

# Estimation de variabilité pour le démixage non-supervisé d'images hyperspectrales

Pierre-Antoine THOUVENIN, Nicolas DOBIGEON, Jean-Yves TOURNERET

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

Travail soutenu par la Direction Générale de l'Armement (DGA) et par le projet ANR Hypanema n°ANR-12-BS03-003.

prénom.nom@enseeiht.fr

**Résumé** – Le démixage d'images hyperspectrales vise à identifier les signatures spectrales d'un milieu imagé ainsi que leurs proportions dans chacun des pixels. Toutefois, les signatures extraites présentent en pratique une variabilité qui peut compromettre la fiabilité de cette identification. En supposant ces signatures potentiellement affectées par le phénomène de variabilité, nous proposons d'estimer les paramètres d'un modèle de mélange linéaire à l'aide d'un algorithme de minimisation alternée (Proximal alternating linearized minimization, PALM) dont la convergence a été démontrée pour une classe de problèmes non-convexes qui contient précisément le problème du démixage d'images hyperspectrales. La méthode proposée est évaluée sur des données synthétiques et réelles.

**Abstract** – Hyperspectral unmixing aims at identifying the reference spectral signatures contained in a hyperspectral image and their abundance fractions. In practice, the variability of the identified signatures can induce significant abundance estimation errors. This paper proposes to address this issue by interpreting the extracted endmembers as possibly corrupted versions of the true endmembers under a linear mixing model. The parameters of this model are then estimated using an optimization algorithm based on the proximal alternating linearized minimization (PALM), whose convergence has been established for the problem considered in this paper. The performance of the proposed unmixing strategy is evaluated on synthetic and real data, showing promising results.

## 1 Contexte, motivations et contribution

Le démixage d'images hyperspectrales consiste à identifier les signatures spectrales des composants élémentaires présents dans les données étudiées (*spectres purs*) ainsi que leurs proportions dans chacun des pixels (*abondances*) suivant un modèle de mélange prédéfini. Un modèle de mélange linéaire est généralement adapté à la description de données hyperspectrales en l'absence d'importantes différences de relief et d'interactions à l'échelle microscopique entre les éléments imagés [1]. Toutefois, les signatures spectrales des constituants élémentaires peuvent varier d'un pixel à l'autre en raison de la fluctuation des conditions d'acquisition ou de la variabilité naturelle de la scène imagée. Ce phénomène constitue une source non négligeable d'erreur lors de l'estimation des coefficients d'abondance, erreur qui sera propagée dans le cas de procédures non-supervisées de démixage. Différentes approches proposées dans la littérature abordent ce problème d'un point de vue statistique (spectres purs représentés par des vecteurs aléatoires) ou en faisant intervenir une bibliothèque spectrale (disponible *a priori* ou extraite des données étudiées) [2]. Dans cet article, nous introduisons un nouveau modèle de mélange qui permet de rendre compte explicitement de la variabilité spectrale des composants purs de l'image. Tandis que la variabilité n'est supposée résulter que de facteurs d'échelle spatialement variables dans [3], chaque signature est ici représentée par un spectre pur affecté par un terme additif spatialement variable.

La première contribution de cet article est la définition d'un modèle de démixage original prenant en compte la variabilité spectrale des spectres purs d'une image hyperspectrale. La

seconde contribution réside dans l'étude des performances de l'algorithme d'optimisation PALM [4] pour l'estimation des paramètres du modèle introduit. Notons que des garanties théoriques de convergence de cet algorithme existent pour le problème de démixage spectral considéré dans ce travail.

L'article est organisé comme suit. Le modèle proposé est introduit dans la section 2 tandis que la section 3 décrit la stratégie de démixage utilisée. Les résultats des simulations conduites sur données synthétiques et réelles sont rapportés respectivement dans les sections 4 et 5. Enfin, la section 6 vient conclure ce travail.

## 2 Modèle de mélange

### 2.1 Modèle de mélange linéaire perturbé (MMLP)

Chaque pixel  $\mathbf{y}_n$  d'une image hyperspectrale est représenté par une combinaison linéaire de  $K$  spectres purs – notés  $\mathbf{m}_k$  – affectés par un terme additif de variabilité  $\mathbf{d}\mathbf{m}_{n,k}$  dépendant du pixel considéré. Le modèle peut alors s'écrire sous la forme

$$\mathbf{y}_n = \sum_{k=1}^K a_{kn} (\mathbf{m}_k + \mathbf{d}\mathbf{m}_{n,k}) + \mathbf{b}_n \text{ pour } n = 1, \dots, N \quad (1)$$

où  $L$  désigne le nombre de bandes spectrales,  $N$  est le nombre de pixels,  $\mathbf{y}_n$  est le  $n^{\text{e}}$  pixel,  $\mathbf{m}_k$  représente le  $k^{\text{e}}$  spectre pur,  $a_{kn}$  est la proportion du  $k^{\text{e}}$  spectre pur dans le  $n^{\text{e}}$  pixel, et  $\mathbf{d}\mathbf{m}_{n,k}$  est le terme de variabilité du  $k^{\text{e}}$  spectre pur dans le  $n^{\text{e}}$  pixel. Le terme  $\mathbf{b}_n$  représente le bruit résultant de l'acquisition et de la modélisation des données. Sous forme matricielle, le

modèle de mélange linéaire perturbé (MMLP) (1) peut s'écrire

$$\mathbf{Y} = \mathbf{M}\mathbf{A} + \underbrace{\left[ \mathbf{dM}_1 \mathbf{a}_1 \mid \dots \mid \mathbf{dM}_N \mathbf{a}_N \right]}_{\Delta} + \mathbf{B} \quad (2)$$

où  $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_N]$  est une matrice de taille  $L \times N$  contenant les pixels de l'image,  $\mathbf{M}$  est la matrice de taille  $L \times K$  contenant les spectres purs,  $\mathbf{A}$  désigne la matrice de taille  $K \times N$  des coefficients d'abondance,  $\mathbf{dM}_n$  est la matrice de variabilité spectrale de taille  $L \times K$  et  $\mathbf{B}$  représente la matrice de bruit. Les contraintes de positivité et de somme-à-un, qui permettent de traduire des considérations d'ordre physique, sont exprimées comme suit [1]

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

où  $\succeq$  désigne une inégalité terme à terme. Pour contrôler l'énergie du terme de variabilité, nous introduisons la contrainte

$$\|\mathbf{dM}_n\|_F^2 \leq \sigma^2, \quad \text{pour } \sigma^2 > 0 \text{ fixé.} \quad (4)$$

## 2.2 Formulation du problème

L'objectif du démixage consiste ici à estimer les signatures spectrales des spectres purs  $\mathbf{M}$ , ainsi que les termes de variabilités  $\mathbf{dM}_n$  et les vecteurs d'abondance en chaque pixel de l'image. Le modèle (1) associé aux contraintes (3) et (4) conduit naturellement à formuler le démixage sous forme d'un problème d'optimisation. En prenant pour terme d'attache aux données la distance de Frobenius entre les données acquises et reconstruites, nous obtenons un problème non-convexe, de surcroît mal posé. Nous introduisons alors deux fonctions de pénalité  $\Phi$  et  $\Psi$  visant à exploiter toute connaissance *a priori* respectivement disponible sur les abondances et les spectres purs. De ce fait, nous considérons le problème

$$(\mathbf{M}^*, \mathbf{dM}^*, \mathbf{A}^*) \in \arg \min_{\mathbf{M}, \mathbf{dM}, \mathbf{A}} \left\{ \mathcal{J}(\mathbf{M}, \mathbf{dM}, \mathbf{A}) \text{ s.c. (3)} \right\} \quad (5)$$

avec

$$\begin{aligned} \mathcal{J}(\mathbf{M}, \mathbf{dM}, \mathbf{A}) &= \frac{1}{2} \|\mathbf{Y} - \mathbf{M}\mathbf{A} - \Delta\|_F^2 \\ &\quad + \alpha \Phi(\mathbf{A}) + \beta \Psi(\mathbf{M}) \end{aligned} \quad (6)$$

où les paramètres  $\alpha, \beta$  contrôlent les poids relatifs accordés aux pénalités  $\Phi(\mathbf{A})$  et  $\Psi(\mathbf{M})$ , et  $\|\cdot\|_F$  désigne la norme de Frobenius. Nous supposons par ailleurs que les pénalisations sont séparables comme suit

$$\Phi(\mathbf{A}) = \sum_{n=1}^N \phi(\mathbf{a}_n), \quad \Psi(\mathbf{M}) = \sum_{\ell=1}^L \psi(\tilde{\mathbf{m}}_\ell) \quad (7)$$

où  $\phi, \psi$  sont des fonctions différentiables à gradient lipschitzien, et  $\tilde{\mathbf{m}}_\ell$  désigne la  $\ell$ ème ligne de  $\mathbf{M}$ . Cette hypothèse permettra d'exploiter l'algorithme PALM [4] en décomposant par blocs les différents termes de la fonction d'objectif  $\mathcal{J}$ . Les pénalités choisies sont détaillées dans les paragraphes suivants<sup>1</sup>.

1. Pénalités et détails supplémentaires sont disponibles dans [5].

### 2.2.1 Pénalité pour les abondances

La pénalité  $\Phi$  a été choisie de sorte à promouvoir une variation spatiale suffisamment lisse des coefficients d'abondance [6]. Cette pénalité peut s'exprimer sous la forme matricielle

$$\Phi(\mathbf{A}) = \frac{1}{2} \|\mathbf{A}\mathbf{H}\|_F^2 \quad (8)$$

où  $\mathbf{H} \in \mathbb{R}^{N \times 4N}$  permet de calculer les différences entre les abondances d'un pixel donné et celles de ses 4 plus proches voisins.

### 2.2.2 Pénalité pour les spectres purs

La littérature mentionne différentes pénalités visant à contraindre le volume du simplexe formé par les données, simplexe dont les sommets sont les spectres purs. Le critère introduit dans [7] permet d'approcher le volume du  $(K-1)$ -simplexe. Ce critère est exprimé par

$$\Psi(\mathbf{M}) = \frac{1}{2} \sum_{i \neq j} \|\mathbf{m}_i - \mathbf{m}_j\|_2^2 = \frac{1}{2} \sum_{k=1}^K \|\mathbf{M}\mathbf{G}_k\|_F^2 \quad (9)$$

avec

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

pour  $k = 1, \dots, K$ , avec  $\mathbf{I}_K$  la matrice identité,  $\mathbf{1}_K$  le vecteur dont les entrées sont égales à 1 et  $\mathbf{e}_k$  le  $k$ ème vecteur de la base canonique de  $\mathbb{R}^K$ .

## 3 Algorithme de démixage

Le problème (5), globalement non-convexe et différentiable, présente une structure qui permet de mettre en œuvre une minimisation itérative à l'aide de l'algorithme développé dans [4] tout en bénéficiant des mêmes propriétés de convergence (cf. Algo. 1). Chacune des étapes de mise à jour se réduit à un pas de descente de gradient projeté. Ces étapes sont détaillées dans les paragraphes suivants.

### 3.1 Mise à jour des abondances

Par application directe de [4] et exploitation de la séparabilité de  $\mathcal{J}$  par rapport à  $\mathbf{a}_n$ , la mise à jour des abondances sous la contrainte (3) est donnée par

$$\mathbf{a}_n^{(k+1)} = \mathcal{P}_{\mathcal{S}_K} \left( \mathbf{a}_n^{(k)} - \frac{1}{\lambda_n^{(k)}} \nabla_{\mathbf{a}_n} \mathcal{J} \right) \quad (11)$$

$$\lambda_n^{(k)} = \gamma_1 L_{1,n}^{(k)} \quad (12)$$

pour  $n = 1, \dots, N$ , où  $\mathcal{S}_K$  désigne le simplexe unité de  $\mathbb{R}^K$ ,  $\mathcal{P}_{\mathcal{S}_K}$  est le projecteur sur  $\mathcal{S}_K$ ,  $L_{1,n}^{(k)}$  désigne la constante de Lipschitz de  $\nabla_{\mathbf{a}_n} \mathcal{J}$  et  $\gamma_1 > 1$ .

### 3.2 Mise à jour des spectres purs

Avec la pénalité choisie au paragraphe 2.2.2 et par application directe de [4], nous obtenons

$$\begin{aligned} \mathbf{M}^{(k+1)} &= (\mathcal{P}_{1,+} \circ \dots \circ \mathcal{P}_{N,+} \circ \mathcal{P}_+) \left( \mathbf{M}^{(k)} - \frac{1}{\mu^{(k)}} \nabla_{\mathbf{M}} \mathcal{J} \right) \\ \mu^{(k)} &= \gamma_2 L_2^{(k)} \end{aligned} \quad \begin{matrix} (13) \\ (14) \end{matrix}$$

où  $\mathcal{P}_{n,+}$  est le projecteur sur l'espace  $\{\mathbf{X} \mid \mathbf{X} + \mathbf{dM}_n \succeq \mathbf{O}_{L,K}\}$ ,  $\mathcal{P}_+$  le projecteur sur  $\{\mathbf{X} \mid \mathbf{X} \succeq \mathbf{O}_{L,K}\}$ ,  $L_2^{(k)}$  désigne la constante de Lipschitz de  $\nabla_{\mathbf{M}} \mathcal{J}$  et  $\gamma_2 > 1$ .

### 3.3 Mise à jour des termes de variabilité

De même, par application directe de [4],

$$\mathbf{dM}_n^{(k+1)} = (\mathcal{P}_{\sigma^2} \circ \mathcal{P}_+) \left( \mathbf{dM}_n^{(k)} - \frac{1}{\nu_n^{(k)}} \nabla_{\mathbf{dM}_n} \mathcal{J} \right) \quad (15)$$

$$\nu_n^{(k)} = \gamma_3 L_{2,n}^{(k)} \quad (16)$$

pour  $n = 1 \dots N$ , où  $L_{3,n}^{(k)}$  désigne la constante de Lipschitz de  $\nabla_{\mathbf{dM}_n} \mathcal{J}$ ,  $\gamma_3 > 1$ ,  $\mathcal{P}_{\sigma^2}$  est le projecteur sur  $\{\mathbf{X} \mid \|\mathbf{X}\|_F^2 \leq \sigma^2\}$  et  $\mathcal{P}_+$  le projecteur sur  $\{\mathbf{X} \mid \mathbf{X} \succeq \mathbf{O}_{L,K}\}$ .

## 4 Simulation sur données synthétiques

La méthode proposée a été évaluée en l'absence de pixels purs sur une image de taille  $128 \times 64$  obtenue par mélange linéaire de 3 spectres purs (413 bandes spectrales). Les pixels sont affectés par un bruit blanc additif gaussien pour un rapport signal-sur-bruit résultant de 30dB. Les signatures spectrales utilisées pour générer les mélange en chaque pixel ont été obtenues par le produit de spectres de référence avec des fonctions affines par morceaux générées aléatoirement. De telles fonctions ont été générées pour chacun des spectres purs et chacun des pixels afin d'obtenir un modèle réaliste de variabilité.

### 4.1 Méthodes comparées

La méthode proposée a été comparée aux algorithmes VCA [8] / FCLS [9] et AEB [10]. Les résultats de VCA/FCLS ont servi à initialiser les abondances et spectre purs de l'algorithme proposé, tandis que les entrées des matrices de variabilité ont été initialisées avec une valeur égale à la précision relative de MATLAB (format virgule flottante, constante *eps*). La convergence de l'algorithme a été considérée atteinte pour une variation de la fonction d'objectif inférieure à  $\varepsilon = 0.002$ . Enfin, les constantes algorithmiques ont été fixées aux valeurs suivantes :  $\gamma_1 = \gamma_2 = \gamma_3 = 1.1$ ,  $\sigma^2 = 0.12$ ,  $\alpha = 2.4$ ,  $\beta = 3.2 \times 10^{-3}$ .

Les performances de l'algorithme ont été évaluées en termes d'erreur de reconstruction (RE) en plus d'autres critères spécifiques aux quantités estimées. Pour les spectres purs, l'angle spectral moyen (aSAM) a été retenu, tandis que la qualité des abondances et des variabilités estimées est exprimée en terme d'erreur quadratique moyenne (GMSE), d'expression

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

$$\text{aSAM}(\mathbf{Y}) = \frac{1}{N} \sum_{n=1}^N \arccos \frac{\langle \mathbf{y}_n, \widehat{\mathbf{y}}_n \rangle}{\|\mathbf{y}_n\|_2 \|\widehat{\mathbf{y}}_n\|_2}, \text{RE} = \frac{\|\mathbf{Y} - \widehat{\mathbf{Y}}\|_F^2}{LN}$$

$$\text{GMSE}(\mathbf{dM}) = \sum_{n=1}^N \frac{\|\mathbf{dM}_n - \widehat{\mathbf{dM}}_n\|_F^2}{NLK}, \text{GMSE}(\mathbf{A}) = \frac{\|\mathbf{A} - \widehat{\mathbf{A}}\|_F^2}{KN}$$

où  $\widehat{\mathbf{Y}}$  désigne la matrice des pixels reconstruits grâce aux paramètres estimés  $\widehat{\mathbf{A}}, \widehat{\mathbf{M}}, \widehat{\mathbf{dM}}_n$ .

### 4.2 Résultats

Les performances des méthodes de démixage comparées sont fournies dans la table 1. La méthode proposée permet d'améliorer sensiblement la qualité d'estimation des spectres purs et

### Algorithme 1 : Algorithme de dé-mélange (MMLP).

---

**Données :**  $\mathbf{Y}, \mathbf{A}^{(0)}, \mathbf{M}^{(0)}, \mathbf{dM}^{(0)}, \gamma_1 > 1, \gamma_2 > 1, \gamma_3 > 1$   
**Début**

$k \leftarrow 0;$

**tant que** critère d'arrêt non satisfait,

// Mise à jour des abondances

**pour**  $n = 1$  à  $N$  faire

(a)  $\mathbf{a}_n^{(k+1)} \leftarrow \text{Update}(\mathbf{a}_n^{(k)});$  // cf. §3.1

// Mise à jour des endmembers

(b)  $\mathbf{M}^{(k+1)} \leftarrow \text{Update}(\mathbf{M}^{(k)});$  // cf. §3.2

// Mise à jour de la variabilité

**pour**  $n = 1$  à  $N$  faire

(c)  $\mathbf{dM}_n^{(k+1)} \leftarrow \text{Update}(\mathbf{dM}_n^{(k)});$  // cf. §3.3

$k \leftarrow k + 1;$

$\mathbf{A} \leftarrow \mathbf{A}^{(k)}, \mathbf{M} \leftarrow \mathbf{M}^{(k)}, \mathbf{dM} \leftarrow \mathbf{dM}^{(k)};$

---

**Résultats :**  $\mathbf{A}, \mathbf{M}, \mathbf{dM}$

---

TABLE 1 – Résultats de simulation sur données synthétiques (GMSE( $\mathbf{A}$ ) $\times 10^{-2}$ , GMSE( $\mathbf{dM}$ ) $\times 10^{-4}$ , RE  $\times 10^{-4}$ ).

	VCA/FCLS	AEB	Méthode proposée
aSAM( $\mathbf{M}$ ) <sup>(°)</sup>	5.0639	5.1104	<b>4.5442</b>
GMSE( $\mathbf{A}$ )	2.07	2.11	<b>1.52</b>
GMSE( $\mathbf{dM}$ )	/	/	<b>5.43</b>
RE	2.66	2.66	<b>0.52</b>
aSAM( $\mathbf{Y}$ ) <sup>(°)</sup>	2.0580	2.0605	<b>0.5114</b>
time (s)	<b>1</b>	33	1678

des abondances par rapport aux autres méthodes, et présente une erreur de reconstruction plus faible au prix d'un coût calculatoire plus important.

## 5 Expérience sur données réelles

La méthode proposée a enfin été appliquée à des données réelles acquises par AVIRIS en 1997 aux abords de Moffett Field en Californie. La scène considérée ( $50 \times 50$ ) est constituée d'un lac et de sa berge. 189 des 224 bandes spectrales disponibles à l'origine ont été exploitées après retrait des bandes d'absorption de l'eau. Les résultats du démixage sont présentés dans la table 2 et les figures 1 et 2 pour  $\varepsilon = 10^{-3}$ ,  $\alpha = 0$ ,  $\beta = 2.2 \times 10^{-4}$ ,  $\sigma^2 = 28.4$  et  $\gamma_i = 1.1, i \in \{1, 2, 3\}$ .

Notons que l'algorithme proposé permet de caractériser la variabilité spectrale des spectres purs au sein de l'image. En particulier, la variabilité capturée (visible sur les figures 1g à 1i) est plus significative au niveau des zones d'interface entre les différentes espèces présentes, zones pour lesquelles un mélange linéaire n'est probablement pas adapté. Les non-linéarités habituellement observées dans ces régions [11] sont ici assimilées à des termes de variabilité sans nécessiter de modèle particulier de non-linéarité. Enfin, la figure 2 permet d'apprécier le niveau de variabilité qui peut affecter certains spectres purs. Nous pouvons d'ailleurs noter que les pics de variabilité observables dans la figure 2 révèlent que certaines bandes spectrales, trop corrompues pour être exploitables, n'ont pas été retirées avant démixage.

TABLE 2 – Résultats de l’expérience sur données réelles.

	VCA/FCLS	AEB	Méthode proposée
RE ( $\times 10^{-4}$ )	2.70	6.24	<b><math>3.35 \times 10^{-4}</math></b>
aSAM(Y) ( $^\circ$ )	7.31	6.93	<b><math>1.73 \times 10^{-4}</math></b>
time (s)	<b>0.41</b>	10	159

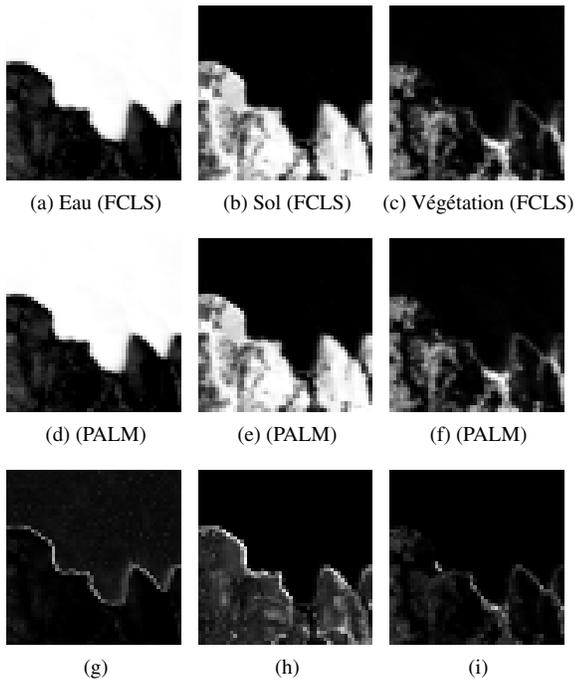


FIGURE 1 – Abondances (Figs. 1a à 1f) et répartition de la variabilité (Figs. 1g à 1i) [noir = 0, blanc = 1 pour les abondances, et  $\max_{k,n} (\|\mathbf{d}\mathbf{m}_{n,k}\|_2 / \sqrt{L}) = 0.08$  pour la variabilité].

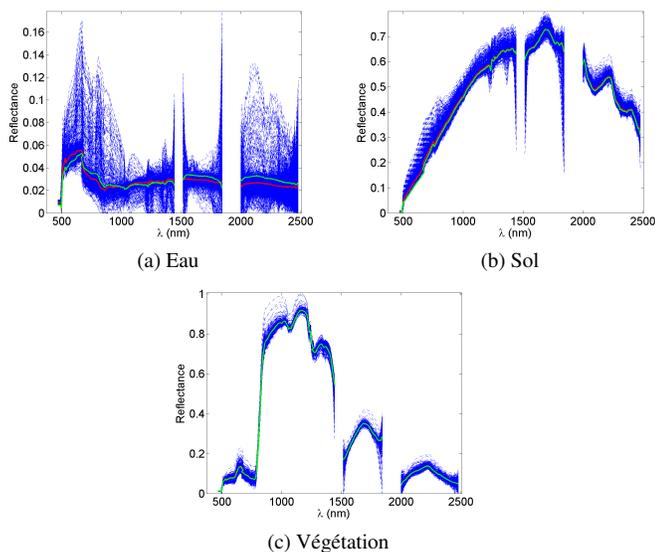


FIGURE 2 – Spectres purs estimés (méthode proposée en rouge, VCA en vert, variabilité en bleu).

## 6 Conclusion

Cet article a introduit un nouveau modèle de mélange de données hyperspectrales permettant de prendre en compte leur variabilité spectrale intrinsèque. Le démélange a été formulé comme un problème d’optimisation dont la résolution a été menée grâce à l’algorithme d’optimisation PALM. L’algorithme proposé, testé sur données réelles et synthétiques, présente une garantie théorique de convergence vers un point critique de la fonction objectif non-convexe [4], ce qui complète le travail effectué dans [5]. Nos prochains travaux consisteront à estimer la variabilité présente dans une série d’images hyperspectrales en vue d’améliorer leur démélange.

## Références

- [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] M. A. Veganzones, L. Drumetz, G. Tochon, M. D. Mura, A. Plaza, J. M. Bioucas-Dias, and J. Chanussot, “A new extended linear mixing model to adress spectral variability,” in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Process. : Evolution in Remote Sens. (WHISPERS)*, Lausanne, Switzerland, June 2014.
- [4] 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.
- [5] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, “Hyperspectral unmixing with spectral variability using a perturbed linear mixing model,” Feb. 2015, submitted. [Online]. Available : <http://arxiv.org/abs/1502.01260/>
- [6] J. Chen, C. Richard, and P. Honeine, “Nonlinear estimation of material abundance in hyperspectral images with  $\ell_1$ -norm spatial regularization,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2654–2665, May 2014.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, “Nonlinear unmixing of hyperspectral images using a generalized bilinear model,” *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.
- [12] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, “Proximal alternating and projection methods for nonconvex problems : An approach based on the Kurdyka-Łojasiewicz inequality,” *Mathematics of Operations Research*, vol. 35, no. 2, pp. 438–457, 2010.
- [13] J. Duchi, S. Shalev-Schwartz, Y. Singer, and T. Chandra, “Efficient projection onto the  $\ell_1$ -ball for learning in high dimensions,” in *Proc. Int. Conf. Machine Learning (ICML)*, Helsinki, Finland, 2008.