

Tests statistiques pour l'analyse d'acquisitions hyperspectrales codées

Trung-Tin DINH^{1,2} Hervé CARFANTAN¹ Antoine MONMAYRANT² Simon LACROIX²

¹ Institut de Recherche en Astrophysique et Planétologie (IRAP) – Université de Toulouse/CNRS/CNES,
14 Avenue Edouard Belin, 31400 Toulouse, France

²LAAS-CNRS, Université de Toulouse, CNRS
7 Avenue du Colonel Roche, 31077 Toulouse, France

Résumé – Des imageurs hyperspectraux co-conçus et programmables, relevant de l'imagerie computationnelle, permettent aujourd'hui de réduire fortement le temps d'acquisition de scènes hyperspectrales. Ils reposent sur la mesure d'une ou de quelques images codées, qui associées à un modèle palliant le manque de données, permettent d'estimer toute la scène hyperspectrale en réduisant la quantité de données et le temps d'acquisition d'un ordre de grandeur ou plus. L'adéquation du modèle à la scène observée est la condition sine qua none d'une estimation fidèle. Nous proposons ici d'étudier un modèle de séparabilité de la scène sur des régions homogènes, similaire à la prise en compte de la variabilité intraclasse en classification spectrale supervisée, et utilisons des tests statistiques pour vérifier ce modèle à partir des données codées.

Abstract – Co-designed and programmable hyperspectral imagers, based on computational imaging, can greatly reduce the acquisition time of hyperspectral scenes. They rely on the measurement of one or a few coded images combined with a model (to compensate for the lack of data) and allow to estimate the hyperspectral scene while reducing the amount of data and the acquisition time by an order of magnitude or more. The adequacy of the model to the observed scene is the sine qua none condition for an accurate estimation. Here we study a separability assumption of the scene over homogeneous regions, which is similar to accounting for intraclass variability in supervised spectral classification, and use statistical tests to verify this model from the coded data.

1 Introduction

L'imagerie hyperspectrale (HS), c'est-à-dire l'acquisition d'une image avec un très grand nombre de bandes spectrales (typiquement plusieurs centaines), a des applications dans des domaines variés. Les procédés technologiques classiques d'imagerie HS procèdent par balayage spatial ou spectral pour remplir le cube de données (avec deux dimensions spatiales et une dimension spectrale) ce qui nécessite un grand nombre d'acquisitions. Récemment sont apparus des dispositifs dits « instantanés » (*snapshot*) nécessitant des dispositifs optiques plus complexes. Nous nous intéressons ici à l'analyse de données acquises avec un imageur à masques codés pilotables [8].

L'objectif de nos travaux est de pouvoir analyser l'image HS à partir d'un faible nombre d'acquisitions par un tel dispositif, pour différentes configurations des masques. Les données codées acquises ne donnent pas accès directement à l'image HS de la scène observée et nécessitent l'utilisation d'algorithmes de reconstruction. Nous avons proposé [1, 6] un algorithme de reconstruction rapide exploitant les propriétés du dispositif et une hypothèse de séparabilité de l'image HS dans des régions spatiales homogènes. Dans ces régions, l'image HS se réduit à un spectre et une carte d'intensité. La difficulté principale devient alors de découper l'image en régions homogènes.

Nous proposons ici d'étudier l'hypothèse de séparabilité de l'image dans des régions homogènes. D'un point de vue modélisation, nous nous retrouvons dans une situation similaire à la modélisation de la variabilité intraclasse en classification ou démixage d'images HS. Notre hypothèse de séparabilité est équivalente à une version simplifiée du modèle de Hapke [2]. Nous étudions ici l'estimation de la carte d'intensité à partir des données et la validité de ce modèle sur des données réelles.

A partir de ce modèle, nous cherchons à déterminer si un

pixel donné appartient ou non à une région homogène de spectre connu, en exploitant les données codées du pixel en question. Nous proposons pour cela d'exploiter des tests statistiques permettant de valider la compatibilité des données au modèle. Dans ce cadre, ce problème se rapproche également des problèmes de classification supervisée où l'on doit décider de l'appartenance d'un spectre à une classe donnée, la difficulté principale provenant ici du fait que l'on ne dispose pas directement du spectre du pixel mais d'un faible nombre de données codées pour ce spectre.

L'ensemble de ces travaux sont validés sur des données réelles obtenues par notre dispositif.

2 Dispositif et reconstruction SA

2.1 Dispositif et modèle associé

Notre dispositif [8], dont le principe d'acquisition est illustré Fig. 1, est de type DD-CASSI [3] (*Dual Disperser-Coded Aperture Snapshot Spectral Imager*) à masque codé pilotable.

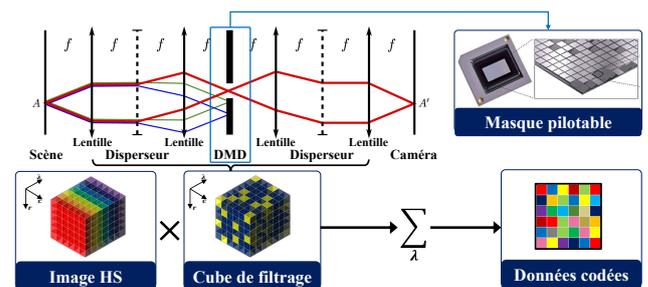


FIGURE 1 : Principe d'acquisitions codées par notre dispositif.

Il possède deux propriétés importantes : i) il est pilotable via le contrôle d'une matrice de micro-miroirs (*Digital Micromirror Device*, DMD); ii) il possède une propriété de co-localisation : les acquisitions conservent les structures spatiales de l'image HS. Le DMD est un masque spatial programmable où chaque micro-miroir bloque ou transmet la lumière. L'ensemble optique complet du DMD et des deux disperses se comporte comme un cube de filtrage hyperspectral programmable contrôlant les composantes spectrales intégrées sur chaque pixel de la caméra lors d'une acquisition codée.

Un exemple de données codées obtenues et du masque DMD utilisé pour une scène constituée de briques de LEGO® sont présentés Fig. 2. Notons qu'avec ce dispositif, grâce à la propriété de co-localisation, l'acquisition effectuée lorsque tous les miroirs sont ouverts correspond à l'image panchromatique de la scène. Cette image panchromatique a la même résolution spatiale et est acquise avec la même réponse spectrale des détecteurs que les acquisitions codées.

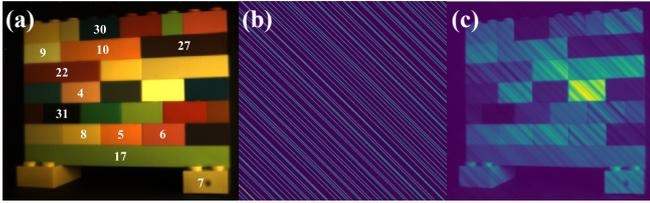


FIGURE 2 : [D'après [6]] (a) Image RGB de la scène observée labélisée²; (b) Masque DMD utilisé; (c) Acquisition codée.

En outre, il permet également l'acquisition complète du cube hyperspectral en mesurant chaque composante spatio-spectrale indépendamment, en programmant des séquences de fentes sur le DMD [7], mais cela nécessite, comme les imageurs classiques, un grand nombre d'acquisitions.

Le dispositif d'acquisition étant linéaire, il se modélise aisément sous forme matricielle. En notant \mathbf{o} le vecteur de dimension RCW correspondant à l'image HS de dimensions spatiales ($R \times C$) et de W longueurs d'onde et \mathbf{d} le vecteur de dimension RCS correspondant aux données codées acquises lors de S acquisitions successives (pour différentes configurations du DMD), le modèle s'écrit sous la forme :

$$\mathbf{d} = \mathbf{H}\mathbf{o} + \mathbf{b}, \quad (1)$$

où la matrice \mathbf{H} modélise la dispersion, le filtrage spatio-spectral pour chaque configuration du DMD et l'intégration sur les longueurs d'onde et \mathbf{b} représente le bruit de mesure.

2.2 Reconstruction SA

L'intérêt d'un tel dispositif est de pouvoir analyser la scène à partir d'un faible nombre d'acquisitions, sans avoir à balayer tout le cube ($S < W$). Dans cette situation, la reconstruction de l'objet à partir des données est un problème inverse sous-déterminé. Cette reconstruction nécessite alors la prise compte d'informations *a priori* sur la scène. Ainsi, nous avons proposé des algorithmes de reconstruction rapide à partir de dix fois moins d'acquisitions que de longueurs d'onde [1, 5, 6].

La méthode SA [1, 6] (pour *Separability Assumption*) est particulièrement simple et rapide. Comme son nom l'indique,

²Les numéros apparaissant sur les briques sont des labels qui seront mentionnés par la suite et non des éléments de la scène.

elle s'appuie sur une hypothèse de séparabilité de la scène. Plus précisément, elle suppose que la scène HS peut être découpée en régions spatiales homogènes contenant des pixels de même spectre à un coefficient d'intensité près et que ce coefficient peut se déduire aisément de l'image panchromatique. Ainsi, pour les pixels d'une telle région homogène r , l'image HS s'écrit sous la forme du produit tensoriel $\mathbf{o}_r = \mathbf{p}_r \otimes \mathbf{s}_r$ où \mathbf{s}_r correspond au spectre de la région r et \mathbf{p}_r rassemble les intensités panchromatiques de l'ensemble des pixels de cette région. En exploitant cette hypothèse de séparabilité et la propriété de co-localisation du dispositif, on peut écrire le modèle de l'équation (1) pour la région r sous la forme $\mathbf{d}_r = \mathbf{H}_r \mathbf{o}_r + \mathbf{b}_r = \mathbf{G}_r \mathbf{s}_r + \mathbf{b}_r$ où \mathbf{d}_r rassemble l'ensemble des acquisitions pour tous les pixels de la région r et \mathbf{H}_r contient les lignes et colonnes de la matrice \mathbf{H} correspondant à cette région. Pour une région avec suffisamment de pixels (supérieur à W/S) le problème est sur-déterminé et il est alors aisé d'estimer le spectre \mathbf{s}_r à partir de ces données, par exemple par moindres carrés, éventuellement régularisés.

Un des problèmes principaux de cette méthode est de déterminer les régions spatialement homogènes dans l'image HS, comme cela est illustré sur des résultats typiques présentés dans la Fig. 3. Nous avons proposé pour cela une segmentation en régions de l'image panchromatique [1, 6] (Fig. 3 (b)). Mais il est possible que des pixels de la scène aient la même intensité sur l'image panchromatique tout en ayant des spectres différents. De plus, réaliser une telle segmentation est un problème complexe. Pour s'en affranchir, une autre approche a été proposée [6], dans laquelle la reconstruction est réalisée par blocs de taille fixée (Fig. 3 (b)), l'hypothèse de séparabilité n'étant alors pas valide pour certains blocs.

Comme illustré dans la Fig. 3, même si la majorité de la scène a été bien reconstruite, il reste certaines petites régions ou pixels isolés qui nécessitent une correction, quelle que soit l'approche utilisée. Nous pouvons alors envisager de réattribuer ces pixels via une approche de type classification supervisée, en utilisant comme spectres de référence ceux obtenus par reconstruction des grandes régions homogènes.

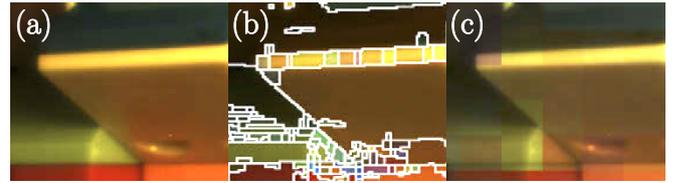


FIGURE 3 : [D'après [6]] Images RGB de la scène (a) et de la reconstruction par SA par régions (b) et par bloc (c).

3 Vers un problème de classification supervisé

Notre objectif est d'associer un spectre de référence à chaque pixel de l'image hyperspectrale à partir d'un faible nombre S d'acquisitions codées (typiquement $S = 10 \ll W = 110$). Nous disposons pour cela des acquisitions codées, de l'image panchromatique et d'une reconstruction initiale fournissant un ensemble de spectres de référence $\{\bar{\mathbf{s}}^k\}_k$. Pour résoudre ce problème, nous proposons d'étudier plus en avant le modèle utilisé et l'estimation de ses paramètres, le modèle de bruit

associé, puis d'exploiter des tests statistiques pour valider ou non les hypothèses considérées.

3.1 Hypothèse de séparabilité et variabilité intraclasse

Le modèle de Hapke [4] donne la réflectance en fonction de la longueur d'onde λ et des conditions d'observation et d'illumination d'un objet. Cela se traduit par une variabilité des spectres d'un même matériau. Une approximation de ce modèle [2] permet d'établir une relation linéaire entre le spectre \mathbf{o}_n d'un pixel n de l'image HS et un spectre de référence $\bar{\mathbf{s}}$ d'un autre pixel du même matériau : $\mathbf{o}_n \approx \psi_n \bar{\mathbf{s}}$. On retrouve donc un modèle équivalent à l'hypothèse de séparabilité exploité dans la méthode SA.

Si on dispose de l'image HS et donc du spectre \mathbf{o}_n en tout point n , il est simple d'estimer le paramètre ψ_n pour un spectre de référence $\bar{\mathbf{s}}$, par exemple par moindres carrés :

$$\widehat{\psi}_n^{(o)} = (\bar{\mathbf{s}}^T \bar{\mathbf{s}})^{-1} \bar{\mathbf{s}}^T \mathbf{o}_n. \quad (2)$$

Dans notre situation, on ne dispose que des données codées \mathbf{d} suivant le modèle (1). En écrivant ce modèle pour les données correspondant à un pixel n , et en exploitant le modèle de variabilité ci-dessus, on obtient :

$$\mathbf{d}_n = \mathbf{H}_n \mathbf{o}_n + \mathbf{b}_n = \mathbf{H}_n \bar{\mathbf{s}} \psi_n + \mathbf{b}_n. \quad (3)$$

Le coefficient ψ_n peut alors être estimé par moindres carrés :

$$\widehat{\psi}_n^{(d)} = (\bar{\mathbf{s}}^T \mathbf{H}_n^T \mathbf{H}_n \bar{\mathbf{s}})^{-1} \bar{\mathbf{s}}^T \mathbf{H}_n^T \mathbf{d}_n. \quad (4)$$

Enfin, nous pouvons relier le coefficient ψ_n à l'image panchromatique p_n au pixel n . Celle-ci correspondant à la somme suivant les longueurs d'onde du vecteur \mathbf{o}_n , nous obtenons donc $p_n = \mathbf{1}^T \mathbf{o}_n = \mathbf{1}^T \bar{\mathbf{s}} \psi_n$, avec $\mathbf{1}$ un vecteur contenant uniquement des 1. Ceci donne un estimateur de ψ_n à partir de l'image panchromatique :

$$\widehat{\psi}_n^{(p)} = \frac{p_n}{\mathbf{1}^T \bar{\mathbf{s}}}. \quad (5)$$

Nous avons étudié ces estimateurs à partir de données réelles obtenues par notre dispositif. Pour cela, nous avons fait l'acquisition, sur la scène illustrée Fig.2, de l'image panchromatique p , de données codées \mathbf{d} (pour $S = 10$ acquisitions) et de l'image HS \mathbf{o} (pour $W = 110$ longueurs d'onde). Les résultats d'estimation sont illustrés Fig. 4 sur l'ensemble des pixels d'une région homogène de l'image (délimitée arbitrairement) pour un spectre de référence correspondant à la moyenne sur cette région des spectres \mathbf{o}_n .

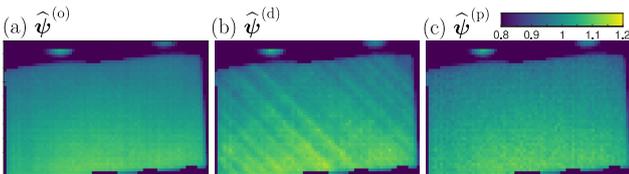


FIGURE 4 : Estimation de ψ avec différents estimateurs.

Les trois estimateurs sont cohérents en terme d'amplitude et de répartition spatiale des coefficients ψ_n . On note cependant des rayures diagonales sur $\widehat{\psi}_n^{(d)}$ sans doute liés à la configuration du DMD (voir Fig. 2b) et à une modélisation imparfaite

de la matrice \mathbf{H} . L'image obtenue par $\widehat{\psi}_n^{(p)}$ n'a pas de motifs particuliers mais elle est moins lisse que celle obtenue par $\widehat{\psi}_n^{(o)}$, ce qui est dû à la présence du bruit sur l'image panchromatique acquise. Pour conclure, l'objectif étant d'éviter de faire l'acquisition de toute l'image HS \mathbf{o} , l'estimateur $\widehat{\psi}_n^{(o)}$ n'est pas exploitable en pratique ; l'acquisition de l'image panchromatique se faisant aisément l'exploitation de $\widehat{\psi}_n^{(p)}$, comme utilisé dans la méthode SA est donc pertinente.

3.2 Modèle de bruit et analyse statistique des résidus

Il est classique, lors de la résolution des problèmes inverses, de considérer que le bruit sur les données est stationnaire, centré, Gaussien, indépendant, de variance σ^2 . Cependant, d'un point de vue théorique, la plus grande source de bruit est liée au détecteur (bruit de grenaille) et suit une loi de Poisson. Il faut néanmoins également prendre en compte le bruit de modèle et autres bruits instrumentaux. Le bruit de Poisson est bien souvent approximé par un bruit Gaussien quand le flux est suffisant, mais la stationnarité nécessite de plus que le flux varie peu. Un tel modèle Gaussien est donc peu plausible sur les images HS \mathbf{o} , puisque le flux est très variable aux différentes longueurs d'onde, mais il peut-être plus vraisemblable sur les données codées \mathbf{d} , puisqu'elles correspondent en chaque pixel à un mélange de plusieurs composantes spectrales.

Pour étudier ce bruit, nous avons étudié les erreurs résiduelles sur l'image HS $\mathbf{r}_n^{(o)} = \mathbf{o}_n - \bar{\mathbf{s}} \psi_n$ et sur les données codées $\mathbf{r}_n^{(d)} = \mathbf{d}_n - \mathbf{H}_n \bar{\mathbf{s}} \psi_n$ pour des pixels de régions homogènes connues. Plus précisément, nous avons estimé les spectres de référence (spectres moyens) de ces régions à partir de l'image HS \mathbf{o} et des coefficients $\widehat{\psi}_n^{(p)}$, calculé l'erreur résiduelle et estimé leurs écarts-types $\widehat{\sigma}_o$ et $\widehat{\sigma}_d$. Nous affichons Fig. 5 les fonctions de répartition empirique des résidus normalisés $\mathbf{r}_n^{(o)}/\widehat{\sigma}_o$ et $\mathbf{r}_n^{(d)}/\widehat{\sigma}_d$ ainsi que celle d'une loi gaussienne centrée réduite. Il est clair que le modèle blanc gaussien sta-

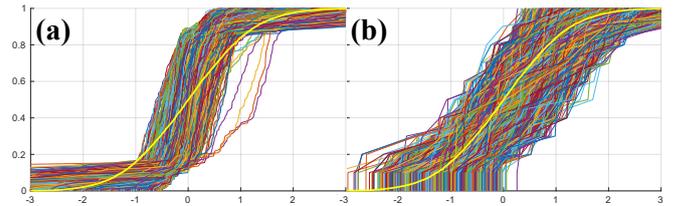


FIGURE 5 : Fonctions de répartition empiriques des résidus normalisés sur l'image HS (a), sur les données codées (b) et de la loi gaussienne $\mathcal{N}(0, 1)$ (en jaune).

tionnaire est plus vraisemblable sur les données codées \mathbf{d} que sur l'image HS \mathbf{o} mais seule l'utilisation de tests statistiques permettra de quantifier cela.

3.3 Validation par tests statistiques

Un test statistique permet de rejeter ou non une hypothèse statistique \mathcal{H}_0 avec un certain niveau de confiance $\alpha \in [0, 1]$. Ainsi, si $\alpha = 5\%$, on accepte que le test rejette \mathcal{H}_0 dans 5% des cas lorsque cette hypothèse est valide. Nous avons utilisé un tel test statistique pour tenter de (1) valider nos hypothèses de variabilité et notre modèle de bruit et (2) détecter si l'on attribue un spectre de référence erroné à un pixel.

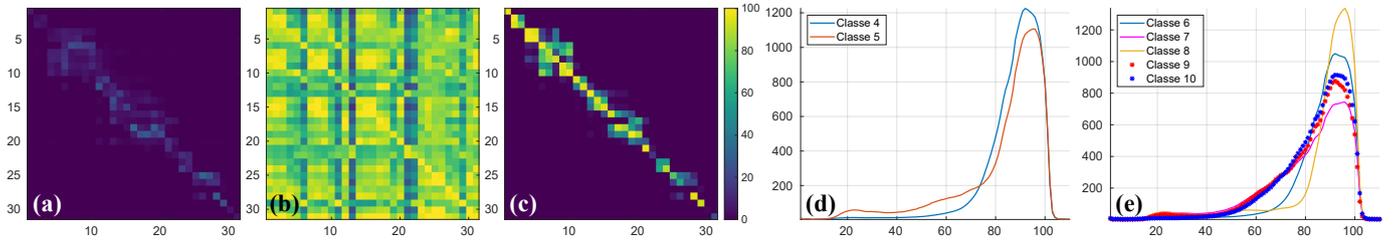


FIGURE 6 : Carte des pourcentages de validation du test de KS sur les résidus $\mathbf{r}_n^{k,\ell}$: (a) sans variabilité ; (b) $\widehat{\psi}_n^{(d)}$ et (c) $\widehat{\psi}_n^{(p)}$; Spectres de référence des classes où la confusion est la plus forte : (d) (\bar{s}_4, \bar{s}_5) et (e) $(\bar{s}_6, \bar{s}_7, \bar{s}_8, \bar{s}_9, \bar{s}_{10})$.

Nous avons choisi le test de Kolmogorov-Smirnov (KS) (avec $\alpha = 5\%$) afin de valider ou non le fait que les résidus normalisés $\mathbf{r}_n/\widehat{\sigma}$ suivent une loi Gaussienne centrée réduite lorsque le spectre de référence est correct. Nous proposons également d'utiliser ce test pour détecter la situation où le spectre de référence est erroné, soit l'attribution du pixel n à la classe ℓ alors qu'il appartient à la classe k . Le résidu est alors calculé à partir d'un spectre de référence \bar{s}^ℓ alors que les données correspondent à la région de spectre de référence \bar{s}^k :

$$\mathbf{r}_n^{k,\ell} = \mathbf{d}_n^k - \mathbf{H}_n \bar{s}^\ell \psi_n^\ell. \quad (6)$$

Nous souhaiterions donc que le test soit accepté à 95% quand $\ell = k$ et qu'il soit rejeté dans les autres cas.

Comme vu en § 3.1, nous pouvons utiliser deux estimateurs de ψ_n (sans faire l'acquisition de tout le cube \mathcal{O}) : $\widehat{\psi}_n^{(d)}$ à partir des données et $\widehat{\psi}_n^{(p)}$ à partir de l'image panchromatique.

Pour 31 régions homogènes (correspondant aux 31 briques du mur délimitées arbitrairement) de la scène observée Fig. 2, nous avons effectué les tests sur les résidus $\mathbf{r}_n^{k,\ell}$ en chaque pixel n . Les taux d'acceptation de ces tests (en %) sont représentés Fig. 6 pour l'utilisation des estimateurs $\widehat{\psi}_n^{(d)}$ et $\widehat{\psi}_n^{(p)}$, ainsi que dans le cas où l'on néglige la variabilité (soit $\psi_n = 1$). Les lignes correspondent aux classes k des régions de l'image HS et les colonnes aux classes ℓ testées. Notons que les régions sont indexées par ordre décroissant d'intensité moyenne de l'image panchromatique.

On constate clairement que négliger la variabilité (Fig. 6.a) ou utiliser de l'estimateur $\widehat{\psi}_n^{(d)}$ (Fig. 6.b) ne donne pas satisfaction. Dans le premier cas, le taux de validation de la bonne classe (validation du modèle), pour $k = \ell$ est inférieur à 35% pour toutes les classes. Dans le second, les classes erronées ne sont pratiquement jamais détectées, le taux d'acceptation pour $k \neq \ell$ étant supérieur à 80% pour la plupart des classes.

L'utilisation de l'estimateur $\widehat{\psi}_n^{(p)}$ (Fig. 6.c) est beaucoup plus satisfaisante avec la validation du modèle à plus de 80% dans la plupart des classes et une détection des classes erronées plutôt bonne. Notons cependant que pour certaines classes (17, 22, 27, 30 et 31, les labels apparaissant Fig. 2) le modèle n'est pratiquement jamais validé : la première (17) correspond à une longue brique à forte variation d'intensité, tandis que les autres sont très sombres (faible RSB). D'autre part, il y a une très forte confusion (>80% de validation) entre certaines classes (les couples (4,5) et (9,10)) qui correspondent à des briques de spectres très proches et également des confusions (entre 60% et 80%) entre le groupe (6,7,8) et le couple (9,10) de spectres assez proches (voir Fig. 6d et e). Cette dernière situation de taux d'acceptation de 60% à 80% nous alerte sur le fait que l'information apportée par un faible nombre d'acquisitions codées n'est pas suffisante pour distinguer ces spectres proches.

4 Conclusion

L'analyse d'une scène hyperspectrale à partir d'un faible nombre d'acquisitions codées est un défi qui nécessite de prendre en compte un modèle *a priori* cohérent. Nous avons montré ici, sur des données réelles, que l'hypothèse de séparabilité, équivalente à une approximation du modèle de Hapke pour la variabilité intraclasse, et la prise en compte d'un bruit blanc Gaussien permettaient de distinguer correctement les spectres à partir de dix fois moins de données codées que de longueurs d'onde. Bien sûr, le problème est particulièrement difficile pour les matériaux sombres (RSB faible) ou quand deux matériaux ont des spectres proches. Des études supplémentaires doivent cependant être menées pour une exploitation pratique de ces résultats. Pour ne citer qu'une piste, notons que pour prendre en compte l'effet convolutif de la PSF (*Point Spread Function*) optique du dispositif il semble intéressant d'envisager un modèle de type mélange spectral.

Références

- [1] I. ARDI : *Reconstruction d'images pour un imageur hyperspectrale configurable*. Thèse de doctorat, Université Toulouse III - Paul Sabatier, Toulouse, France, octobre 2020.
- [2] L. DRUMETZ : *Endmember Variability in hyperspectral image unmixing*. Thèse de doctorat, Université Grenoble Alpes, Grenoble, France, octobre 2016.
- [3] M. E. GEHM, R. JOHN, D. J. BRADY, R. M. WILLET et T. J. SCHULZ : Single-shot compressive spectral imaging with a dual-disperser architecture. *Optic Express*, 15(21), octobre 2007.
- [4] B. HAPKE : *Theory of Reflectance and Emittance Spectroscopy*. Cambridge University Press, 2ème édition, juillet 2012.
- [5] E. HEMSLEY, I. ARDI, S. LACROIX, H. CARFANTAN et A. MONMAYRANT : Optimized coded aperture for frugal hyperspectral image recovery using a dual-disperser system. *Journal of the Optical Society of America A*, 37(12):1916–1926, novembre 2020.
- [6] E. HEMSLEY, I. ARDI, T. ROUVIER, S. LACROIX, H. CARFANTAN et A. MONMAYRANT : Fast reconstruction of hyperspectral images from coded acquisitions using a separability assumption. *Optics Express*, 30(5), février 2022.
- [7] E. HEMSLEY, S. LACROIX, H. CARFANTAN et A. MONMAYRANT : Calibration of programmable spectral imager with dual disperser architecture. *Optics Communications*, 468:125767, août 2020.
- [8] S. MCGREGOR, S. LACROIX et A. MONMAYRANT : Adaptive hyperspectral imager : design, modeling, and control. *Journal of Optics*, 17(8):085607, août 2015.