AN NCM-BASED BAYESIAN ALGORITHM FOR HYPERSPECTRAL UNMIXING

Olivier Eches, Nicolas Dobigeon and Jean-Yves Tourneret

University of Toulouse, IRIT/INP-ENSEEIHT, 2 rue Camichel, 31071 Toulouse cedex 7, France

{olivier.eches,jean-yves.tourneret,nicolas.dobigeon}@enseeiht.fr

ABSTRACT

This paper studies a new Bayesian algorithm to unmix hyperspectral images. The algorithm is based on the recent normal compositional model introduced by Eismann. Contrary to the standard linear mixing model, the endmember spectra are assumed to be random signatures with know mean vectors. Appropriate prior distributions are assigned to the abundance coefficients to ensure the usual positivity and sum-to-one constraints. However, the resulting posterior distribution is too complex to obtain a closed form expression for the Bayesian estimators. A Markov chain Monte Carlo algorithm is then proposed to generate samples distributed according to the full posterior distribution. These samples are used to estimate the unknown model parameters. Several simulations are conducted on synthetic and real data to illustrate the performance of the proposed method.

Index Terms— hyperspectral imagery, spectral unmixing, Normal Compositional Model, Bayesian estimation.

1. INTRODUCTION

A major problem when analyzing hyperspectral images consists of decomposing a measured pixel spectrum as a mixture of macroscopic components referred to as *endmembers*. Most unmixing algorithms proceed in two steps: the first step estimates the endmembers using an endmember extraction procedure (such as the well-known N-finder (N-FINDR) algorithm developed by Winter [1] or the vertex component analysis (VCA) introduced by Nascimento [2]) and the second step estimates the fractions of this mixture (referred to as *abundances*) assuming the endmembers are known (see for instance [3] and references therein).

The most frequent observation model used for spectral unmixing is the linear mixing model (LMM). The LMM assumes that the spectrum of a given pixel is a linear combination of endmembers. Due to obvious physical considerations, the LMM abundances have to satisfy positivity and sum-to-one constraints. However, the LMM have some drawbacks when used on real hyperspectral images [3]. For instance, the endmember extraction procedures based on the LMM can be inefficient when the image does not contain enough pure pixels [4]. A new model referred to as normal compositional model (NCM) was recently proposed in [5]. The NCM allows one to alleviate some problems mentioned above by assuming that the pixels of the hyperspectral image are linear combinations of random endmembers (as opposed to deterministic endmmebers for the LMM) with known means (e.g. resulting from the N-FINDR or the VCA algorithms). Thus, it allows more flexibility regarding the observed pixels and the endmembers. The NCM assumes the spectrum of a mixed pixel can be written

$$\boldsymbol{y} = \sum_{r=1}^{R} \alpha_r \boldsymbol{\varepsilon}_r \tag{1}$$

where α_r is the abundance coefficient of the r^{th} material in the pixel, R is the number of endmembers present in the observed scene and $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_R$ are independent Gaussian vectors with known means, e.g., extracted from a spectral library or estimated by an appropriate method (N-FINDR or VCA). First, we assume that the covariance matrix of each endmember can be written $\sigma^2 \boldsymbol{I}_L$, where $\boldsymbol{I}_L$ is the $L \times L$ identity matrix, L is the number of spectral bands and σ^2 is the endmember variance in any spectral band (an extension to endmembers with different variances will be studied in the second part of the paper). The resulting supervised unmixing problem consists of estimating the endmember variance σ^2 and the abundance coefficients α_r ($r = 1, \ldots, R$) under the following positivity and sum-to-one constraints

$$\begin{cases} \alpha_r \geq 0, \ \forall r = 1, \ldots, R \\ \sum_{r=1}^{R} \alpha_r = 1. \end{cases} \tag{2}$$

This paper studies a Bayesian unmixing strategy derived from the NCM. The proposed algorithm is based on a hierarchical Bayesian model combined with a Markov chain Monte Carlo sampling strategy. It is organized as follows. Section 2 derives the posterior distribution of the unknown parameter vector resulting from the NCM and the proposed hierarchical model. This posterior distribution is too complex to obtain closed-form expressions for the usual Bayesian estimators (such as the maximum *a posteriori* estimator or the minimum mean square error estimator). As a consequence, we resort to a Markov chain Monte Carlo (MCMC) method that generates samples distributed according to the posterior. The generated samples are then used to estimate the unknown model parameters (all details about this MCMC method are available in [4]). Section 3 extends the proposed model to endmembers with different variances. Simulation results conducted on synthetic and real data are presented in Sections 4 and 5. Conclusions are reported in Section 6.

2. HIERARCHICAL BAYESIAN MODEL

This section presents the likelihood and the priors inherent to the proposed NCM for the spectral unmixing of hyperspectral images. A particular attention is devoted to the abundance prior distribution satisfying positivity and sum-to-one constraints.

2.1. Likelihood

Assuming the endmember spectra are *a priori* independent and using the NCM in (1), the likelihood of the pixel $\boldsymbol{y}$ is defined as

$$f(\boldsymbol{y}\,|\boldsymbol{\alpha}, \sigma^2) = \frac{1}{\left[2\pi c(\boldsymbol{\alpha}^+)\right]^{L/2}} \exp\left(-\frac{\|\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{\alpha}^+)\|^2}{2c(\boldsymbol{\alpha}^+)}\right) \tag{3}$$

where $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^T\boldsymbol{x}}$ is the standard ℓ^2 norm, $\boldsymbol{\alpha}^+ = [\alpha_1, \ldots, \alpha_R]^T$, and

$$\boldsymbol{\mu}(\boldsymbol{\alpha}^+) = \sum_{r=1}^{R} \alpha_r \boldsymbol{m}_r, \quad c(\boldsymbol{\alpha}^+) = \sigma^2 \sum_{r=1}^{R} \alpha_r^2. \tag{4}$$

Note that $\boldsymbol{m}_r = [m_{r,1}, \ldots, m_{r,L}]^T$ and $\sigma^2 \boldsymbol{I}_L$ are the mean vector and the covariance matrix of $\boldsymbol{\varepsilon}_r$.

2.2. Parameter priors

2.2.1. Abundance prior

Due to the constraints in (2), the abundance vector can be rewritten as $\boldsymbol{\alpha}^+ = [\boldsymbol{\alpha}^T, \alpha_R]^T$ where $\alpha_R = 1 - \sum_{r=1}^{R-1} \alpha_r$ and $\boldsymbol{\alpha}$ lives in a simplex defined by

$$\mathbb{S} = \left\{ \boldsymbol{\alpha} \;\middle|\; \alpha_r \geq 0, \forall r = 1, \ldots, R-1, \sum_{r=1}^{R-1} \alpha_r \leq 1 \right\}. \tag{5}$$

A uniform distribution on this simplex is chosen as prior for $\boldsymbol{\alpha}$

$$f(\boldsymbol{\alpha}) \propto \mathbf{1}_{\mathbb{S}}(\boldsymbol{\alpha}), \tag{6}$$

where $\propto$ means "proportional to" and $\mathbf{1}_{\mathbb{S}}(.)$ is the indicator function defined on the set $\mathbb{S}$. This prior ensures the positivity and sum-to-one constraints and reflects the absence of other knowledge regarding the abundances.

2.2.2. Endmember variance prior

A conjugate exponential distribution is chosen as prior for $1/\sigma^2$

$$\sigma^{-2} | \delta \sim \mathcal{E}(\delta). \tag{7}$$

In (7), δ is an adjustable hyperparameter that will be estimated with the other model parameters using a hierarchical Bayesian algorithm. Hierarchical Bayesian algorithms require to define prior distributions for the hyperparameters. This paper assumes that the prior of δ is a non-informative Jeffreys' prior

$$f(\delta) \propto \frac{1}{\delta} \mathbf{1}_{\mathbb{R}^+}(\delta). \tag{8}$$

This prior reflects the absence of knowledge about δ.

2.3. Posterior distribution of $\boldsymbol{\theta}$

The likelihood and the priors defined above allow us to determine the posterior distribution of $\boldsymbol{\theta} = \left\{\boldsymbol{\alpha}, \sigma^2\right\}$

$$f(\boldsymbol{\theta} \,|\boldsymbol{y}) \propto \frac{1}{\sigma^{2R} \left[c(\boldsymbol{\alpha})\right]^{L/2}} \exp\left(-\frac{\|\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{\alpha})\|^2}{2c(\boldsymbol{\alpha})} \right) \mathbf{1}_{\mathbb{S}}(\boldsymbol{\alpha}). \tag{9}$$

The posterior distribution (9) is too complex to derive the MMSE or MAP estimators of $\boldsymbol{\theta}$. In such case, MCMC algorithm provide efficient strategies to generate samples distributed according to the distribution of interest [6]. We propose here to generate abundances and variances distributed according to the full posterior (9) with a hybrid Gibbs sampler detailed in [4]. The generated samples will then be used to approximate the standard Bayesian estimators by empirical averages.

3. EXTENSION TO ENDMEMBER SPECTRA WITH DIFFERENT VARIANCES

This section presents an extended version of the previous model to endmembers with different variances. A new vector $\boldsymbol{\sigma} = \left[\sigma_1^2, \ldots, \sigma_R^2\right]^T$ is introduced, where σ_r^2 is the rth endmember variance. By denoting as $\boldsymbol{m}_r = [m_{r,1}, \ldots, m_{r,L}]^T$ the mean of the rth endmember, the following result can be obtained

$$\boldsymbol{\varepsilon}_r | \boldsymbol{m}_r, \sigma_r^2 \sim \mathcal{N}\left(\boldsymbol{m}_r, \sigma_r^2 \boldsymbol{I}_L\right). \tag{10}$$

This new model is interesting because it yields more flexibility. The price to pay is an increased computational complexity.

3.1. Identifiability problems

If the prior distributions chosen for σ_r^2 ($r = 1, \ldots, R$) are not sufficiently informative, indeterminacy issues lead to very poor mixing properties of the Gibbs sampler [4]. To alleviate this problem, we propose to consider jointly the unmixing of P pixels, with $P \geq R$. These pixels are assumed to share the same endmember characteristics (same mean $\boldsymbol{m}_r$ and same variance σ_r^2) with different abundance vectors, i.e.,

$$\boldsymbol{y}_p = \sum_{r=1}^{R} \boldsymbol{\varepsilon}_r \alpha_{r,p}, \qquad p = 1, \ldots, P.$$

A standard matrix formulation yields

$$\boldsymbol{Y} = \boldsymbol{E}\boldsymbol{A} \tag{11}$$

where $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_P]$, $\boldsymbol{E} = [\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_R]$ and $\boldsymbol{A} = \left[\boldsymbol{\alpha}_1^+, \ldots, \boldsymbol{\alpha}_P^+\right]$. The likelihood and priors for this new model are described below.

3.2. Likelihood

The likelihood for one pixel is

$$f(\boldsymbol{y}_p | \boldsymbol{\alpha}_p, \boldsymbol{\sigma}) \propto \frac{1}{\left[c(\boldsymbol{\alpha}_p)\right]^{L/2}} \exp\left[-\frac{\|\boldsymbol{y}_p - \boldsymbol{\mu}(\boldsymbol{\alpha}_p)\|^2}{2c(\boldsymbol{\alpha}_p)} \right]$$

with $c(\boldsymbol{\alpha}_p) = \sum_{r=1}^{R} \sigma_r^2 \alpha_{r,p}^2$ and $\boldsymbol{\mu}(\boldsymbol{\alpha}_p) = \sum_{r=1}^{R} \boldsymbol{m}_r \alpha_{r,p}$. Assuming the pixel spectra $\boldsymbol{y}_p$ ($p = 1, \ldots, P$) are *a priori* independent, the joint likelihood for the set of P pixels can be written

$$f\left(\boldsymbol{Y} \,|\boldsymbol{A}, \boldsymbol{\sigma}\right) = \prod_{p=1}^{P} f(\boldsymbol{y}_p | \boldsymbol{\alpha}_p, \boldsymbol{\sigma}). \tag{12}$$

3.3. Prior distributions

Uniform distributions on the simplex (5) are chosen as prior distributions for the partial abundance vectors $\boldsymbol{\alpha}_p$. Assuming *a priori* independence between the pixels, the following result is obtained

$$f(\boldsymbol{\alpha}_p) \propto \prod_{p=1}^{P} \mathbf{1}_{\mathbb{S}}(\boldsymbol{\alpha}_p). \tag{13}$$

The prior distributions for the endmember variances are conjugate inverse Gamma distributions with a common hyperparameter δ (as in (7)) for which a Jeffreys' prior is elected (as in (8)).

4. SIMULATION RESULTS ON SYNTHETIC DATA

This section illustrates the performance of the proposed unmixing algorithms via simulations on synthetic data. The results have been obtained with pixels observed in $L = 276$ spectral bands ranging from wavelength $0.4\mu m$ to $2.5\mu m$ (from visible to near infrared).

4.1. NCM algorithm with a single endmember variance

A synthetic mixture of $R = 3$ endmembers is considered in this experiment. The means of these endmembers $\boldsymbol{m}_r$ $(r = 1, \ldots, 3)$ have been extracted from the spectral libraries distributed with the ENVI package [7]. These spectra correspond to construction green paint, bare red brick and galvanized steel metal and are depicted in Fig. 1. The endmember variance is $\sigma^2 = 0.015$. The linear mixture considered in this section is defined by $\boldsymbol{\alpha}^+ = [0.46,\ 0.34,\ 0.2]^T$. Fig. 2 shows the posterior distributions of the abundances generated by the proposed Gibbs sampler. These distributions are in good agreement with the actual values of the abundances. Fig. 3 shows the estimated posterior distribution of σ^2 that is also in good agreement with the actual endmember variance.

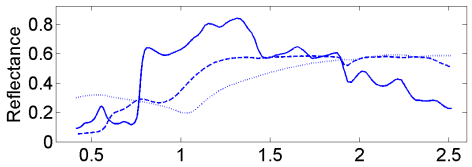

Fig. 1. Endmember spectra: construction concrete (solid line), green grass (dashed line).

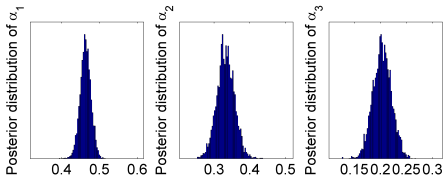

Fig. 2. Estimated posterior distributions of the abundances.

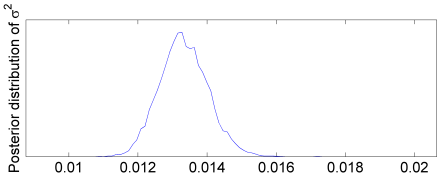

Fig. 3. Estimated posterior distribution of the variance σ^2.

4.2. NCM algorithm with different variances

The performance of the extended NCM algorithm introduced in Section 3 (and detailed in [4]) is first illustrated with synthetic data. We consider $P = 3$ pixels with $R = 3$ endmembers. The actual parameter values are

- Pixel 1: $\boldsymbol{\alpha}_1^+ = [0.1,\ 0.3,\ 0.6]^T$, $\sigma_1^2 = 0.004$,
- Pixel 2: $\boldsymbol{\alpha}_2^+ = [0.21,\ 0.52,\ 0.27]^T$, $\sigma_2^2 = 0.002$,
- Pixel 3: $\boldsymbol{\alpha}_3^+ = [0.47,\ 0.38,\ 0.15]^T$, $\sigma_2^2 = 0.0035$.

Fig. 4 shows the estimated posterior distributions of the variances σ_r^2 $(r = 1, \ldots, R)$ that are clearly centered around the actual values. The histograms of the abundances generated for each pixel by the proposed hybrid Gibbs sampler are also shown in Fig. 5. The mean square errors (MSEs) of the abundance estimates obtained in the case of one or multiple variances are reported in Table 1. These results indicate that the additional flexibility induced by considering multiple variances leads to better estimation performance.

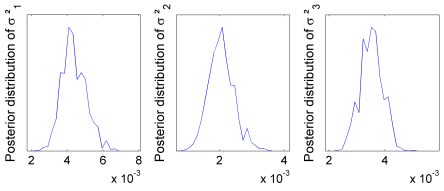

Fig. 4. Estimated posterior distribution of the variances for $P = 3$ pixels

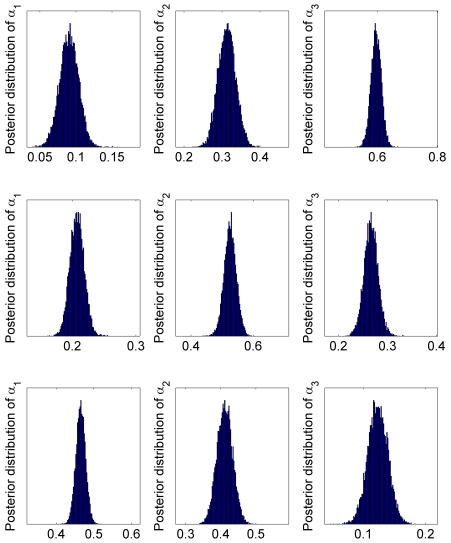

Fig. 5. Estimated posterior distributions of the abundances for each pixel (top: pixel 1, center: pixel 2, bottom: pixel 3).

Table 1. Global MSE of the abundance vector for the NCM with unique variance and with distinct variances.

NCM with single variance	NCM with multiple variances
1.72×10^{-2}	1.54×10^{-2}

5. SPECTRAL UNMIXING OF AN AVIRIS IMAGE

This section considers a real hyperspectral image of size 50×50 depicted in Fig. 6 to evaluate the performance of the different algorithms. This image has been extracted from a larger image acquired in 1997 by the airborne visible infrared imaging spectrometer (AVIRIS) over Moffett Field in Californa. The data set has been reduced from the original 224 bands to $L = 189$ bands by removing water absorption bands. First, the image has been pre-processed by a PCA to determine the number of endmembers present in the scene

as explained in [3]. Then, the N-FINDR algorithm has been applied to this image to estimate the endmember spectra. The $R = 3$ extracted endmembers (whose spectra can be found in [4]) correspond to vegetation, water and soil, and have been used as the endmember means $\boldsymbol{m}_1$, $\boldsymbol{m}_2$ and $\boldsymbol{m}_3$.

5.1. NCM algorithm with a single endmember variance

The image fraction maps estimated by the standard NCM algorithm (for the $R = 3$ pure materials) are depicted in Fig. 7 (bottom). Note that a white (resp. black) pixel in the map indicates a large (resp. small) value of the abundance coefficient. Thus, the lake area (represented by white pixels in the water fraction map and by black pixels in the other maps) can be clearly recovered. These results have been compared to the fraction maps estimated with the LMM algorithm (extracted from [8]). As depicted in Fig. 7, the fraction maps obtained with the two algorithms are clearly in good agreement. Some results regarding the estimation of the endmember variance σ^2 are also presented. Fig. 8 shows the estimated posterior distributions of σ^2 for the pixels $\#(35, 43)$ (left) and $\#(43, 35)$ (right) of the image as well as their MAP estimates. The algorithm performs satisfactorily for this example.

5.2. NCM algorithm with distinct endmember variances

The hyperspectral image has also been analyzed by considering multiple endmember variances. For this, the image has been divided into blocks of 3×3 pixels. Thus, the analyzed area has been reduced to 48×48. The estimated variances for the endmembers associated to the block centered around the pixel $\#(35, 43)$ are shown in Table 2. These results indicate that the variances can be estimated with good performance.

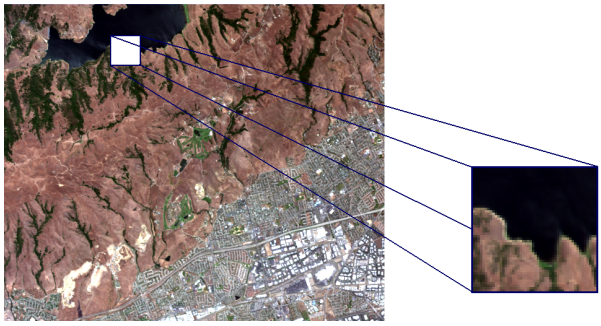

Fig. 6. Real hyperspectral data: Moffett field acquired by AVIRIS in 1997 (left) and the region of interest shown in true colors (right).

Table 2. MMSE estimates of σ_r^2 $(r = 1, \ldots, R)$.

	Soil	Vegetation	Water
MMSE estimates	1×10^{-4}	6.9×10^{-3}	1×10^{-4}

6. CONCLUSIONS

This paper studied hierarchical Bayesian algorithms based on the normal compositional model to unmix hyperspectral images. Any pixel of the image was decomposed as a random mixture of endmember spectra. Appropriate priors for the abundances were chosen to satisfy the positivity and sum-to-one constraints. Two different models were investigated. The first model assumed that the endmember spectra have the same variance. The second model introduced more flexibility by considering different variances for the endmembers. The results obtained with these two models on synthetic and real data are promising. In practical applications, we think the algorithm involving different variances should be preferred when the associated computational cost is not too prohibitive. Perspectives include the generalization of the proposed strategies to models involving spatial correlation between the pixels of the image.

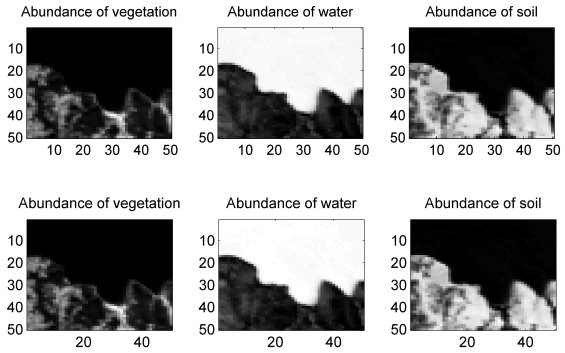

Fig. 7. Top: fraction maps estimated by the LMM algorithm (from [8]). Bottom: fraction maps estimated by the proposed algorithm (black (resp. white) means absence (resp. presence) of the material).

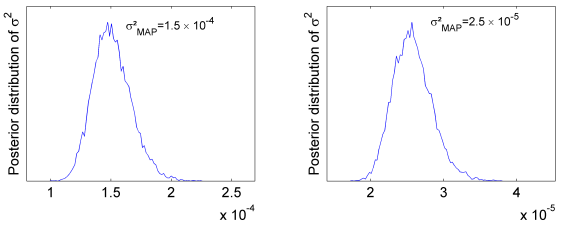

Fig. 8. Posterior distributions of the variance σ^2 for the pixels $\#(35, 43)$ (left) and $\#(43, 35)$ (right).

7. REFERENCES

[1] M. E. Winter, "Fast autonomous spectral endmember determination in hyperspectral data," in *Proc. 13th Int. Conf. on Applied Geologic Remote Sensing*, vol. 2, Vancouver, April 1999, pp. 337–344.

[2] J. M. Nascimento and J. M. B. 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.

[3] N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, pp. 44–56, Jan. 2002.

[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," University of Toulouse, Tech. Rep., March 2009. [Online]. Available: eches.perso.enseeiht.fr

[5] M. T. Eismann and D. Stein, "Stochastic mixture modeling," in *Hyperspectral Data Exploitation: Theory and Applications*, C.-I. Chang, Ed. Wiley, 2007, ch. 5.

[6] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York: Springer Verlag, 2004.

[7] RSI (Research Systems Inc.), *ENVI User's guide Version 4.0*, Boulder, CO 80301 USA, Sept. 2003.

[8] N. Dobigeon, J.-Y. Tourneret, and C.-I Chang, "Semi-supervised linear spectral using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Processing*, vol. 56, no. 7, pp. 2684–2696, July 2008.