UNMIXING HYPERSPECTRAL IMAGES USING THE GENERALIZED BILINEAR MODEL

Abderrahim Halimi, Yoann Altmann, Nicolas Dobigeon and Jean-Yves Tourneret

University of Toulouse, IRIT/INP-ENSEEIHT/TéSA, Toulouse, France

{abderrahim.halimi, yoann.altmann, nicolas.dobigeon, jean-yves.tourneret}@enseeiht.fr

ABSTRACT

Nonlinear models have recently shown interesting properties for spectral unmixing. This paper considers a generalized bilinear model recently introduced for unmixing hyperspectral images. Different algorithms are studied to estimate the parameters of this bilinear model. The positivity and sum-to-one constraints for the abundances are ensured by the proposed algorithms. The performance of the resulting unmixing strategy is evaluated via simulations conducted on synthetic and real data.

Index Terms— hyperspectral imagery, spectral unmixing, bilinear model, Bayesian inference, MCMC methods, gradient descent algorithm, least square algorithm.

1. INTRODUCTION

Spectral unmixing is one of the major issues when analyzing hyperspectral images. Unmixing hyperspectral images is based on the assumption that a pixel spectrum is a combination of pure spectral components (referred to as endmembers). The underlying mixture model can be linear or nonlinear. The linear mixture model (LMM) has been widely used in the literature and has shown promising results [1]. However, the LMM can be inappropriate for some hyperspectral images where the detected photons interact with multiple components before they reach the sensor. In this case, nonlinear models can be more interesting for abundance estimation, e.g., for scenes including mixtures of orchards [2] or vegetation [3].

This paper considers a generalized bilinear model (GBM) introduced in [4] for nonlinear unmixing of hyperspectral images. This model has shown good properties for modeling non-linear interactions between pure spectral components. However, estimating the abundances associated with this nonlinear model is a challenging problem. This paper studies two approaches to address this estimation problem. The first approach is based on a Bayesian model constructed from appropriate prior distributions for the abundances and the non-linear mixing coefficients. The parameter priors are chosen to ensure positivity and sum-to-one constraints for the abundances and physical constraints about the non-linear mixing coefficients. The joint posterior distribution of the model parameters is then derived. The minimum mean square error (MMSE) estimator is then computed from samples distributed according to this posterior generated by Markov chain Monte Carlo (MCMC) methods. The second approach is based on the minimization of a cost function under appropriate constraints for the abundances and non-linear mixing coefficients. Two optimization algorithms are considered. The first one is inspired from the works of [5] and [6] and relies on a Taylor series expansion of the model nonlinearity. The second algorithm uses a constrained gradient descent method coupled with a line search technique as in [7].

The paper is organized as follows. Section 2 introduces the linear and bilinear models considered in this work. Section 3 presents algorithms for estimating the parameters of these models. Simulation results for synthetic and real images are analyzed in Sections 4 and 5. Conclusions and future works are reported in Section 6.

2. UNMIXING MODELS

The physical assumption underlying the LMM is that each incident photon interacts with only one earth surface component. In this case, the L-spectrum $\boldsymbol{y} = [y_1, \ldots, y_L]^T$ of an observed pixel can be expressed as a mixture of R endmembers $\boldsymbol{m}_k$ with additive noise [8]

$$\boldsymbol{y} = \sum_{k=1}^{R} \alpha_k \boldsymbol{m}_k + \boldsymbol{n} = \boldsymbol{M}\boldsymbol{\alpha} + \boldsymbol{n} \tag{1}$$

where $\boldsymbol{M}$ is the $L \times R$ matrix whose columns are the $L \times 1$ endmember spectra $\boldsymbol{m}_k = [m_{1,k}, \ldots, m_{L,k}]^T$, $k = 1, \ldots, R$, $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_R]^T$ is the $R \times 1$ abundance vector and $\boldsymbol{n} = [n_1, \ldots, n_L]^T$ is assumed to be an independent and identically distributed (i.i.d.) zero-mean Gaussian sequence with variance σ^2.

The GBM considered in this work is a nonlinear model that accounts for the presence of second order interactions between endmember $\#i$ and endmember $\#j$ (for $i, j = 1, \ldots, R$ and $i \neq j$) (see [4] for motivations about this model). The corresponding mixed pixel $\boldsymbol{y}$ is then expressed as

$$\boldsymbol{y} = \boldsymbol{M}\boldsymbol{\alpha} + \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \gamma_{i,j} \alpha_i \alpha_j \boldsymbol{m}_i \odot \boldsymbol{m}_j + \boldsymbol{n} \tag{2}$$

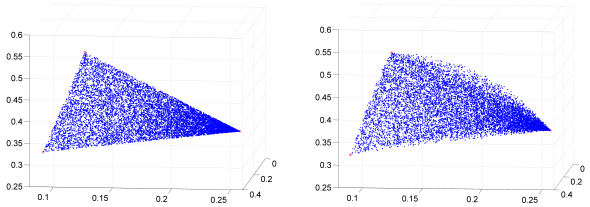
where $\odot$ denotes the Hadamard (term-by-term) product operation, i.e, $\boldsymbol{m}_i \odot \boldsymbol{m}_j = [m_{1,i}m_{1,j}, \cdots, m_{L,i}m_{L,j}]^T$, and with the following constraints for the different parameters

$$\alpha_k \geq 0, \, \forall k \in \{1, \ldots, R\} \quad \text{and} \quad \sum_{k=1}^{R} \alpha_k = 1 \tag{3}$$

$$0 \leq \gamma_{i,j} \leq 1, \, \forall i \in \{1, \ldots, R-1\}, \, \forall j \in \{i+1, \ldots, R\}. \tag{4}$$

Eq. (3) expresses the usual positivity and sum-to-one constraints for the abundances. Eq. (4) introduces constraints for the non-linear mixing coefficient $\gamma_{i,j}$ that controls the interaction between endmembers $\#i$ and $\#j$. Precisely, the parameter $\gamma_{i,j}$ is upper bounded by 1, reflecting the fact that the interaction term $\gamma_{i,j}\alpha_i\alpha_j$ is always smaller than the product of the individual abundances $\alpha_i\alpha_j$ as explained in [4]. Moreover, it makes sense to assume $\gamma_{i,j} \geq 0$ to consider constructive interferences and ensure positive observations in $\boldsymbol{y}$ (a similar assumption has been made in [5] and [2]). Note that the proposed GBM defined in (2) reduces to the standard LMM for $\gamma_{i,j} = 0, \forall (i,j)$ and to the non-linear model studied in [5] for $\gamma_{i,j} = 1, \forall (i,j)$. As a consequence, the proposed GBM can be viewed as a generalization of these two models. It is also important to note that for $\alpha_i = 1$ and $\alpha_j = 0$ for any $j \neq i$, the second order

terms of the proposed model disappear and the observed vector $\boldsymbol{y}$ reduces to the ith endmember $\boldsymbol{m}_i$. Thus, the pure materials (endmembers) are present in the data generated by the proposed GBM. Moreover, when the non linearity coefficients γ_{ij} are small (it will be the case in most real applications), the data generated by the proposed model lie into a modified simplex whose extremities are the pure materials. To illustrate the effect of endmember interactions, synthetic pixels have been generated from $R = 3$ endmembers following the proposed GBM. The resulting pixels are depicted in the hyperspectral space as blue points in Fig. 1 (right figure). From this figure, it appears that the pixels lie on a curved simplex whose vertices are the endmembers. For comparison, we also show the simplex associated with the LMM obtained by removing the second-order terms (left figure).



(a) LMM. (b) GBM.

Fig. 1. Data generated according to the mixing models.

3. UNMIXING ALGORITHMS

This section studies two approaches for estimating the GBM parameters. We denote as $\boldsymbol{\theta} = \left(\boldsymbol{\alpha}^T, \boldsymbol{\gamma}^T\right)^T$ the unknown GBM parameter vector, where $\boldsymbol{\alpha}$ is the abundance vector and $\boldsymbol{\gamma} = [\gamma_{1,2}, \cdots, \gamma_{R-1,R}]^T$ is the nonlinearity coefficient vector.

3.1. Bayesian Algorithm

The Bayesian algorithm studied in this Section has been initially introduced in [4]. Bayesian estimators are computed from the posterior distribution of $\boldsymbol{\theta}$ denoted as $f(\boldsymbol{\theta}|\boldsymbol{y})$. This posterior distribution is related to the likelihood of the observations $f(\boldsymbol{y}|\boldsymbol{\theta})$ and the parameter prior distribution $f(\boldsymbol{\theta})$ via Bayes' theorem

$$f(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{f(\boldsymbol{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{f(\boldsymbol{y})} \propto f(\boldsymbol{y}|\boldsymbol{\theta})f(\boldsymbol{\theta}) \tag{5}$$

where $\propto$ stands for "proportional to". As a consequence, computing the Bayes estimators require to define prior distribution for the unknown parameter vector (summarized in $f(\boldsymbol{\theta})$) as well as the likelihood related to the observation.

3.1.1. Likelihood

The observation model defined in (2) and the Gaussian properties of the noise sequence $\boldsymbol{n}$ yield the following likelihood function

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{L}{2}} \exp\left[-\frac{||\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{\theta})||^2}{2\sigma^2}\right] \tag{6}$$

where

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \boldsymbol{M}\boldsymbol{\alpha} + \sum_{i=1}^{R-1}\sum_{j=i+1}^{R} \gamma_{i,j}\alpha_i\alpha_j\boldsymbol{m}_i \odot \boldsymbol{m}_j \tag{7}$$

and $||\cdot||$ denotes the standard l_2 norm such that $||\boldsymbol{x}||^2 = \boldsymbol{x}^T\boldsymbol{x}$.

3.1.2. Parameter priors

To satisfy the sum-to-one constraint (3), one can express one abundance as a function of the others, e.g., $\alpha_R = 1 - \sum_{k\neq R}\alpha_k$. A uniform distribution on the simplex defined as

$$\mathcal{S} = \left\{ \boldsymbol{\alpha}_{1:R-1} \middle| \alpha_k \geq 0, \forall k \neq R \text{ and } \sum_{k\neq R}\alpha_k \leq 1 \right\}$$

is then assigned to the reduced abundance vector $\boldsymbol{\alpha}_{1:R-1} = (\alpha_1, ..., \alpha_{R-1})^T$. Moreover, to satisfy the constraints (4), uniform priors on $(0, 1)$ have been considered for the non-linearity coefficients $\gamma_{i,j}$ assumed to be a priori independent (see [4]).

3.1.3. Posterior distribution

The posterior distribution of the unknown parameter vector $\boldsymbol{\theta}$ can be computed from Bayes theorem according to (5). Unfortunately, this posterior distribution is too complex to derive closed-form expressions for standard Bayesian estimators such as the minimum mean square error (MMSE) estimator or the maximum a posteriori (MAP) estimator. Instead, we can use Markov chain Monte Carlo (MCMC) methods to generate samples according to the posterior distribution of $\boldsymbol{\theta}$. These simulated samples are then used to approximate the Bayesian estimators of $\boldsymbol{\theta}$ (see [4] for more details).

3.2. Optimization algorithms

The main problem with the MCMC method advocated in the previous Section is its computational complexity. This section studies alternative estimation strategies based on optimization algorithms allowing computational cost to be significantly reduced. More precisely, estimating the GBM parameter vector can be formulated as the following optimization problem

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} ||\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{\theta})||^2 \tag{8}$$

subject to (s.t.) the constraints given by (3) and (4), where $\boldsymbol{\mu}(\boldsymbol{\theta})$ has been defined in (7). This paper proposes two algorithms to solve this minimization problem.

3.2.1. Fan-FCLS algorithm (FFA)

The optimization problem (8) reveals two main difficulties: i) the nonlinearities of the mixture model and ii) the constraints on the parameters. These difficulties will be addressed consecutively in the following paragraphs.

a) Nonlinearity: to overcome the difficulty related to the nonlinearity defined in (2), we propose to linearize the objective criterion by using a first order Taylor series expansion of $\boldsymbol{\mu}(\boldsymbol{\theta})$. This approach leads to an iterative algorithm similar to the algorithm proposed by Fan *et al.* in [5]. At iteration t of the algorithm, i.e., at the given point estimate $\boldsymbol{\theta}^{(t)} = \left(\alpha_1^{(t)}, \cdots, \alpha_R^{(t)}, \gamma_{1,2}^{(t)}, \cdots, \gamma_{R-1,R}^{(t)}\right)$, the linearization yields

$$\boldsymbol{\mu}(\boldsymbol{\theta}) \approx \boldsymbol{\mu}\left(\boldsymbol{\theta}^{(t)}\right) + \left.\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(t)}} \left(\boldsymbol{\theta} - \boldsymbol{\theta}^{(t)}\right), \tag{9}$$

where $\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\theta}}$ is the gradient of $\boldsymbol{\mu}$, i.e., a matrix of size $(L \times R(R+1)/2)$ containing the derivatives $\frac{\partial\mu_i}{\partial\theta_j}$. Straightforward computations lead to the following iterative updating rule for the parameter vector $\boldsymbol{\theta}$

$$\boldsymbol{\theta}^{(t+1)} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \left|\left|\boldsymbol{h}^{(t)} - \boldsymbol{P}\boldsymbol{\theta}\right|\right|^2 \text{ s.t. (3) and (4)} \tag{10}$$

where the elements of the matrix $\boldsymbol{P}$ are $p_{i,j} = \frac{\partial \mu_i}{\partial \theta_j}$ and

$$\boldsymbol{h}^{(t)} = \left(\boldsymbol{y} - \boldsymbol{\mu}\left(\boldsymbol{\theta}^{(t)}\right) + \boldsymbol{P}\boldsymbol{\theta}^{(t)}\right). \tag{11}$$

b) Constraints: to solve (10), a strategy inspired by [6] is adopted. First, we introduce a vector with positive components $\boldsymbol{w} = [w_1, \ldots, w_S]^T$ such that $\boldsymbol{w} = \mathbf{1}_S - \boldsymbol{\gamma}$, where $\mathbf{1}_S = [1, \ldots, 1]^T$ is the $S \times 1$ vector of ones, and $S = R(R-1)/2$ is the number of nonlinearity coefficients[1] γ_s. The constraints (4) can then be expressed as

$$\left\{ \begin{array}{ll} \gamma_s \geq 0, & \forall s = 1, \ldots, S \\ w_s \geq 0, & \text{and} \quad w_s + \gamma_s = 1, \quad \forall s = 1, \ldots, S. \end{array} \right. \tag{12}$$

The optimization problem (10) is finally reformulated to apply the fully constrained least squares (FCLS) algorithm of [6]. The $(S+1)$ constraints $\sum_{i=1}^{R} \alpha_i = 1$ and $\boldsymbol{w} + \boldsymbol{\gamma} = \mathbf{1}_S$ can be handled by introducing the extended matrix $\tilde{\boldsymbol{P}}$

$$\tilde{\boldsymbol{P}} = \left(\begin{array}{ccc} \delta \boldsymbol{P} & & \mathbf{O}_{L,S} \\ \mathbf{O}_{S,R} & \mathbf{I}_S & \mathbf{I}_S \\ \mathbf{1}_R^T & \mathbf{0}_S^T & \mathbf{0}_S^T \end{array} \right) \tag{13}$$

and the extended vectors $\tilde{\boldsymbol{h}}^{(t)}$ and $\tilde{\boldsymbol{\theta}}$

$$\tilde{\boldsymbol{h}}^{(t)} = \left[\begin{array}{c} \delta \boldsymbol{h}^{(t)} \\ \mathbf{1}_S \\ 1 \end{array} \right], \quad \tilde{\boldsymbol{\theta}} = \left[\begin{array}{c} \boldsymbol{\theta} \\ \boldsymbol{\gamma} \\ \boldsymbol{w} \end{array} \right] \tag{14}$$

where $\mathbf{O}_{m,n}$ is the $m \times n$ matrix of zeroes and $\mathbf{I}_S$ is the $S \times S$ identity matrix. The parameter δ in (13) and (14) controls the impact of the equality constraints. The final updating rule is given by

$$\tilde{\boldsymbol{\theta}}^{(t+1)} = \underset{\tilde{\boldsymbol{\theta}}}{\operatorname{argmin}} \left|\left| \tilde{\boldsymbol{h}}^{(t)} - \tilde{\boldsymbol{P}}\tilde{\boldsymbol{\theta}} \right|\right|^2$$

subject to the non-negativity constraints

$$\left\{ \begin{array}{ll} \alpha_i \geq 0, & i = 1, \ldots, R \\ \gamma_s \geq 0, & s = 1, \ldots, S \\ w_s \geq 0, & s = 1, \ldots, S. \end{array} \right. \tag{15}$$

The solution of this minimizing problem is computed iteratively using the nonnegativity constrained least-square algorithm (NCLS) [6]. The stopping rule of the NCLS is $\left|\left| \boldsymbol{\theta}^{(t+1)} - \boldsymbol{\theta}^{(t)} \right|\right|^2 < \rho$, where ρ is a given threshold. The iterative procedure can be initialized by

$$\boldsymbol{\theta}^{(0)} = \left(\alpha_1^{(0)}, \ldots, \alpha_R^{(0)}, 0, \ldots, 0\right), \tag{16}$$

which is the solution obtained when considering the LMM.

3.2.2. Gradient descent algorithm (GDA)

An alternative of the linearization-based approach described above consists of using a gradient descent algorithm (GDA). The sum-to-one constraint for the abundances can be handled by expressing one abundance as a function of the others, e.g., $\alpha_R = 1 - \sum_{k \neq R} \alpha_k$. Then the GDA is used to minimize a cost function

$$J\left(\boldsymbol{\theta}_R\right) = \left\| \boldsymbol{y} - \boldsymbol{\mu}\left(\boldsymbol{\theta}\right) \right\|^2$$

[1]Note that, for conciseness, the nonlinearity coefficients $\gamma_{i,j}$ are now indexed as γ_s with $s = 1, \ldots, S$.

with respect to $\boldsymbol{\theta}_R = \left(\boldsymbol{\alpha}_{1:R-1}^T, \boldsymbol{\gamma}^T\right)^T$, using the following iterative scheme

$$\boldsymbol{\theta}_R^{(t+1)} = \boldsymbol{\theta}_R^{(t)} - \lambda^{(t)} \frac{\partial J}{\partial \boldsymbol{\theta}_R}. \tag{17}$$

The step parameter $\lambda^{(t)}$ is adjusted by a constrained line search procedure as in [7]. The principle of the line search is to determine the parameter $\lambda^{(t)}$ that minimizes the function $J(\cdot)$ in the gradient direction. The constraints (3) and (4) will be satisfied by enforcing the parameter $\lambda^{(t)}$ to belong to a bounded set. Note that the line search procedure has been performed by using the golden section method [7]. Note also that the algorithm has been initialized by (16).

4. SIMULATION RESULTS FOR SYNTHETIC IMAGES

Four synthetic images have been used to evaluate the three algorithms introduced above. The first image denoted as I_1 has been generated according to the LMM. The two images I_2 and I_3 have been obtained from the non-linear mixing models defined in [5] and (2). The last image denoted as I_4 has been decomposed into two different regions whose pixels satisfy the LMM and the GBM. The abundance vectors $\boldsymbol{\alpha}(p)$ $(p = 1, \ldots, 100)$ have been generated uniformly on the simplex. The nonlinearity coefficients are uniformly drawn in the set $[0, 1]$ for the GBM. All images (of size 10×10) have been corrupted by an additive Gaussian noise of variance $\sigma^2 = 2.8 \times 10^{-3}$. To compare the different algorithms, we propose three measures of performance. First, the reconstruction error (RE) is used to measure the distance between the measured pixel $\boldsymbol{y}(p)$ and the estimated spectrum $\hat{\boldsymbol{y}}(p)$

$$\text{RE} = \sqrt{\frac{1}{n\,L} \sum_{p=1}^{n} \left\| \hat{\boldsymbol{y}}(p) - \boldsymbol{y}(p) \right\|^2}. \tag{18}$$

The spectral angle mapper (SAM) is also used to estimate the performance of the unmixing procedure [1]

$$\text{SAM} = \frac{1}{n} \sum_{p=1}^{n} \arccos\left(\frac{\langle \boldsymbol{y}(p), \hat{\boldsymbol{y}}(p) \rangle}{\|\boldsymbol{y}(p)\| \ \|\hat{\boldsymbol{y}}(p)\|} \right) \tag{19}$$

where $\arccos(\cdot)$ denotes the inverse cosine operator. The third criterion compares the estimated and actual abundances by using the root mean square error (RMSE) [9]

$$\text{RMSE} = \sqrt{\frac{1}{n\,R} \sum_{p=1}^{n} \left\| \boldsymbol{\alpha}(p) - \hat{\boldsymbol{\alpha}}(p) \right\|^2} \tag{20}$$

where $\boldsymbol{\alpha}(p)$ and $\hat{\boldsymbol{\alpha}}(p)$ are the actual and estimated pth abundance vectors of the image and n is the number of pixels (here $n = 100$).

Table 1 (left) shows the REs and SAMs obtained when unmixing the synthetic images with the GBM. The different algorithms (FFA, GDA and Bayesian) perform very similarly for these examples. The corresponding RMSEs between the actual and estimated abundances are provided in Table 2. Even if the unmixing performances are very similar for all algorithms, the FFA seems to provide better performance for abundance estimation. Finally, it is interesting to note that the Bayesian algorithm provides posterior distributions for the unknown parameters that can be used to determine uncertainties about the estimations (such as standard deviations or confidence intervals). However, the price to pay with the Bayesian approach is its high computational complexity as illustrated in Table 1 (right).

Table 1. RE and SAM obtained with the 3 algorithms.

	RE ($\times 10^{-2}$)			SAM ($\times 10^{-2}$)		
	Bay	FFA	GDA	Bay	FFA	GDA
I_1	5.75	**5.47**	5.49	16.12	**15.55**	15.58
I_2	**5.44**	**5.44**	5.53	**13.93**	**13.93**	14.15
I_3	5.55	**5.47**	5.50	14.87	**14.70**	14.78
I_4	5.65	**5.49**	5.51	15.42	**15.08**	15.14

Table 2. RMSEs (left) between actual and estimated abundances and computational times (right).

	RMSEs ($\times 10^{-2}$)			Time (s.)		
	Bay.	FFA	GDA	Bay.	FFA	GDA
I_1	1.86	**1.59**	2.76	236	**16.54**	32.95
I_2	7.73	**1.49**	5.34	258	**19.90**	26.74
I_3	4.02	**1.81**	2.93	240	18.67	**17.40**
I_4	3.42	**1.98**	2.82	241	**17.96**	23.10

5. SPECTRAL UNMIXING OF AN AVIRIS IMAGE

This section considers a real hyperspectral image acquired over Moffett Field (CA, USA) in 1997 by the JPL spectro-imager AVIRIS. A 50×50 subimage (that has received much attention in the literature) has been unmixed using the proposed approaches. The subimage is mainly composed of $R = 3$ spectral components (water, soil and vegetation). These spectral components have been extracted by the vertex component analysis (VCA) [10]. Table 3 shows the unmixing results when the AVIRIS image has been processed using the proposed methods. Note that the unmixing performance is related to the computational complexity (e.g., see the values of SAM). Fig. 2 shows examples of abundance maps obtained with the Bayesian algorithm. The GBM allows the abundances of the three pure materials (top figures) to be recovered. However, it also provides interactions between the materials as illustrated in the three bottom figures.

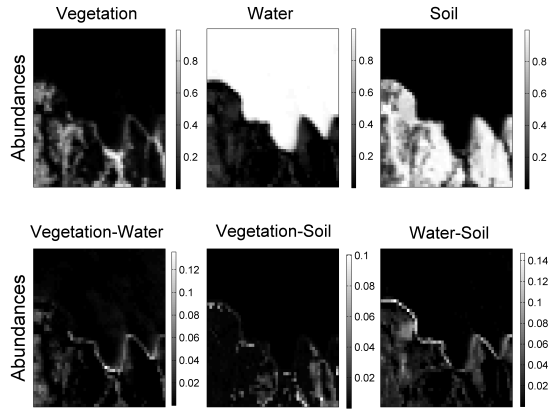


Fig. 2. Fraction maps estimated using GBM for Moffett Field.

Table 3. Unmixing performance for a real AVIRIS image.

	RE $\times 10^{-2}$	SAM $\times 10^{-1}$	Time (s.)
FFA	1.85	1.849	**49.78**
GDA	**1.80**	1.841	143.58
Bayesian	1.82	**1.837**	4469.1

6. CONCLUSION

This paper studied different estimation algorithms for the generalized bilinear model. The first algorithm was based on the Bayesian estimation principle. It provided point estimates for the unknown model parameters as well as information about estimation uncertainties. The other algorithms relied on optimization theory. They only provided point estimates of the parameters but at the price of a reduced computational cost. The performance in terms of spectral unmixing show to be very similar for the different algorithms. Future works include the introduction of spatial correlation in the generalized bilinear model. Using the correlations between adjacent pixels of an image has shown interesting properties for the *linear* unmixing of hyperspectral images. Determining the impact of spatial information on *non-linear* unmixing is clearly an interesting problem.

7. REFERENCES

[1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, pp. 44–57, Jan. 2002.

[2] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. V. der Zande, W. W. Verstraeten, and P. Coppin, "Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards," *Remote Sensing of Environment*, vol. 113, pp. 1183–1193, Feb. 2009.

[3] J. M. Bioucas-Dias and J. M. P. Nascimento, "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Conference on Image and signal processing for remote sensing*, L. Bruzzone, C. Notarnicola, and F. Posa, Eds., vol. 7477. SPIE, 2009.

[4] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. and Remote Sensing*, 2011, to appear.

[5] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *International Journal of Remote Sensing*, vol. 30, no. 11, pp. 2951–2962, June 2009.

[6] 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.

[7] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear programming: Theory and algorithms*. Canada: John Wiley and Sons, 1993.

[8] N. Dobigeon, J.-Y. Tourneret, and C.-I Chang, "Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Processing*, vol. 56, no. 7, pp. 2684–2695, July 2008.

[9] J. Plaza, A. Plaza, R. Pérez, and P. Martinez, "Joint linear/nonlinear spectral unmixing of hyperspectral image data," in *Proc. IEEE Int. Conf. Geosci. and Remote Sensing (IGARSS)*, 2007, pp. 4037–4040.

[10] 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.