

# Analyse d'images hyperspectrales à l'aide d'un modèle de mélange de spectres aléatoires

Olivier ECHES, Nicolas DOBIGEON, Jean-Yves TOURNERET

Université de Toulouse, IRIT/INP-ENSEEIH/TéSA  
2 rue Camichel, 31071 Toulouse cedex 7, France

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

**Résumé** – Cet article présente un nouvel algorithme de démixage appliqué aux images hyperspectrales. Chaque pixel de l'image est modélisé par une combinaison linéaire de spectres purs, appelés *endmembers*, supposés aléatoires. Plus précisément, ces spectres purs suivent des lois normales multivariées de moyennes connues (résultats d'un algorithme d'extraction d'endmembers tel que N-FINDR ou VCA). Cet article propose d'estimer les coefficients de mélange (nommés *abondances*) avec un algorithme Bayésien hiérarchique. Afin de respecter les contraintes de positivité et d'additivité, des lois *a priori* sont judicieusement choisies pour les abondances, alors qu'une loi *a priori* conjuguée est prise pour la variance. Un algorithme de Gibbs hybride permet de générer les échantillons d'abondances et de variance suivant leurs lois *a posteriori* respectives. Les performances de la méthode proposée sont évaluées à l'aide des résultats de simulations sur des images synthétiques et réelles.

**Abstract** – This paper studies a new unmixing algorithm for hyperspectral images. Each pixel of the image is modeled as a linear combination of *endmembers* which are supposed to be random in order to model uncertainties regarding their knowledge. More precisely, endmembers are modeled as Gaussian vectors with known means (resulting from an endmember extraction algorithm such as the N-FINDR and VCA). In this paper, the mixture coefficients (referred to as *abundances*) are estimated using a Bayesian algorithm. Suitable priors are assigned to the abundances in order to satisfy positivity and additivity constraints whereas a conjugate prior is chosen for the variance. The computational complexity of the resulting Bayesian estimators requires the use of a hybrid Gibbs algorithm. It generates abundance and variance samples distributed according to the posterior distribution of the unknown parameters. The performance of the proposed method is evaluated thanks to simulation results conducted on synthetic and real images.

## 1 Introduction

Le problème du démixage spectral a fait l'objet d'une attention particulière dans la littérature de traitement du signal et de l'image (voir par exemple [1] et les références citées). Les algorithmes classiques de démixage linéaire d'images hyperspectrales supposent qu'un pixel de l'image est la combinaison linéaire d'un nombre donné de spectres purs (supposés déterministes) appelés *endmembers*, pondérée par des coefficients appelés *abondances*. Ce modèle, classiquement appelé modèle de mélange linéaire (MML), a fait l'objet de nombreuses études dont certaines sont rappelées dans [1]. La plupart des algorithmes d'estimation adaptés à ce modèle MML commencent par estimer les composantes spectrales (endmembers) du modèle à l'aide d'algorithmes d'extraction comme l'algorithme N-finder (N-FINDR [2]) ou l'algorithme Vertex Component Analysis (VCA [3]). Dans une seconde étape, les algorithmes estiment les coefficients du mélange (abondances) à l'aide d'une procédure qualifiée de démixage linéaire.

Le modèle MML a donné de bons résultats dans de nombreuses applications mais il est mal adapté aux images hyperspectrales contenant peu de pixels purs. Nous proposons d'étudier un modèle de mélange de spectres aléatoires Gaussiens introduit récemment dans [4] et appelé modèle NCM (« normal

compositional model »). Ces spectres aléatoires ont des moyennes connues qui sont par exemple obtenues à l'aide des algorithmes d'extraction de spectres cités ci-dessus. Le fait de considérer des spectres aléatoires permet une plus grande flexibilité qui s'avère intéressante en l'absence de pixels purs dans l'image. Le modèle NCM suppose que le vecteur des observations associé à un pixel  $\mathbf{y}$  de l'image s'écrit :

$$\mathbf{y} = \sum_{r=1}^R \alpha_r \mathcal{E}_r, \quad (1)$$

où  $\alpha_r$  est le coefficient d'abondance du  $r$ ème spectre,  $R$  est le nombre de spectres présents dans le pixel et  $\mathcal{E}_1, \dots, \mathcal{E}_R$  sont des vecteurs Gaussiens indépendants de moyennes connues. La matrice de covariance de chaque spectre est égale à  $\sigma^2 \mathbf{I}_L$ , où  $\mathbf{I}_L$  est la matrice identité de taille  $L \times L$  et où  $\sigma^2$  est un paramètre inconnu. Les coefficients d'abondances  $\alpha_1, \dots, \alpha_R$  sont inconnus mais soumis à des contraintes de positivité et d'additivité liées à la physique du modèle (ce sont des proportions) :

$$\begin{cases} \alpha_r \geq 0, \forall r = 1, \dots, R, \\ \sum_{r=1}^R \alpha_r = 1. \end{cases} \quad (2)$$

Cet article étudie un nouvel algorithme de démixage spectral basé sur une approche Bayésienne hiérarchique. Cet algorithme permet de respecter les contraintes de positivité et

d'additivité concernant les coefficients d'abondance. Il suppose également que les spectres du mélange sont aléatoires en vertu du modèle NCM défini ci-dessus. Afin d'estimer les paramètres inconnus du modèle NCM, nous proposons d'utiliser des méthodes de Monte Carlo par chaînes de Markov (MCMC) qui permettent de simuler des vecteurs suivant la loi *a posteriori* de ces paramètres. Ces vecteurs sont alors utilisés dans une procédure d'estimation statistique Bayésienne.

## 2 Modèle Bayésien hiérarchique

### 2.1 Vraisemblance

Les paramètres inconnus à estimer sont le vecteur des abondances  $\alpha^+ = [\alpha_1, \dots, \alpha_R]^T$  et la variance  $\sigma^2$  des spectres (qui est la même pour toutes les composantes spectrales comme dans la plupart des articles sur le sujet). Ces paramètres sont regroupés dans le vecteur  $\theta = \{\alpha^+, \sigma^2\}$ . Sachant que les vecteurs  $\mathcal{E}_r$ ,  $r = 1, \dots, R$ , sont Gaussiens indépendants de moyennes respectives  $\mathbf{m}_r = [m_{r,1}, \dots, m_{r,L}]^T$  ( $r = 1, \dots, R$ ) et de matrice de covariance  $\sigma^2 \mathbf{I}_L$ , la vraisemblance s'écrit

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

où  $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}}$  est la norme  $\ell^2$  de  $\mathbf{x}$ ,  $\boldsymbol{\mu}(\alpha^+) = \sum_{r=1}^R \mathbf{m}_r \alpha_r$  et  $c(\alpha^+) = \sigma^2 \sum_{r=1}^R \alpha_r^2$ . Les lois *a priori* des paramètres inconnus vont être définies avec soin dans ce qui suit.

### 2.2 Lois a priori des paramètres

**Abondances** : Comme le vecteur des abondances est soumis aux contraintes de positivité et d'additivité, il peut se réécrire suivant la forme  $\alpha^+ = [\alpha, 1 - \sum_{r=1}^{R-1} \alpha_r]^T$ , où  $\alpha = [\alpha_1, \dots, \alpha_{R-1}]^T$ . Le vecteur  $\alpha$  appartient donc au simplexe défini par :

$$\mathbb{S} = \left\{ \alpha \mid \alpha_r \geq 0, \forall r = 1, \dots, R-1, \sum_{r=1}^{R-1} \alpha_r \leq 1 \right\}. \quad (3)$$

Nous proposons de choisir une loi uniforme sur  $\mathbb{S}$  comme loi *a priori* pour le vecteur  $\alpha$  car aucune information autre que la positivité et l'additivité n'est connue pour ce vecteur de paramètres :

$$f(\alpha) \propto \mathbf{1}_{\mathbb{S}}(\alpha), \quad (4)$$

où  $\mathbf{1}_{\mathbb{S}}(\cdot)$  est la fonction indicatrice défini sur l'ensemble  $\mathbb{S}$ . Il est à noter que n'importe quelle abondance du vecteur  $\alpha^+$ , et pas seulement la dernière  $\alpha_R$ , peut être mise de côté pour construire le vecteur  $\alpha$ .

**Variance des spectres** : La loi *a priori* de  $\sigma^2$  est une loi conjuguée inverse-Gamma :

$$\sigma^2 | \delta \sim \mathcal{IG}(\nu, \delta), \quad (5)$$

où  $\nu$  et  $\delta$  sont deux hyperparamètres ajustables. Il est admis dans cet article que  $\nu = 1$  (comme dans [5]). L'hyperparamètre  $\delta$  est estimé en utilisant un algorithme Bayésien hiérarchique. Il convient donc de définir une lois *a priori* pour cet hyperparamètre. Ainsi,  $\delta$  est muni d'une loi non-informative de Jeffreys :

$$f(\delta) \propto \frac{1}{\delta} \mathbf{1}_{\mathbb{R}^+}(\delta), \quad (6)$$

où  $\propto$  signifie « proportionnel à ». Ce choix, classique dans les problèmes d'estimation Bayésienne, reflète le manque de connaissance *a priori* sur  $\delta$ .

### 2.3 Loi a posteriori de $\theta$

La vraisemblance et les lois *a priori* définies ci-dessus permettent de déterminer la loi *a posteriori* de  $\theta = \{\alpha, \sigma^2\}$  après intégration de  $\delta$  :

$$f(\theta | \mathbf{y}) \propto \frac{1}{\sigma^{2R} [c(\alpha)]^{L/2}} \exp\left(-\frac{\|\mathbf{y} - \boldsymbol{\mu}(\alpha)\|^2}{2c(\alpha)}\right) \mathbf{1}_{\mathbb{S}}(\alpha). \quad (7)$$

La forme de cette loi est trop complexe pour pouvoir calculer les estimateurs Bayésiens classiques MMSE et MAP (notons que le numérateur et le dénominateur de terme contenu dans l'exponentielle dépendent tous les deux de  $\alpha$ ). Dans ces conditions, il est alors très courant de faire appel aux méthodes MCMC pour générer des échantillons asymptotiquement distribués suivant (7). L'échantillonneur de Gibbs requis pour cette génération est détaillé dans la section suivante. Il itère des étapes de simulation suivant  $f(\alpha | \mathbf{y}, \sigma^2, \mathbf{m}_r)$ ,  $f(\sigma^2 | \mathbf{y}, \alpha)$  et  $f(\delta | \sigma^2)$ .

## 3 Algorithme de Gibbs hybride

Il est possible de générer des échantillons suivant la loi  $f(\theta | \mathbf{y})$  grâce à un échantillonneur de Gibbs et un algorithme de Metropolis-within-Gibbs. L'échantillonneur de Gibbs est une stratégie d'échantillonnage itérative qui consiste à générer des échantillons (que l'on note  $\cdot^{(t)}$  où  $t$  est l'indice de l'échantillon) distribués suivant la loi conditionnelle de chaque paramètre. Cependant, il arrive que des échantillons soient très difficiles à générer, dû à la forme de leurs lois conditionnelles. Dans ce cas, une méthode hybride appelé Metropolis-within-Gibbs est employé : elle combine l'algorithme de Metropolis-Hastings [6] avec l'échantillonneur de Gibbs.

### 3.1 Loi conditionnelle de $\alpha$

D'après Bayes,  $f(\alpha | \mathbf{y}, \sigma^2, \mathbf{m}_r) \propto f(\mathbf{y} | \theta) f(\alpha)$ . Ce qui donne facilement :

$$f(\alpha | \mathbf{y}, \sigma_r^2, \mathbf{m}_r) \propto \left(\frac{1}{c(\alpha)}\right)^{L/2} \exp\left(-\frac{\|\mathbf{y} - \boldsymbol{\mu}(\alpha)\|^2}{2c(\alpha)}\right) \mathbf{1}_{\mathbb{S}}(\alpha), \quad (8)$$

Cependant, il est difficile de simuler  $\alpha$  directement suivant cette loi. L'utilisation d'une étape de Metropolis-within-Gibbs paraît donc appropriée. La génération d'échantillons distribués

suivant (8) est alors effectuée à l'aide d'un algorithme d'acceptation/rejet avec comme choix de loi de proposition la loi *a priori* de  $\alpha$ .

### 3.2 Loi conditionnelle de la variance $\sigma^2$

La loi conditionnelle de  $\sigma^2$  se détermine par la formule de Bayes :  $f(\sigma^2 | \mathbf{y}, \alpha) \propto f(\mathbf{y} | \theta) f(\sigma^2 | \nu, \delta)$ . Ce qui donne finalement :

$$\sigma^2 | \mathbf{y}, \alpha, \delta \sim \text{IG} \left( \frac{L}{2} + 1, \frac{\|\mathbf{y} - \mu(\alpha)\|^2}{2 \sum_{r=1}^R \alpha_r^2} + \delta \right). \quad (9)$$

### 3.3 Loi conditionnelle de l'hyperparamètre

La loi conditionnelle de  $\delta$  s'obtient de la même manière que les autres lois :  $f(\delta | \sigma^2) \propto f(\sigma^2 | \delta) f(\delta)$  :

$$\delta | \sigma^2 \sim \mathcal{G} \left( R + 1, 1 + \frac{R}{\sigma^2} \right). \quad (10)$$

## 4 Résultats de simulations

L'algorithme de démixage proposé a été testé sur des données synthétiques puis réelles. Toutes les simulations ont été obtenues pour des pixels observés dans  $L = 276$  bandes spectrales allant de  $0.4 \mu\text{m}$  à  $2.5 \mu\text{m}$  en longueur d'onde (du visible au proche infrarouge). Il est d'usage d'ignorer les premiers éléments générés par l'échantillonneur de Gibbs, appartenant aux itérations de chauffage. Dans ces simulations tout les résultats ont été obtenus avec  $N_{MC} = 25000$  dont  $N_{bi} = 5000$  itérations de chauffage.

### 4.1 Sur des données synthétiques

Afin de tester notre algorithme, un pixel synthétique a été généré avec deux spectres, i.e.  $R = 2$  (cet exemple est volontairement très simple pour limiter le nombre de lois *a posteriori* des différentes abondances). Les moyennes des spectres  $m_r$  supposées connues ont été extraites d'une bibliothèque spectrale fournie avec le logiciel ENVI [7] et représentent respectivement du béton de construction et de l'herbe verte. Le pixel synthétique à analyser a été créé à partir du modèle NCM en prenant  $\alpha^+ = [0.3 \ 0.7]^T$  et  $\sigma^2 = 0.01$  (SNR = 18dB). La figure 1 montre les lois *a posteriori* des divers paramètres inconnus estimées à l'aide des vecteurs générés par l'échantillonneur de Gibbs. Ces lois *a posteriori* sont clairement en accord avec les valeurs réelles de ces paramètres.

Les performances du modèle NCM ont ensuite été comparées avec celles du modèle MML (à spectres déterministes). Pour effectuer cette comparaison, nous avons estimé les spectres du mélange à l'aide de l'algorithme VCA. Pour le modèle MML, nous supposons que les observations sont un mélange de ces spectres auquel se rajoute un bruit additif blanc Gaussien. Plusieurs algorithmes d'estimation des abondances sont alors disponibles dans la littérature pour ce modèle. Nous nous sommes

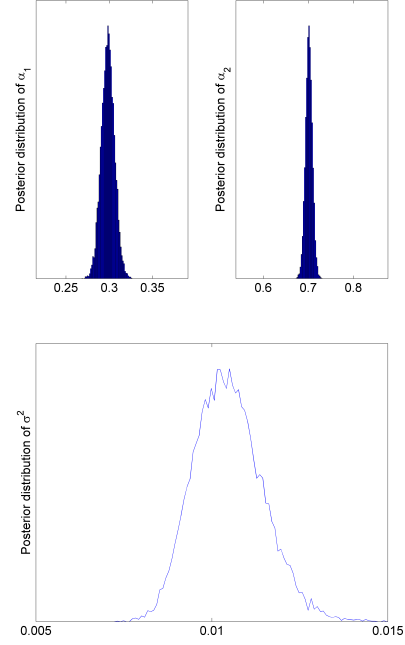


FIG. 1 – Lois *a posteriori* de :  $\alpha_r$  (en haut),  $\sigma^2$  (en bas).

concentrés sur l'algorithme FCLS (« fully constrained least squares ») étudié dans [8] et un algorithme Bayésien récemment publié dans [9]. Pour le modèle NCM, nous supposons que les spectres issus de l'algorithme VCA sont les vecteurs moyennes des vecteurs aléatoires  $\mathcal{E}_r$  pour  $r = 1, \dots, R$ . Les coefficients du mélange sont alors estimés à l'aide de l'algorithme proposé dans cet article. Les simulations ont été effectuées avec  $P = 625$  pixels contenant  $R = 6$  spectres (avec  $L = 276$  bandes spectrales). L'erreur quadratique moyenne (EQM) des vecteurs d'abondance déterminée pour chacun des algorithmes est donnée dans le tableau 1. Ces résultats indiquent que le modèle NCM donne de meilleurs résultats que le modèle MML pour cette simulation. Ceci peut s'expliquer par le fait que l'image considérée ici contient peu de pixels purs. En effet, en attribuant une composante aléatoire aux spectres, une plus grande flexibilité est permise par le modèle NCM.

### 4.2 Sur des images hyperspectrales réelles

Dans ce paragraphe sont donnés les résultats des simulations sur une image hyperspectrale de taille  $50 \times 50$  représentée sur la Fig. 2(en haut à gauche). Cette image provient d'une scène acquise en 1997 par le spectro-imageur AVIRIS au dessus de Moffett Field en Californie (Etats-Unis). Le nombre de bandes a été réduit de  $L = 224$  bandes à  $L = 189$  bandes en supprimant les bandes d'absorption d'eau. Dans un premier temps, un pré-traitement par une analyse en composantes principales (ACP) a été mis en oeuvre pour connaître le nombre de composants purs présents dans l'image considérée, comme décrit dans [1]. Puis l'algorithme du N-FINDR est appliqué à l'image

TAB. 1 – Erreur quadratique moyenne des vecteurs d’abondances.

	Modèle MML		Modèle NCM
	Estimation Bayésienne	FCLS	Estimation Bayésienne
EQM <sup>2</sup>	$8.91 \times 10^{-2}$	$6.10 \times 10^{-2}$	$5.51 \times 10^{-2}$

afin de déterminer les spectres de ces composants purs. Dans notre cas, les  $R = 3$  composants purs extraits correspondent à la végétation, l’eau et la terre utilisés comme vecteurs moyennes  $m_1$ ,  $m_2$  et  $m_3$ . Les cartes d’abondances estimés par l’algorithme proposé pour les  $R = 3$  matériaux purs sont donnés Fig. 2 (en haut à droite et en bas). Un pixel blanc (respectivement noir) sur la carte indique une grande (respectivement petite) valeur du coefficient d’abondance. Ainsi la zone du lac (représenté par des pixels blancs dans la carte d’abondance de l’eau et par des pixels noirs dans les autres) peut être clairement retrouvée.

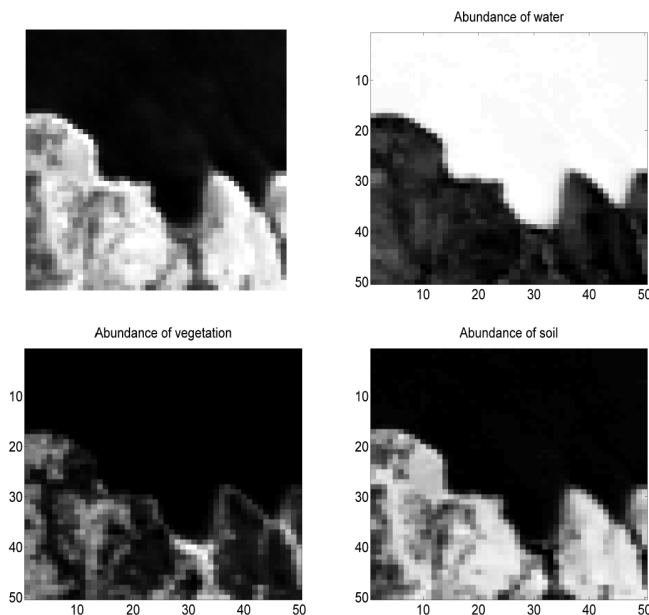


FIG. 2 – En haut à gauche : région d’intérêt (Moffett field) en niveaux de gris (longueur d’onde  $\lambda = 0.66 \mu m$ ). En haut à droite et en bas : cartes d’abondances estimés par l’algorithme proposé.

## 5 Conclusions

Cet article présente un algorithme Bayésien hiérarchique pour un modèle de mélange de spectres aléatoires proposé initialement dans [4]. L’algorithme permet d’estimer la loi *a posteriori* des coefficients de ce mélange et de la variance des spectres du mélange. Cette loi peut ensuite être utilisée pour estimer les paramètres inconnus du modèle et déterminer des mesures d’incertitude associées à ces estimations (par exemple des intervalles de confiance). Les simulations effectuées dans cet ar-

ticle montrent que le modèle proposé est plus robuste en l’absence de pixels purs dans l’image que le modèle de mélange linéaire classiquement utilisé. Des simulations obtenues à partir d’une image réelle étudiée dans [9] ont également donnés des résultats satisfaisants. Des études plus poussées dans l’avenir incluent une généralisation de la méthode proposée vers un algorithme semi-supervisé où le nombre de composants purs est inconnu, similaire à celui présenté dans [9].

## Références

- [1] N. Keshava and J. Mustard, “Spectral unmixing,” *IEEE Signal Processing Magazine*, pp. 44–56, Jan. 2002.
- [2] 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.
- [3] 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.
- [4] M. T. Eismann and D. Stein, “Stochastic mixture modeling,” in *Hyperspectral Data Exploitation : Theory and Applications*, C.-I. Chang, Ed. Wiley, 2007, ch. 5.
- [5] E. Punskeya, C. Andrieu, A. Doucet, and W. Fitzgerald, “Bayesian curve fitting using MCMC with applications to signal segmentation,” *IEEE Trans. Signal Processing*, vol. 50, no. 3, pp. 747–758, March 2002.
- [6] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, “Introducing Markov chain Monte Carlo,” in *Markov Chain Monte Carlo in Practice*, W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, Eds. London : Chapman & Hall, 1996, pp. 5–8.
- [7] RSI (Research Systems Inc.), *ENVI User’s guide Version 4.0*, Boulder, CO 80301 USA, Sept. 2003.
- [8] 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. 39, no. 3, pp. 529–545, March 2001.
- [9] 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.