

Algorithme primal-dual de points intérieurs pour l'estimation pénalisée des cartes d'abondances en imagerie hyperspectrale

Emilie CHOUZENOUX¹, Saïd MOUSSAOUI², Jérôme IDIER²

¹LIGM, CNRS UMR 8049, 5, Boulevard Descartes, Champs sur Marne, 77454 Marne La Vallée Cedex 2, France
emilie.chouzenoux@univ-mlv.fr

²IRCCYN, CNRS UMR 6597, 1, rue de la Noë, BP 92101, F-44321 Nantes Cedex 03, France
prenom.nom@irccyn.ec-nantes.fr

Résumé – L'estimation des cartes d'abondances en imagerie hyperspectrale nécessite de résoudre un problème d'optimisation sous des contraintes de positivité et d'additivité (somme à un). Nous proposons un algorithme de points intérieurs primal-dual pour l'estimation contrainte de ces cartes, en nous plaçant dans le cadre où les signatures spectrales des composants présents au sein de l'image ont été préalablement estimées par un algorithme d'extraction des pôles de mélange. En comparaison avec la méthode de référence FCLS, l'algorithme proposé présente l'avantage d'un coût de calcul réduit, notamment dans un contexte d'images de grande taille, et de pouvoir traiter le cas d'un critère pénalisé favorisant la régularité spatiale des cartes d'abondances. Des exemples sur des données synthétiques illustrent les performances de cet algorithme.

Abstract – The estimation of abundance maps in hyperspectral imaging requires the resolution of an optimization problem under non-negativity and full-additivity (sum to one) constraints. Assuming that the spectral signatures of the image components have been previously determined by an endmember extraction algorithm, we propose in this paper a primal-dual interior point algorithm for the estimation of their fractional abundances. In comparison with the reference method FCLS, our algorithm has the advantage of a reduced computational cost, especially in the context of large scale images. Moreover, it allows to deal with a penalized criterion favorizing the spatial smoothness of abundance maps.

1 Introduction

L'imagerie hyperspectrale est une technique de mesure qui permet d'accéder à des informations liées à la composition d'une surface, en exploitant ses propriétés de réflexion (ou d'absorption) de la lumière dans plusieurs bandes de fréquences. Comme chaque pixel d'une image hyperspectrale correspond à une surface composée de plusieurs constituants, chacun ayant sa propre signature spectrale, l'analyse de toutes les images permet d'identifier ces constituants et de déterminer leurs proportions. Le modèle le plus couramment utilisé ([1, 2]) suppose que le spectre de réflectance dans chaque pixel de l'image est la combinaison linéaire d'un nombre fini de spectres purs des constituants, appelés aussi pôles de mélange, pondérés par des coefficients, appelés abondances, qui sont liés à la proportion de chaque constituant dans ce pixel.

Plus précisément, considérons N pixels d'une image hyperspectrale acquise dans K bandes spectrales. Selon le modèle de mélange linéaire, le spectre $\mathbf{y}_n = [y_{n,1}, \dots, y_{n,K}]^t$ du $n^{\text{ème}}$ pixel s'exprime comme la combinaison linéaire de P signatures spectrales \mathbf{s}_p entachée d'un bruit additif ϵ_n , qui sera supposé gaussien i.i.d, de moyenne nulle et de variance inconnue :

$$\mathbf{y}_n = \sum_{p=1}^P \mathbf{s}_p c_{n,p} + \epsilon_n, \quad (1)$$

où $\mathbf{s}_p = [s_{p,1}, \dots, s_{p,K}]^t$ est le spectre du $p^{\text{ème}}$ constituant,

$c_{n,p}$ est le coefficient d'abondance du $p^{\text{ème}}$ constituant dans le $n^{\text{ème}}$ pixel et P est le nombre de constituants recherchés. Si l'on considère tous les pixels de l'image hyperspectrale, le système d'équations (1) s'écrit sous forme matricielle

$$\mathbf{Y} = \mathbf{S}\mathbf{C} + \mathbf{E}, \quad (2)$$

avec $\mathbf{Y} \in \mathbb{R}^{K \times N}$ les observations associées aux pixels de l'image, $\mathbf{S} \in \mathbb{R}^{K \times P}$ les signatures spectrales, $\mathbf{C} \in \mathbb{R}^{P \times N}$ les coefficients d'abondances et $\mathbf{E} \in \mathbb{R}^{K \times N}$ les bruits associés aux observations. L'objectif du traitement est donc d'estimer \mathbf{S} et \mathbf{C} à partir de \mathbf{Y} , en utilisant le modèle (2).

Afin de résoudre ce problème, plusieurs stratégies peuvent être adoptées [3, 4]. L'article [5] peut être consulté pour une revue plus détaillée des méthodes existantes. Celles-ci se déclinent en deux grandes familles : les méthodes d'estimation séquentielle et les méthodes d'estimation conjointe. Dans la première approche, une procédure d'extraction des pôles du mélange est d'abord employée pour estimer les spectres des constituants des images avant d'appliquer un algorithme d'estimation des abondances. La deuxième approche se fonde sur une estimation conjointe des spectres purs et des abondances. Dans ce travail, nous nous plaçons dans le cadre de la première approche et nous nous intéressons au développement d'une méthode d'estimation rapide des cartes d'abondances.

Les vecteurs d'abondances $\mathbf{c}_n = [c_{n,1}, \dots, c_{n,P}]^t$ dans (2) étant reliés aux proportions des constituants de la surface, leur

estimation doit être réalisée sous les contraintes de positivité et d'additivité suivantes :

$$c_{n,p} \geq 0, \quad \forall p = 1, \dots, P, \quad (3a)$$

$$\sum_{p=1}^P c_{n,p} = 1. \quad (3b)$$

Les algorithmes usuels sont basés sur la minimisation du critère des moindres carrés sous la contrainte (3a) (NNLS [6], *non-negative least squares*) ou (3b) (SCLS [7], *sum-to-one constrained least squares*) ou encore sous les deux contraintes (3a) et (3b) (FCLS [8], pour *fully constrained least squares*). Cependant, ces méthodes sont caractérisées par un coût de calcul important, notamment dans le cas d'images de grande taille. De plus, celles-ci ne s'appliquent pas dans le cas d'un critère prenant en compte des corrélations inter-pixels.

Nous proposons dans cet article un algorithme rapide basé sur une stratégie de points intérieurs permettant de minimiser un critère strictement convexe quelconque en intégrant les contraintes de positivité et d'additivité. Par conséquent, cet algorithme peut s'appliquer au cas d'un critère de moindres carrés pénalisé par un terme convexe favorisant la régularité spatiale des cartes.

2 Approche proposée

Nous adoptons une approche itérative de type point intérieurs pour traiter le problème d'optimisation contraint

$$\min_{\mathbf{C}} F(\mathbf{C}) \quad \text{sous les contraintes (3a) – (3b),} \quad (4)$$

où F est une fonction strictement convexe de $\mathbb{R}^{P \times N}$.

2.1 Intégration de la contrainte d'égalité

Tout d'abord, le problème (4) est transformé, à l'aide d'un changement de variable, en un nouveau problème faisant apparaître des contraintes d'inégalité uniquement. Comme proposé dans [9], pour tout vecteur initial $\mathbf{c}_n^{(1)}$ vérifiant la contrainte (3b), le vecteur défini par $\mathbf{c}_n = \mathbf{c}_n^{(1)} + \mathbf{Z}\mathbf{u}_n$, avec $\mathbf{u}_n \in \mathbb{R}^{P-1}$, respecte aussi cette contrainte d'additivité si $\mathbf{Z} \in \mathbb{R}^{P \times P-1}$ est une matrice dont les colonnes forment l'espace nul de $\mathbf{1}_{1 \times P}$. Cette matrice étant donnée par

$$Z_{ij} = \begin{cases} 1 & \text{si } i = j, \\ -1 & \text{si } i = j + 1, \\ 0 & \text{sinon,} \end{cases}$$

le problème (4) est réécrit sous la forme

$$\min_{\mathbf{U} \in \mathbb{R}^{P-1 \times N}} F(\mathbf{C}^{(1)} + \mathbf{Z}\mathbf{U}) \quad (5)$$

sous les contraintes

$$\mathbf{Z}\mathbf{u}_n + \mathbf{c}_n^{(1)} \geq 0, \quad \forall n = 1, \dots, N, \quad (6)$$

ou, de façon équivalente,

$$\min_{\mathbf{u} \in \mathbb{R}^{N \times P-N}} L(\mathbf{u}), \quad \text{sous les contraintes } \mathbf{T}\mathbf{u} + \mathbf{t} \geq \mathbf{0}. \quad (7)$$

Le critère $L(\cdot)$ se déduit de $F(\cdot)$ par $L(\mathbf{u}) = F(\mathbf{C}^{(1)} + \mathbf{Z}\mathbf{U})$ en notant $\mathbf{u} = \text{vect}(\mathbf{U})$, où l'opérateur vect correspond à la transformation matrice-vecteur dans l'ordre lexicographique. De plus, \mathbf{T} et \mathbf{t} sont définis respectivement par $\mathbf{t} = \text{vect}(\mathbf{C}^{(1)})$ et $\mathbf{T} = \mathbf{I}_N \otimes \mathbf{Z}$ où \otimes est le produit de Kronecker et \mathbf{I}_N la matrice identité de taille $N \times N$.

2.2 Algorithme de points intérieurs primal-dual

Principe. Les conditions de Karush-Kuhn-Tucker (KKT) permettent de caractériser la solution \mathbf{u} de (7) et les multiplicateurs de Lagrange associés $\boldsymbol{\lambda}$:

$$\begin{cases} \nabla L(\mathbf{u}) - \mathbf{T}^t \boldsymbol{\lambda} = \mathbf{0} \\ \boldsymbol{\Lambda}(\mathbf{T}\mathbf{u} + \mathbf{t}) = \mathbf{0} \\ \mathbf{T}\mathbf{u} + \mathbf{t} \geq \mathbf{0} \\ \boldsymbol{\lambda} \geq \mathbf{0} \end{cases} \quad (8)$$

en notant $\boldsymbol{\Lambda} = \text{Diag}(\boldsymbol{\lambda})$. L'approche primale-duale consiste à estimer de façon jointe \mathbf{u} et $\boldsymbol{\lambda}$ en résolvant une séquence de problèmes, correspondant à des versions perturbées des conditions (8) paramétrées par une suite de paramètres positifs (μ_k) convergeant vers 0 :

$$\begin{cases} \nabla L(\mathbf{u}) - \mathbf{T}^t \boldsymbol{\lambda} = \mathbf{0} \\ \boldsymbol{\Lambda}(\mathbf{T}\mathbf{u} + \mathbf{t}) = \boldsymbol{\mu}_k \\ \mathbf{T}\mathbf{u} + \mathbf{t} \geq \mathbf{0} \\ \boldsymbol{\lambda} \geq \mathbf{0} \end{cases} \quad (9)$$

où $\boldsymbol{\mu}_k = \mu_k \mathbf{1}_{NP \times 1}$.

Algorithme. Une itération k de l'algorithme se décompose en deux étapes. Tout d'abord, $(\mathbf{u}_{k+1}, \boldsymbol{\lambda}_{k+1})$ est calculé en fonction de $(\mathbf{u}_k, \boldsymbol{\lambda}_k)$ en résolvant (9) de façon approchée. Puis le paramètre μ_{k+1} est déterminé selon une règle de mise à jour permettant de garantir la convergence de l'algorithme. Nous proposons d'utiliser la stratégie décrite dans [9] pour mettre en œuvre ces deux étapes de calcul.

La solution approchée de (9) est obtenue par une itération de Newton associée à une recherche de pas, selon le schéma :

$$(\mathbf{u}_{k+1}, \boldsymbol{\lambda}_{k+1}) = (\mathbf{u}_k + \alpha_k \mathbf{d}_k^u, \boldsymbol{\lambda}_k + \alpha_k \mathbf{d}_k^\lambda).$$

Les directions de Newton primales et duales $(\mathbf{d}_k^u, \mathbf{d}_k^\lambda)$ sont calculées en résolvant le système linéaire

$$\begin{pmatrix} \nabla^2 L(\mathbf{u}_k) & -\mathbf{T}^t \\ \boldsymbol{\Lambda}_k \mathbf{T} & \text{Diag}(\mathbf{T}\mathbf{u}_k + \mathbf{t}) \end{pmatrix} \begin{pmatrix} \mathbf{d}_k^u \\ \mathbf{d}_k^\lambda \end{pmatrix} = -r_{\mu_k}(\mathbf{u}_k, \boldsymbol{\lambda}_k), \quad (10)$$

où $r_{\mu_k}(\mathbf{u}, \boldsymbol{\lambda})$ est le résidu défini par les deux premières équations de (9) :

$$r_{\mu_k}(\mathbf{u}, \boldsymbol{\lambda}) = \begin{pmatrix} \nabla L(\mathbf{u}) - \mathbf{T}^t \boldsymbol{\lambda} \\ \boldsymbol{\Lambda}(\mathbf{T}\mathbf{u} + \mathbf{t}) - \boldsymbol{\mu}_k \end{pmatrix}.$$

Notons que le coût de calcul de l'algorithme est fortement dépendant du coût de la résolution de ce système.

Le pas α_k est déterminé par une stratégie d'Armijo appliquée à la fonction de mérite primale-duale $\phi_{\mu_k}(\mathbf{u}_k + \alpha \mathbf{d}_k^u, \boldsymbol{\lambda}_k +$

αd_k^λ), définie par :

$$\begin{aligned} \phi_\mu(\mathbf{u}, \boldsymbol{\lambda}) = & L(\mathbf{u}) - \mu \sum_{i=1}^{NP} \ln([\mathbf{T}\mathbf{u} + \mathbf{t}]_i) \\ & + \boldsymbol{\lambda}^t (\mathbf{T}\mathbf{u} + \mathbf{t}) - \mu \sum_{i=1}^{NP} \ln(\lambda_i [\mathbf{T}\mathbf{u} + \mathbf{t}]_i), \end{aligned}$$

et permettant notamment d'intégrer les deux contraintes d'inégalité de (9). Enfin, le paramètre μ_{k+1} est mis à jour selon la condition de μ -criticité [9] :

$$\mu_{k+1} = \frac{\delta_{k+1}}{NP} \min \left(\frac{1}{2}, \frac{\|r_0(\mathbf{u}_{k+1}, \boldsymbol{\lambda}_{k+1})\|}{2NP - N} \right),$$

où $\delta_{k+1} = (\mathbf{T}\mathbf{u}_{k+1} + \mathbf{t})^t \boldsymbol{\lambda}_{k+1}$ est le saut de dualité.

3 Résultats expérimentaux

Nous analysons tout d'abord les performances de l'algorithme en terme de temps de calcul pour la minimisation du critère des moindres carrés sous contraintes d'additivité et de positivité, puis nous discutons de l'influence de l'intégration d'un terme de régularité spatiale dans le critère.

3.1 Moindres carrés contraints

Dans le cas d'un critère des moindres carrés non pénalisé

$$F(\mathbf{C}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{S}\mathbf{C}\|_F^2, \quad (11)$$

où $\|\cdot\|_F$ représente la norme de Frobenius, l'algorithme FCLS permet d'intégrer efficacement les contraintes d'additivité et de positivité. L'objectif de cette section est de comparer l'algorithme primal-dual à cette méthode de référence, en terme de temps de calcul. Des cubes hyperspectraux synthétiques d'images de taille N obtenues dans $K = \sqrt{N}$ bandes de fréquences sont simulés. Des cartes d'abondances synthétiques sont générées à partir d'une loi de Dirichlet de paramètre unitaire pour chaque variable. Le nombre de constituants est pris égal à $P = 3, 5$ ou 10 et un bruit gaussien i.i.d de moyenne nulle et de variance adéquate est ajouté pour obtenir un rapport signal sur bruit de 15dB au niveau de chaque pixel de l'image.

Pour l'ensemble des tests effectués, les solutions fournies par les deux algorithmes sont similaires en terme de valeur de résidu et de qualité de reconstruction. Cependant, comme illustré par la figure 1, le temps de calcul nécessaire à l'optimisation du critère par l'algorithme primal-dual est très réduit. A titre d'exemple, un gain de temps d'un facteur de l'ordre de 4 est obtenu pour une image de taille 256×256 dans le cas de $P = 10$ composantes. Pour la même taille d'image, ce facteur est de l'ordre de 7 pour $P = 5$ et 12 pour $P = 3$.

3.2 Intégration d'une pénalisation spatiale

Nous proposons de considérer la minimisation d'un critère des moindres carrés pénalisé

$$F(\mathbf{C}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{S}\mathbf{C}\|_F^2 + \eta R(\mathbf{C}), \quad (12)$$

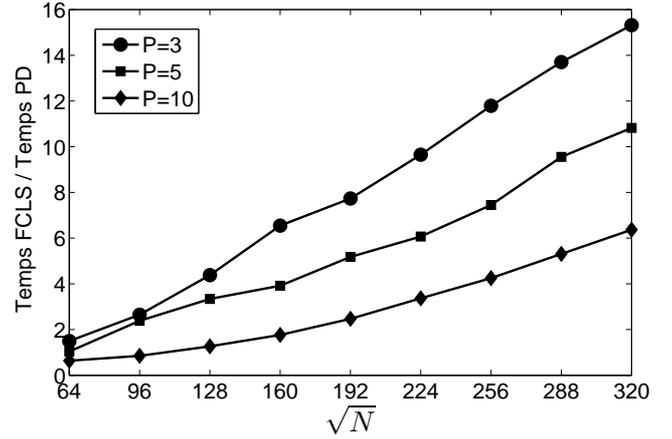


FIGURE 1 – Rapport entre les temps de calcul des algorithmes FCLS et primal-dual (PD) en fonction du nombre de pixels et du nombre de composantes.

sous les contraintes (3a) et (3b), de manière à renforcer la régularité de la répartition spatiale des coefficients d'abondances : . Le terme de régularisation est choisi de la forme

$$R(\mathbf{C}) = \sum_{p=1}^P \left(\psi(\Delta^V \mathbf{c}_p) + \psi(\Delta^H \mathbf{c}_p) \right),$$

avec Δ^V et Δ^H respectivement les opérateurs de gradient spatial vertical et horizontal, et ψ une fonction de pondération quadratique favorisant l'apparition de zones lisses dans les cartes.

Nous considérons un exemple synthétique d'un cube hyperspectral d'images de taille $N = 256 \times 256$ pixels. Les cartes simulées contiennent une dizaine d'atomes gaussiens générés aléatoirement et normalisés pour satisfaire la contrainte (3b). Les $P = 5$ spectres de taille $K = 224$ sont choisis arbitrairement à partir de la bibliothèque AVIRIS (figure 2).

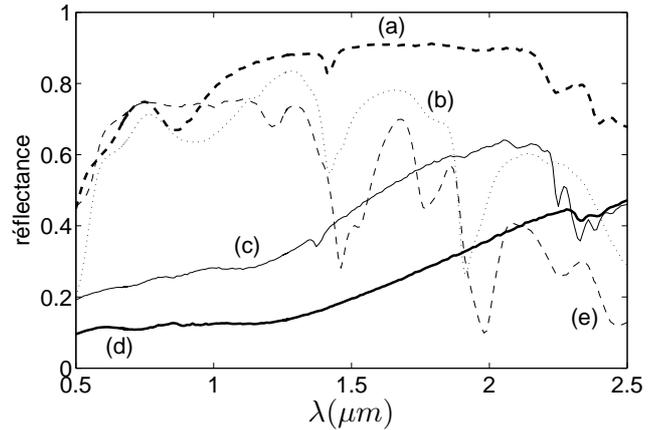


FIGURE 2 – Signatures spectrales sélectionnées dans la bibliothèque AVIRIS. (a) Andradite. (b) Erionite. (c) Chlorite. (d) Biotite. (e) Carnallite.

Les tableaux 1 et 2 présentent les résultats obtenus en terme de qualité d'estimation des cartes d'abondance et de temps de

calcul, par l'algorithme primal dual non pénalisé (PD) ou pénalisé (PD-P). Nous distinguons le cas où les pôles du mélange sont obtenus à l'aide de l'algorithme NFINDR [10] et le cas où ces pôles sont parfaitement connus (LIB). La qualité des cartes d'abondances estimées \hat{c}_p est évaluée au sens de l'erreur quadratique moyenne normalisée

$$\text{EQMn} = \frac{1}{P} \sum_{p=1}^P (\|c_p - \hat{c}_p\|_2^2 / \|c_p\|_2^2), \quad (13)$$

par rapport aux cartes de référence c_p . Le paramètre de régularisation η est choisi égal à 100 afin de minimiser l'erreur de reconstruction. Nous pouvons constater que le rajout de la régularisation permet de réduire l'erreur d'estimation des cartes, quels que soient les spectres retenus pour la reconstruction. Cette amélioration est d'autant plus notable que le niveau de bruit est élevé. Cependant, la minimisation du critère pénalisé nécessite un temps de calcul beaucoup plus élevé. Cette augmentation est liée à l'inversion du système (10) qui, dans le cas pénalisé, ne possède pas une structure permettant de mettre en œuvre une stratégie d'inversion rapide.

RSB (dB)	NFINDR-PD	NFINDR-PD-P
20	$1,3 \cdot 10^{-1} / 2,7 \text{ s}$	$1,2 \cdot 10^{-1} / 160 \text{ s}$
15	$1,4 \cdot 10^{-1} / 2,5 \text{ s}$	$1,2 \cdot 10^{-1} / 167 \text{ s}$
10	$1,9 \cdot 10^{-1} / 3 \text{ s}$	$1,5 \cdot 10^{-1} / 181 \text{ s}$
5	$2,4 \cdot 10^{-1} / 3,4 \text{ s}$	$1,9 \cdot 10^{-1} / 180 \text{ s}$

TABLE 1 – Résultats d'estimation et temps de calcul (EQMn/Temps) obtenus dans le cas de spectres estimés par NFINDR.

RSB (dB)	LIB-PD	LIB-PD-P
20	$2,5 \cdot 10^{-2} / 2,8 \text{ s}$	$2,5 \cdot 10^{-2} / 160 \text{ s}$
15	$7,1 \cdot 10^{-2} / 2,8 \text{ s}$	$2,5 \cdot 10^{-2} / 170 \text{ s}$
10	$1,7 \cdot 10^{-1} / 3 \text{ s}$	$2,4 \cdot 10^{-2} / 191 \text{ s}$
5	$3,6 \cdot 10^{-1} / 3,5 \text{ s}$	$2,5 \cdot 10^{-2} / 231 \text{ s}$

TABLE 2 – Résultats d'estimation et temps de calcul (EQMn/Temps) obtenus dans le cas de spectres issus de la bibliothèque.

4 Conclusion

Nous avons proposé un algorithme rapide pour l'estimation des cartes d'abondances en imagerie hyperspectrale sous contraintes de positivité et additivité. Cette approche pourrait s'appliquer au cas d'un bruit non gaussien ou de contraintes de somme inférieure ou égale à un. En outre, l'utilisation d'une méthode de type points intérieurs permet de traiter simultanément tous les pixels de l'image et par conséquent d'intégrer simplement une information spatiale sur les coefficients. Les prochaines étapes de ce travail concernent l'accélération de la

minimisation dans le cas pénalisé, à travers par exemple la proposition d'une stratégie d'inversion approchée du système (10) ou le remplacement de l'algorithme de Newton par une méthode de premier ordre.

Références

- [1] C.-I. Chang, *Hyperspectral Data Exploitation*, Wiley Interscience, 2007.
- [2] N. Keshava et J. Mustard, « Spectral unmixing », *IEEE Signal Processing Mag.*, vol. 19, n° 1, pp. 44–57, jan. 2002.
- [3] N. Keshava, « A survey of spectral unmixing », *Lincoln Lab. J.*, vol. 14, n° 1, pp. 55–78, 2003.
- [4] N. Dobigeon, S. Moussaoui, M. Coulon, J. Tourneret et A. O. Hero, « Algorithmes bayésiens pour le démixage supervisé, semi-supervisé et non-supervisé images hyperspectrales », *Trait. du Signal*, vol. 27, n° 1, pp. 79–108, août 2010.
- [5] M. Parente et A. Plaza, « Survey of geometric and statistical unmixing algorithms for hyperspectral images », in *Proceedings of the 2nd Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, juin 2010, pp. 1–4.
- [6] C. Lawson et R. Hanson, *Solving Least-Squares Problems*, Prentice-Hall, Englewood Cliffs, New Jersey, 1974.
- [7] J. J. Settle et N. A. Drake, « Linear mixing and the estimation of ground cover proportions », *Int. J. Remote Sens.*, vol. 14, n° 6, pp. 1159–1177, 1993.
- [8] D. Heinz et C.-I. Chang, « Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery », *IEEE Trans. Geosci. Remote Sensing*, vol. 39, n° 3, pp. 529–545, mars 2001.
- [9] P. Armand, J. C. Gilbert et S. Jan-Jégou, « A feasible BFGS interior point algorithm for solving strongly convex minimization problems », *SIAM J. Optimization*, vol. 11, pp. 199–222, 2000.
- [10] M. E. Winter, « N-FINDR : an algorithm for fast autonomous spectral endmember determination in hyperspectral data », in *proc. of SPIE - Image Spectrometry V*, 1999, vol. 3753, pp. 266–277.