$$

with

$$\hat{g}_t(\mathbf{M}) = \frac{1}{2t} \sum_{i=1}^{t} \|\mathbf{Y}_i - (\mathbf{M} + \mathbf{dM}_i)\mathbf{A}_i\|_{\mathrm{F}}^2 + \beta\Psi(\mathbf{M})$$

$$= \frac{1}{t} \left[\frac{1}{2} \mathrm{Tr}(\mathbf{M}^{\mathrm{T}}\mathbf{M}\mathbf{C}_t) + \mathrm{Tr}(\mathbf{M}^{\mathrm{T}}\mathbf{D}_t) \right] + \beta\Psi(\mathbf{M})$$

$$\mathbf{C}_t = \sum_{i=1}^{t} \mathbf{A}_i\mathbf{A}_i^{\mathrm{T}}, \quad \mathbf{D}_t = \sum_{i=1}^{t} (\mathbf{dM}_i\mathbf{A}_i - \mathbf{Y}_i)\mathbf{A}_i^{\mathrm{T}}.$$

where k denotes the current iteration of the endmember update subproblem, $\mathcal{P}_+$ is the projector on $\{\mathbf{X} : \mathbf{X} \succeq \mathbf{0}_{L,K}\}$ and L_{3t} denotes the Lipschitz constant of $\nabla_{\mathbf{M}}\hat{g}_t$.

3.3. Convergence

The convergence of the generated endmember sequence $(\mathbf{M}_t)_t$ towards a critical point of the problem (7) can be ensured provided the function $f(\mathbf{Y}_t, \mathbf{M}, \cdot, \cdot)$ exclusively admits locally unique stationary points, which is the case for the function f proposed in this paper. Thereby slightly adapting the arguments used in [10] to our problem, we recover a similar convergence result. This point is detailed in the extended paper [19]. Note that adding a forgetting factor ξ as in Algorithm 1 slightly modifies the expression of the function actually minimized, but does not alter the convergence proof.

4. EXPERIMENT WITH SYNTHETIC DATA

The proposed method has been first evaluated on a dataset composed of 15 images of size 31×30, where each image is a linear mixture of $K = 3$ corrupted endmembers – generated as detailed in [8] – composed of $L = 413$ spectral bands. All the mixtures have been corrupted by an additive white Gaussian noise ensuring a signal-to-noise ratio of 30dB. Note that each image of the series does not necessarily fulfill the pure pixel assumption in order to evaluate the method in a challenging situation.

4.1. Sate-of-the-art methods

The proposed strategy was compared to two classical algorithms VCA [20] / FCLS [21, 22] and SISAL [23]/ FCLS independently applied to each image of the time series. For the proposed approach, the variability terms were initialized with all their entries equal to zero. The endmembers were initialized with the VCA algorithm applied to the union of the sets of points belonging to the convex envelope of the $(K-1)$-simplices containing the image pixels. The abundances were initialized using the FCLS algorithm.

Algorithm 1 was run for 50 cycles through the whole dataset (see [10]), and the inner PALM and proximal gradient descent algorithms were respectively stopped after $N_{\mathrm{iter}} = 50$ iterations. The values of all the other parameters are $\xi = 0.99$, $\gamma_i = 1.1$ for $i \in \{1, 2, 3\}$, $\alpha = 3.9 \times 10^{-2}$, $\beta = 5.4 \times 10^{-4}$, $\gamma = 3.2 \times 10^{-4}$, $\sigma^2 = 12.4$, $\kappa^2 = 1.9$. Note that (α, β, γ) was chosen such that each initial penalty term of the first image represents 10% of the initial data fitting term, whereas (σ^2, κ^2) was selected by cross-validation.
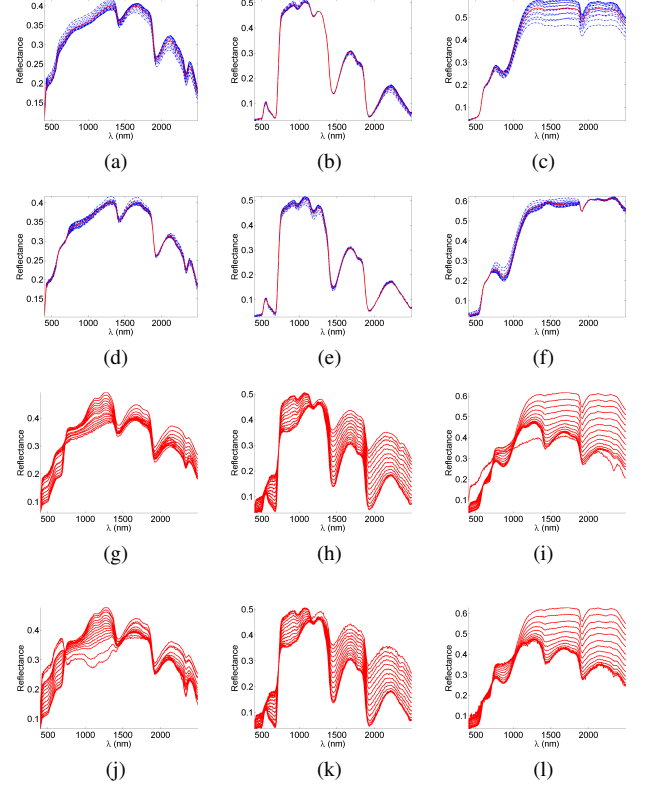

Fig. 1. Endmembers (red lines) and variability (blue dotted lines) [true endmembers and variability on the first line, endmembers recovered by the proposed method on the second line, VCA-extracted endmembers on the third line, SISAL-extracted endmembers on the last line].

The performance of the algorithm is assessed in this paper in terms of average spectral angle mapper (aSAM) for the endmembers, and in terms of global mean square errors (GMSEs) for the estimated abundances and perturbations. Finally, the reconstruction error (RE) providing a measure of fit between the actual images and their estimations is also considered. All these performance measures are defined below

$$\mathrm{aSAM}(\mathbf{M}) = \frac{1}{K} \sum_{k=1}^{K} \arccos \left(\frac{\langle \mathbf{m}_k | \widehat{\mathbf{m}}_k \rangle}{\|\mathbf{m}_k\|_2 \|\widehat{\mathbf{m}}_k\|_2} \right)$$

$$\mathrm{GMSE}(\mathbf{A}) = \frac{1}{TKN} \sum_{t=1}^{T} \|\mathbf{A}_t - \widehat{\mathbf{A}}_t\|_{\mathrm{F}}^2$$

$$\mathrm{GMSE}(\mathbf{dM}) = \frac{1}{TLK} \sum_{t=1}^{T} \|\mathbf{dM}_t - \widehat{\mathbf{dM}}_t\|_{\mathrm{F}}^2$$

$$\mathrm{RE} = \frac{1}{TLN} \sum_{t=1}^{T} \|\mathbf{Y}_t - \widehat{\mathbf{Y}}_t\|_{\mathrm{F}}^2$$

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

4.2. Results

The performance of the different unmixing methods are provided in Table 1. The proposed approach provides competitive results when
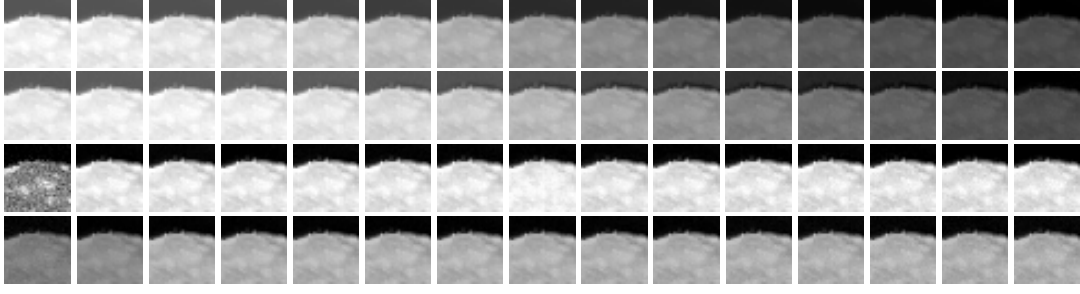
Fig. 2. First endmember abundance maps [0: black, 1: white] (true maps on the first row, proposed method on the second, VCA/FCLS on the third, SISAL/FCLS on the fourth). Each column corresponds to a time instant t.
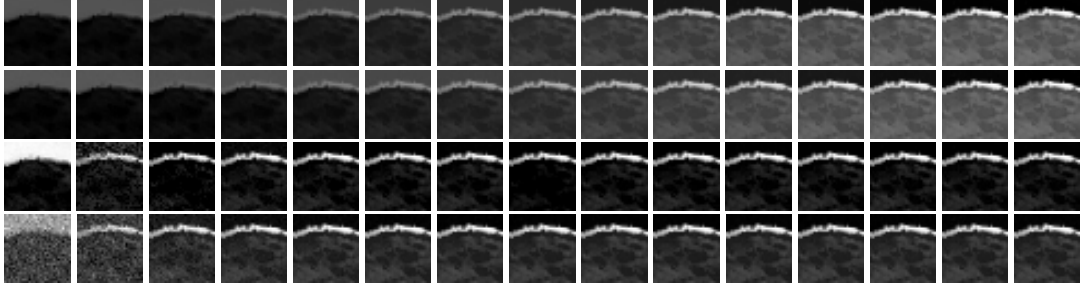

Fig. 3. Second endmember abundance maps [0: black, 1: white] (true maps on the first row, proposed method on the second, VCA/FCLS on the third, SISAL/FCLS on the fourth). Each column corresponds to a time instant t.

Fig. 4. Third endmember abundance maps [0: black, 1: white] (true maps on the first row, proposed method on the second, VCA/FCLS on the third, SISAL/FCLS on the fourth). Each column corresponds to a time instant t.
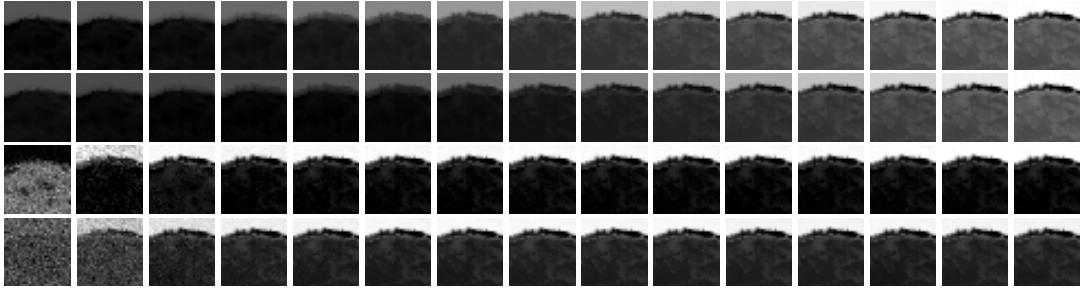
compared to other methods and exhibits lower aSAMs and GMSEs, at the price of a higher computational cost. As expected, the abundance maps returned by the proposed approach are more convincing than the ones obtained by VCA/FCLS or SISAL/FCLS. Indeed, the proposed unmixing method accurately renders the smoothness of the abundance evolution by sequentially exploiting all the images and accounting for variability, whereas the two classical methods do not. Note that the improved estimation performance reported in Table 1 mainly results from the multi-temporal information exploited by the online unmixing strategy.

5. CONCLUSION AND FUTURE WORK

This paper introduced an online hyperspectral unmixing procedure accounting for endmember temporal variability based on the perturbed linear model proposed in [8]. The resulting unmixing problem was formulated as a two-stage stochastic program solved by an online algorithm to avoid the use of batch procedures when con-

fronted to large hyperspectral images. Simulations conducted on synthetic data allowed the interest of the proposed approach to be appreciated. Applying this method to large real data and addressing abrupt endmember changes are currently investigated. Another possible perspective would consist in incorporating spatial variability into the proposed method.

Table 1. Simulation results obtained with synthetic data ($GMSE(\mathbf{A}) \times 10^{-2}$, $GMSE(\mathbf{dM}) \times 10^{-4}$, $RE \times 10^{-5}$).

	VCA/FCLS	SISAL/FCLS	Prop. method
aSAM($\mathbf{M}$) (°)	8.9792	8.6685	**1.9898**
GMSE($\mathbf{A}$)	6.67	3.90	**0.47**
GMSE($\mathbf{dM}$)	/	/	**3.07**
RE	9.59	**9.49**	9.63
time (s)	**2**	2.2	561

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 Observ. in 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] 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 Observ. in Remote Sens.*, vol. 5, no. 2, pp. 396–408, April 2012.

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

[5] 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. in Remote Sens.*, vol. 7, no. 6, pp. 1994–2003, June 2014.

[6] A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Unsupervised unmixing of hyperspectral images accounting for endmember variability," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4904–4917, Dec. 2015.

[7] S. Henrot, J. Chanussot, and C. Jutten, "Dynamical spectral unmixing of multitemporal hyperspectral images," Oct. 2015, arXiv preprint. [Online]. Available: http://arxiv.org/pdf/1510.04238v1.pdf

[8] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 525–538, Jan. 2016.

[9] D. Ralph and H. Xu, "Convergence of stationary points of sample average two-stage stochastic programs: A generalized equation approach," *Mathematics of Operations Research*, vol. 36, no. 3, pp. 568–592, Aug. 2011.

[10] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *J. Mach. Learning Research*, vol. 11, pp. 19–60, Jan. 2010.

[11] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Mathematical Programming*, vol. 1-2, no. 146, pp. 459–494, July 2013.

[12] A. Halimi, N. Dobigeon, J.-Y. Tourneret, S. McLaughlin, and P. Honeine, "Unmixing hyperspectral images accounting for temporal and spatial endmember variability," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Nice, France, 2015, pp. 1686–1690.

[13] 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. Remote Sens.*, vol. 42, no. 10, pp. 2085–2095, Oct. 2004.

[14] M. Arngren, M. N. Schmidt, and J. Larsen, "Unmixing of hyperspectral images using Bayesian nonnegative matrix factorization with volume prior," *J. Signal Process. Sys.*, vol. 65, no. 3, pp. 479–496, Nov. 2011.

[15] L. Condat, "Fast projection onto the simplex and the ℓ_1 ball," *Math. Program., Ser. A*, pp. 1–11, Sept. 2015. [Online]. Available: http://dx.doi.org/10.1007/s10107-015-0946-6

[16] J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra, "Efficient projection onto the ℓ_1-ball for learning in high dimensions," in *Proc. Int. Conf. Machine Learning (ICML)*, Helsinki, Finland, 2008.

[17] J. P. Boyle and R. L. Dykstra, "A method for finding projections onto the intersection of convex sets in hilbert spaces," in *Advances in Order Restricted Statistical Inference*. Springer New York, 1986, pp. 28–47.

[18] R. Heylen, A. Atker, and P. Scheunders, "On using projection onto convex sets for solving the hyperspectral unmixing problem," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1522–1526, 2013.

[19] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Online unmixing of multitemporal hyperspectral images accounting for spectral variability," Oct. 2015, submitted. [Online]. Available: http://arxiv.org/abs/1510.05893

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

[21] 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 Sens.*, vol. 29, no. 3, pp. 529–545, March 2001.

[22] 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)*, Reykjavik, Iceland, June 2010.

[23] J. M. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Process.: Evolution in Remote Sens. (WHISPERS)*, Grenoble, France, Aug. 2009.