

Méthode de démixage et dématricage conjoints fondée sur la complétion de rang un pour les images multispectrales « snapshot »

Kinan ABBAS, Matthieu PUIGT, Gilles DELMAIRE, Gilles ROUSSEL

Univ. Littoral Côte d'Opale, LISIC – UR 4491, F-62219 Longuenesse, France *

{kinan.abbas,matthieu.puigt,gilles.delmaire,gilles.rousseau}@univ-littoral.fr

Résumé – L'imagerie spectrale « snapshot » (SSI) est une technologie qui permet aux caméras miniaturisées d'acquérir le contenu spatio-spectral de scènes dynamiques à partir d'une seule exposition. Les systèmes SSI accomplissent cela en associant chaque pixel à une bande spectrale spécifique, introduisant en conséquence un compromis crucial entre la résolution spatiale et spectrale. Dans cet article, nous concevons une méthode conjointe pour effectuer le dématricage et le démixage d'images hyperspectrales acquises par la caméra SSI. Nous formulons le problème comme un problème de factorisation et de complétion matricielle de faible rang. A cette fin, nous proposons une méthode dédiée qui suppose que les données observées sont localement de rang 1 dans certains patchs SSI à trouver. Les expériences montrent que l'approche proposée offre des performances de dématricage et de démixage respectivement légèrement et bien supérieures aux méthodes de la littérature.

Abstract – Snapshot Spectral Imaging (SSI) is a technology that allows miniaturized cameras to acquire the spatio-spectral content of dynamic scenes from a single exposure. SSI systems accomplish this by associating each pixel with a specific spectral band, accordingly introducing a crucial trade-off between spatial and spectral resolution. In this paper, we design a joint method for performing “demaicing” and “unmixing” for the hyperspectral images acquired by the SSI camera. We formulate the problem as a low-rank matrix factorization and completion problem. To that end, we propose a dedicated method which assumes the observed data to be locally rank-1 in some SSI “patches” to find. The experiments show our proposed approach provides a slightly better demaicing performance than state-of-the-art methods and a much higher unmixing enhancement.

1 Introduction

La plupart des architectures matérielles d'acquisition d'images multi- et hyperspectrales font appel à un scan – qui peut être par exemple spatial ou spectral – pour générer un cube de données [1]. Pour répondre à ce problème des caméras de type « snapshot » (SSI pour *Snapshot Spectral Imaging* en anglais) permettent en une prise l'acquisition d'un cube de données, comme un appareil photo RVB classique [2]. Cependant, la miniaturisation de ces caméras entraîne un compromis entre résolution spectrale et spatiale qui peut être levé par des approches de dématricage [3]. Pour cela, de nombreuses méthodes ont été proposées, utilisant la corrélation spatiale et/ou spectrale et basées sur l'interpolation (WB) [4], l'arborescence binaire (BTES) [5], la différence spectrale (ItSD) [6], une image pseudo-panchromatique (PPID) [7], la complétion de faible rang (GRMR) [3] ou l'apprentissage profond [8].

Par ailleurs, le démixage hyperspectral décompose la signature spectrale observée sur chaque pixel spatial du cube en une combinaison des signatures des matériaux (endmembers), pondérée par un poids proportionnel à leur surface sur le pixel (abondance). De nombreux algorithmes ont été proposés, dont des méthodes fondées sur l'hypothèse du « pixel pur » [9], la

factorisation matricielle non-négative (FMN) [9], ou sur l'hypothèse de parcimonie des abondances (SCA) [10].

Selon [3], classifier des données après le dématricage d'une image SSI n'est pas performant. Or, il a été montré pour une application d'étalonnage *in situ* de capteurs [11] qu'une approche couplée réalisant une factorisation et une complétion de matrices était plus performant qu'une approche en deux étapes consistant d'abord en une complétion de matrice de faible rang suivie d'une factorisation matricielle. Dans cet article, nous étudions aussi cet effet appliqué aux données SSI dématricées et nous proposons une nouvelle méthode conjointe de démixage et dématricage. La structure de l'article est organisée ainsi : nous introduisons le problème à traiter dans la section 2. Nous décrivons ensuite la méthode que nous proposons dans la section 3 et nous étudions ses performances dans la section 4. Enfin, nous concluons en section 5.

2 Position du problème

Nous présentons d'abord le système d'acquisition SSI et nous formalisons le problème à résoudre. La technologie « snapshot » fait appel à une mosaïque de filtres de Fabry-Perot [2] qui bloquent la lumière en dehors d'une bande spectrale limitée propre à chaque filtre. Ainsi, dans un cas *idéal*, chaque pixel de la caméra ne capture qu'une unique bande étroite de longueurs d'ondes, et estimer le cube complet de données à partir de ces

*K. Abbas remercie la région Hauts-de-France qui finance en partie ses travaux. Les expériences présentées dans cet article ont été réalisées sur la plateforme de calcul scientifique CALCULCO, gérée par le SCoSI de l'ULCO.

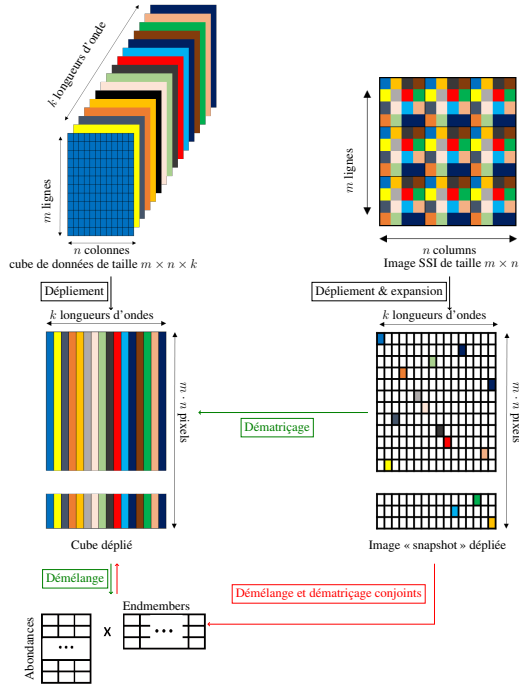


FIGURE 1 – Principes du démélange post-dématriçage (en vert) et du dématriçage et démélange conjoints (en rouge).

quelques données peut être revu comme un problème de complétion de matrice de faible rang [3] (voir fig. 1).

D’un point de vue formel, une caméra SSI acquiert une image en 2 dimension de $m \times n$ pixels pour chaque acquisition, où m et n sont respectivement le nombre de pixels selon l’axe horizontal et vertical. Chacun de ces pixels est associé à une longueur d’onde parmi les k acquises, ce qui signifie qu’une image SSI peut être vue comme une projection 2D d’une image 3D théorique de taille $m \times n \times k$ que nous allons maintenant modéliser.

En utilisant la stratégie de dépliement classiquement utilisée en démélange, nous pouvons ré-écrire le cube de données hyperspectrales 3D en une matrice notée Y de taille $(m \cdot n) \times k$, dont le premier indice correspond à un pixel spatial du cube initial et dont le second indice est associé à une longueur d’onde. De la même manière, il est possible [3] de ré-écrire l’image SSI sous la forme d’une matrice X de même taille que Y mais à données manquantes. X et Y sont alors liées par la relation :

$$W \circ X = W \circ Y, \quad (1)$$

où W représente une matrice binaire de poids dont les entrées non-nulles permettent de sélectionner quelle longueur d’onde est observée par la caméra et où le symbole \circ correspond au produit de Hadamard. La matrice W étant connue, l’image SSI peut être transformée sous la forme de la matrice X sans aucune ambiguïté. Cela implique notamment que le dématriçage consiste à estimer Y à partir des données observées dans X , ce qui peut être réalisé en supposant que Y est de faible rang [3].

Nous supposons de plus que chaque ligne de Y peut être

exprimée comme un mélange linéaire des endmembers, c.-à-d.

$$Y \approx G \cdot F, \quad (2)$$

où les matrices F et G sont respectivement de taille $p \times k$ et $(m \cdot n) \times p$ et représentent les matrices d’endmembers et d’abondances. La relation (2) non seulement fournit un modèle très classique de démélange hyperspectral [9] mais aussi un modèle d’approximation de faible rang, en supposant que $p < \min\{(m \cdot n), k\}$. En combinant les équations (1) and (2), on obtient alors le modèle de dématriçage et démélange conjoint :

$$W \circ X \approx W \circ (G \cdot F). \quad (3)$$

En effet, si on peut complètement estimer G et F à partir de la matrice partiellement connue X , le produit $G \cdot F$ apporte une approximation de Y qui permet de compléter X en calculant

$$\hat{Y} = W \circ X + (\mathbb{1}_{(m \cdot n) \times k} - W) \circ (G \cdot F), \quad (4)$$

où $\mathbb{1}_{(m \cdot n) \times k}$ représente la matrice de taille $(m \cdot n) \times k$ dont toutes les valeurs sont un.

Rappelons que le problème (3) a été défini dans le cas idéal où les filtres de Fabry-Perot ne conservent que la longueur d’onde d’intérêt dans chaque pixel du capteur SSI. Cependant, en pratique, ces filtres laissent passer des harmoniques. Plus précisément, chaque pixel SSI représente une version filtrée de l’ensemble du spectre qui est observée au niveau de ce pixel. Heureusement, ces filtres sont connus et fournis par les fabricants de telles caméras. Ce comportement a aussi été modélisé dans [3] et nous ré-utilisons une idée similaire dans cet article. Plus précisément, plutôt que de ne considérer qu’une seule valeur par ligne de X , nous supposons que la valeur mesurée est étalée sur l’ensemble des k longueurs d’ondes de la ligne de X associée. Nous supposons de plus que chacune de ces k valeurs est associée à un « indice de contribution dans le pixel » qui correspond à la valeur normalisée du filtre de Fabry-Perot pour la longueur d’onde correspondante, dans le pixel SSI considéré. Cela implique que les lignes de W ne sont plus binaires mais contiennent des valeurs entre 0 et 1 dont la somme vaut 1. En utilisant cette modélisation, le problème conjoint de dématriçage et démélange dans l’équation (3) reste inchangé, mis à part dans la définition de W .

A cette étape, rappelons que le contenu de G et de F peut être utile dans de nombreuses applications, comme l’apprentissage d’une bibliothèque spectrale (à partir de F) ou l’estimation de l’usage des sols (via G). Dans cet article, nous nous intéressons donc notamment à la comparaison entre une approche conjointe et une approche en 2 étapes, où Y serait dans un premier temps estimée par une méthode de dématriçage, avant d’appliquer sur cette matrice complète une approche de démélange. Ces deux stratégies sont fournies sur la figure 1, en rouge et vert, respectivement. Plus particulièrement, en complément d’une approche naïve de FMN Pondérée (FMNP) [12] faisant appel à une stratégie espérance-maximisation (EM) et des mises à jour par gradient de Nesterov [13] dans l’étape M, nous proposons une approche spécifiquement dédiée aux images SSI dans la section suivante.

3 Méthode proposée

Nous proposons maintenant une méthode spécifiquement conçue pour le traitement des données SSI. Cette approche trouve ses racines dans la manière la plus simple de restaurer un cube de données. En effet, le capteur est divisé en patches de mosaïques de filtres qui sont répétés sur la surface du capteur. Chaque patch est de taille $\sqrt{k} \times \sqrt{k}$ (typiquement, $k = 16$ ou 25 , de sorte que la taille d'un patch est 4×4 ou 5×5). On peut alors supposer que chaque patch correspond à un « super-pixel », c.-à-d. que chaque patch est lié à un unique endmember. En pratique, une telle hypothèse est bien évidemment fautive, d'où le fait que de nombreuses méthodes de dématricage aient été proposées. Cependant, elle peut être valide pour quelques patches à trouver, où un endmember domine les autres, notamment en observation en champ proche. Cela peut être vu comme une extension de l'hypothèse de pixel pur [9] ou un cas particulier de parcimonie des abondances utilisées en SCA [10]. Cependant, la principale différence par rapport aux configurations de démixage classique réside dans le fait que nous n'observons que partiellement le cube de données sur le patch, et que nous cherchons à estimer les données manquantes.

Notre méthode se présente ainsi. Nous notons X_i , Y_i et W_i les sous-matrices de taille $k \times k$ extraites de X , Y et W correspondant au patch i , respectivement. Nous estimons alors l'approximation de rang 1 de Y_i à partir de X_i par une approche de FMNP basée sur le gradient de Nesterov [12], c.-à-d.

$$W_i \circ X_i \approx W_i \circ (\underline{g}_i \cdot \mathbf{f}_i), \quad (5)$$

où \underline{g}_i est un vecteur colonne et \mathbf{f}_i un vecteur ligne. Si Y_i est approximativement de rang 1, alors

$$\|W_i \circ X_i - W_i \circ (\underline{g}_i \cdot \mathbf{f}_i)\|_F^2 \approx 0, \quad (6)$$

tandis que, dans le cas contraire,

$$\|W_i \circ X_i - W_i \circ (\underline{g}_i \cdot \mathbf{f}_i)\|_F^2 \gg 0. \quad (7)$$

Une telle erreur peut donc être utilisée comme « mesure de confiance mono-source » similaire à celles utilisées par les approches de SCA. Nous pouvons alors déduire de chaque patch une estimation plus ou moins proche d'un « vrai » endmember ainsi qu'une mesure de confiance mono-source. En ne conservant que les estimées liées aux erreurs les plus faibles, nous obtenons un nuage de spectres organisé en clusters distribués autour des spectres théoriques [14]. Nous pouvons alors utiliser une approche de clustering, par exemple K-means, pour estimer les endmembers finaux et les stocker dans F .

Ensuite, la matrice d'abondances G est ré-estimée à partir de X et F , sous forme de problème de moindres carrés non-négatifs – en utilisant les mises à jours de FMN utilisées dans [12] – et l'estimation finale \hat{Y} du cube de données est obtenue à partir de l'équation (4).

4 Validation expérimentale

La validation expérimentale de la méthode proposée est réalisée par des simulations d'images SSI obtenues à partir d'une

image hyperspectrale synthétique [15] qui consiste en une image de 120×120 pixels et $p = 5$ endmembers. Nous considérons des patches de mosaïques de filtres spectraux de taille 4×4 et 5×5 , à la fois dans le cas de filtres idéaux et de Fabry-Perot réels. Nous sélectionnons aléatoirement 16 ou 25 bandes spectrales de celles disponibles dans chaque image puis nous ajoutons un bruit Gaussien avec divers niveaux de rapports signal-à-bruit (SNR, en dB). Nous répétons 10 fois chaque expérience avec une initialisation différente.

Nous comparons les performances de la méthode de FMNP naïve avec l'approche proposée et cinq approches en deux étapes consistant à d'abord dématricer les images SSI – en utilisant GMRM [3], BTES [5], WB [4], PPID [7], and ItSD [6] – puis en démixant le cube restauré en utilisant la méthode de FMN utilisant le gradient de Nesterov [13].

Nous nous intéressons aux performances de dématricage – qui consistent à comparer l'estimée de Y par rapport à la matrice théorique – que nous mesurons en utilisant le rapport signal-à-bruit de crête (PSNR, en dB) alors que la qualité de démixage – calculée pour chaque ligne de F – est mesurée via le rapport signal-à-interférence (SIR) moyen (en dB). Les résultats obtenus sont présentés dans les tracés de la figure 2 qui montrent respectivement les performances de démixage et de dématricage, en fonction du SNR d'entrée. Il est notable que les performances de démixage atteintes par notre approche proposée est toujours très significativement meilleure que celle des techniques en 2 étapes et de l'approche naïve. Cela est probablement dû au fait que toutes les méthodes de dématricage testées peuvent introduire des erreurs d'estimations dans Y qui ont un impact négatif sur les performances de démixage et sur le fait que la convergence vers un optimum global de l'approche naïve n'est pas garantie. Malgré tout, les performances de dématricage fournies par l'approche proposée sont similaires aux meilleures approches de dématricage (PPID et ItSD) et même supérieures pour les SNR les plus élevés. De manière intéressante, alors que notre approche proposée fournit de bien meilleurs SIR que l'approche naïve, cette dernière fournit de meilleurs PSNR lorsque le niveau de bruit augmente. Cela signifie qu'il existe encore des marges de manœuvre pour améliorer le dématricage avec l'approche proposée.

5 Conclusion et discussion

Dans cet article, nous avons proposé de revisiter le problème de dématricage sous forme de factorisation matricielle pondérée. Cette formalisation nous permet aussi de voir ce problème sous forme de démixage. Outre une approche naïve, nous proposons une méthode spécifique qui est fondée sur une approximation de rang 1 dans des patches du capteur multispectral. L'approche proposée permet d'atteindre de bien meilleures performances de démixage que les approches en deux étapes – consistant en un dématricage suivi d'un démixage – et l'approche naïve. De plus, ses performances de dématricage sont similaires ou meilleures que celles obtenues par les approches

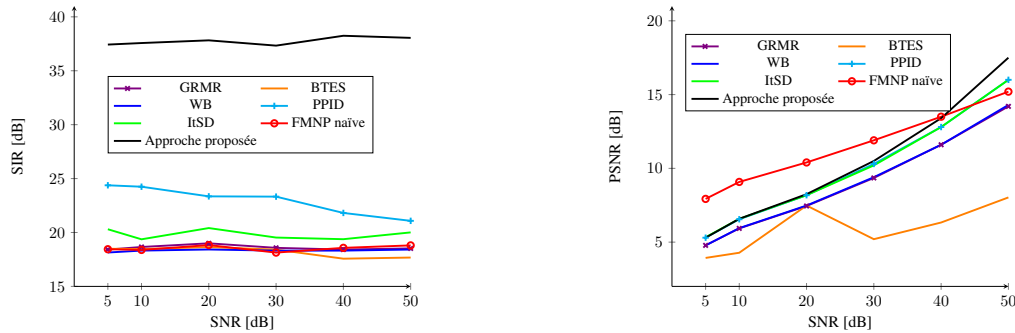


FIGURE 2 – SIR (à gauche) et PSNR (à droite) moyens – obtenus pour les filtres réels et idéaux dans des patches de taille 4×4 et 5×5 – par rapport au SNR d’entrée.

de la littérature et sont seulement dépassées par l’approche naïve lorsque les SNR décroient. Dans nos futurs travaux, nous souhaitons valider les performances obtenues sur de vraies images SSI. Nous souhaitons aussi développer de nouvelles extensions robustes à la variabilité spectrale [16].

Références

- [1] D. Manolakis, R. Lockwood, and T. Cooley, *Hyperspectral Imaging Remote Sensing*. Cambridge University Press, 2016.
- [2] B. Geelen, C. Blanch, P. Gonzalez, N. Tack, and A. Lambrechts, “A tiny VIS-NIR snapshot multispectral camera,” in *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics VIII*, vol. 9374, p. 937414, International Society for Optics and Photonics, 2015.
- [3] G. Tsagkatakis, M. Bloemen, B. Geelen, M. Jayapala, and P. Tsakalides, “Graph and rank regularized matrix recovery for snapshot spectral image demosaicing,” *IEEE Trans. Comput. Imaging*, vol. 5, pp. 301–316, June 2019.
- [4] J. Brauers and T. Aach, “A color filter array based multispectral camera,” in *12. Workshop Farbbildverarbeitung* (G. C. Group, ed.), (Ilmenau), October 5-6 2006.
- [5] L. Miao, H. Qi, R. Ramanath, and W. Snyder, “Binary tree-based generic demosaicking algorithm for multispectral filter arrays,” *IEEE Trans. Image Process.*, vol. 15, pp. 3550–3558, Nov. 2006.
- [6] J. Mizutani, S. S. Ogawa, K. Shinoda, M. Hasegawa, and S. Kato, “Multispectral demosaicking algorithm based on inter-channel correlation,” in *Proc. IEEE VCIP’14*, pp. 474–477, 2014.
- [7] S. Mihoubi, O. Losson, B. Mathon, and L. Macaire, “Multispectral demosaicing using pseudo-panchromatic image,” *IEEE Trans. Comput. Imaging*, vol. 3, pp. 982–995, Dec. 2017.
- [8] K. Feng, Y. Zhao, J. C.-W. Chan, S. Kong, X. Zhang, and B. Wang, “Mosaic convolution-attention network for demosaicing multispectral filter array images,” *IEEE Trans. Comput. Imaging*, vol. 7, pp. 864–878, 2021.
- [9] 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. Remote Sens.*, vol. 5, pp. 354–379, Apr. 2012.
- [10] M. S. Karoui, Y. Deville, S. Hosseini, and A. Ouamri, “Blind spatial unmixing of multispectral images : New methods combining sparse component analysis, clustering and non-negativity constraints,” *Pattern Recognition*, vol. 45, no. 12, pp. 4263–4278, 2012.
- [11] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, “Informed nonnegative matrix factorization methods for mobile sensor network calibration,” *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 4, pp. 667–682, 2018.
- [12] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, “Fast nonnegative matrix factorization and completion using Nesterov iterations,” in *Proc. LVA/ICA’17*, vol. 10169 of *LNCS*, pp. 26–35, Feb. 2017.
- [13] N. Guan, D. Tao, Z. Luo, and B. Yuan, “NeNMF : An optimal gradient method for nonnegative matrix factorization,” *IEEE Trans. Signal Process.*, vol. 60, pp. 2882–2898, June 2012.
- [14] Y. Deville, “Sparse component analysis : A general framework for linear and nonlinear blind source separation and mixture identification,” in *Blind Source Separation : Advances in Theory, Algorithms and Applications* (G. R. Naik and W. Wang, eds.), ch. 6, pp. 151–196, Springer, 2014.
- [15] Grupo de Inteligencia Computacional, UPV/EHU, “Hyperspectral imagery synthesis (EIAs) toolbox.” http://www.ehu.es/ccwintco/index.php/Hyperspectral_Imagery_Synthesis_tools_for_MATLAB, Last accessed : 2022-02-25.
- [16] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, C. Richard, J. Chanussot, L. Drumetz, J.-Y. Tournet, A. Zare, and C. Jutten, “Spectral variability in hyperspectral data unmixing : A comprehensive review,” *IEEE Trans. Geosci. Remote Sens.*, vol. 9, no. 4, pp. 223–270, 2021.