

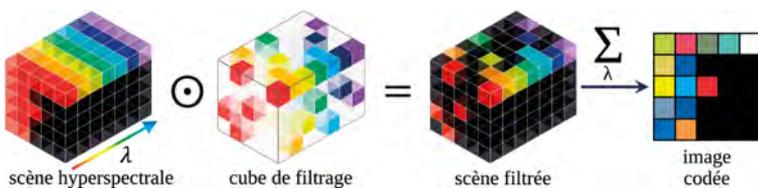
# APPORTS ET DÉFIS DE L'ACQUISITION CODÉE POUR L'IMAGERIE HYPERSPECTRALE

Antoine ROUXEL<sup>1</sup>, Trung-Tin DINH<sup>1,2</sup>, Antoine MONMAYRANT<sup>1</sup>, Simon LACROIX<sup>1</sup>, Hervé CARFANTAN<sup>2</sup>

<sup>1</sup> LAAS-CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse, France

<sup>2</sup> IRAP, Université de Toulouse, UPS/CNRS/CNES, 14 Avenue Édouard Belin, 31400 Toulouse, France

\*antoine.monmayrant@laas.fr



<https://doi.org/10.1051/photon/202312051>

Article publié en accès libre sous les conditions définies par la licence Creative Commons Attribution License CC-BY (<https://creativecommons.org/licenses/by/4.0>), qui autorise sans restrictions l'utilisation, la diffusion, et la reproduction sur quelque support que ce soit, sous réserve de citation correcte de la publication originale.

**L'imagerie hyperspectrale à ouverture codée permet l'acquisition rapide d'une scène hyperspectrale en mesurant des mélanges contrôlés d'informations spatio-spectrales présentes dans la scène. Associées à des algorithmes de traitement spécifiques, les données acquises permettent de reconstruire efficacement l'information nécessaire pour une application donnée. Ces systèmes mêlant la conception optique à la définition d'algorithmes de traitement sont une voie prometteuse pour l'imagerie hyperspectrale.**

**E**n mesurant le spectre lumineux en tous points d'une scène, l'imagerie hyperspectrale produit des données structurées en un cube défini par deux dimensions spatiales et une dimension spectrale.

L'acquisition de ce cube est classiquement réalisée par un balayage spatial d'une ligne, imagée par un capteur matriciel via un élément dispersif : il s'agit d'un processus long et contraignant, qui de plus génère une quantité importante de données à stocker, transmettre et traiter, quantité généralement très supérieure au volume d'information nécessaire aux applications.

Une solution pour accélérer l'acquisition et réduire le volume des données est de mesurer des mélanges de différents éléments du cube hyperspectral : c'est le principe de l'acquisition hyperspectrale à ouverture codée illustré sur la figure en en-tête de cet article.

L'ouverture codée définit un filtrage spatio-spectral complexe, représenté de manière logique par un « cube de filtrage, » qui est appliqué à la scène observée ; le résultat de ce filtrage est ensuite mesuré par un capteur matriciel, dont chaque pixel mesure une somme de différentes composantes spatio-spectrales de la scène.

Cette approche nécessite le développement d'algorithmes de traitement spécifiques pour interpréter les images codées.

## IMAGEURS HYPERSPECTRAUX À ACQUISITION CODÉE

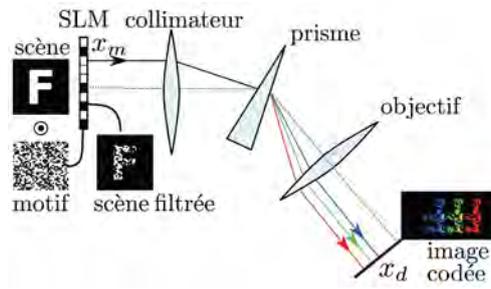
L'imagerie à ouverture codée repose sur la présence au sein du système optique d'un masque qui effectue un filtrage spatial du rayonnement collecté. Selon les arrangements relatifs du masque et des différents éléments optiques, le signal collecté est un mélange d'éléments de la scène imagée, qui permet d'encoder (et donc d'extraire) des informations spécifiques. La possibilité d'acquérir des informations de profondeur à partir d'une seule image codée a ainsi été démontrée avec un tel dispositif. Le masque utilisé pour le codage peut-être fixe ou, mieux encore, programmable avec un modulateur spatial de lumière ou spatial light modulator (SLM), le plus ●●●

souvent une matrice de micro-miroirs orientables : cette capacité à contrôler les acquisitions en définissant le codage de l'ouverture offre de nombreuses possibilités.

Pour l'imagerie hyperspectrale, le premier dispositif d'acquisition avec une ouverture codée a été proposé en 2007 [1], avec l'objectif d'acquérir tout le cube hyperspectral avec une seule acquisition (pour un masque fixe) ou un faible nombre d'acquisitions (pour des masques différents). On dénomme ce type d'imageurs CASSI (*Coded Aperture Spectral Snapshot Imager*). Différentes architectures optiques ont depuis été proposées. La figure 1 présente une des plus simples, le *Single-Disperser* CASSI (SD-CASSI) qui est semblable à un spectromètre imageur (une ligne 4-f) dont la fente d'entrée est remplacée par un SLM. L'objet est directement imagé sur le SLM qui code l'ouverture du dispositif, puis est dispersé spectralement par un prisme ou un réseau de diffraction, et enfin imagé sur un détecteur matriciel.

La présence de l'ouverture codée et d'un élément dispersif font que chaque pixel du détecteur mesure une combinaison linéaire d'intensités provenant de différentes positions et de différentes bandes spectrales de la scène, contrairement aux imageurs hyperspectraux conventionnels à fente, où chaque pixel ne mesure qu'une seule bande spectrale d'un seul élément dans la scène. Le système effectue un mélange spatio-spectral, défini par ses caractéristiques optiques et par le masque spatial réalisé par le SLM. Par conséquent, les images acquises par le détecteur ne sont pas interprétables directement, mais nous verrons que des algorithmes spécifiques permettent d'en extraire des informations.

Une autre architecture d'imageur de type CASSI est présentée sur la figure 2. Elle est définie par deux lignes 4-f contenant chacune un disperser, symétriquement arrangées autour du SLM. La première disperse spectralement l'objet qui est ensuite imagé sur le SLM, et la seconde recombine spatialement et spectralement l'image



**Figure 1 :** Architecture optique d'un imageur hyperspectral à ouverture codée Single-Disperser (SD-CASSI).

filtrée. Cette architecture dite de Double-Disperser CASSI (DD-CASSI) est plus complexe que la précédente, et donc plus encombrante, plus difficile à aligner et à étalonner, mais elle présente une caractéristique intéressante : chaque pixel du détecteur mesure cette fois une combinaison linéaire de bandes spectrales issues de la même position dans la scène observée. L'imageur conserve donc la structure spatiale de la scène, ce qui présente des avantages pour les algorithmes d'extraction d'informations. De plus, l'ouverture de l'ensemble du masque permet d'acquérir simplement l'image panchromatique de la scène (aucune des composantes spectrales n'étant bloquée).

## ALGORITHMES ET TRAITEMENTS DÉDIÉS

Les données acquises avec de tels dispositifs ne peuvent être analysées directement et nécessitent un traitement algorithmique. La plupart des travaux méthodologiques pour cette analyse

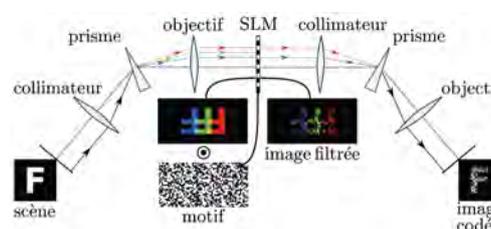
ont visé à reconstruire le cube hyperspectral de la scène observée à partir de quelques acquisitions correspondant à des configurations différentes des masques. Ces cubes reconstruits peuvent alors être analysés, par exemple pour de la classification, avec les algorithmes classiquement utilisés sur les cubes entiers acquis avec les dispositifs classiques à balayage. Plus récemment des méthodes ont été proposées afin d'extraire directement les informations souhaitées, en s'affranchissant de l'étape de reconstruction.

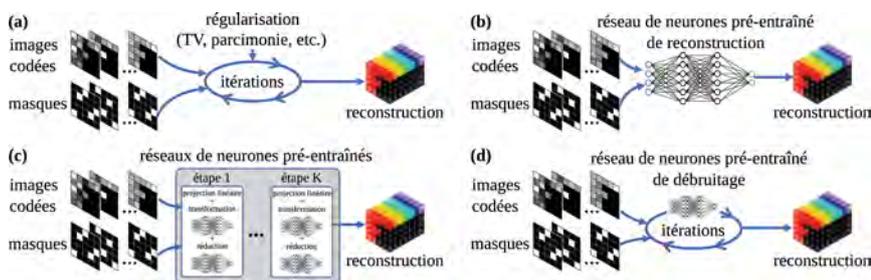
## Reconstruction du cube hyperspectral

La reconstruction du cube à partir d'un faible nombre d'acquisitions codées est un problème inverse sous-déterminé dans le sens où l'on dispose de moins de données que le nombre d'éléments du cube à reconstruire. Il est donc nécessaire d'apporter des informations supplémentaires sur le cube à reconstruire, ce qui est possible en imagerie hyperspectrale car il existe le plus souvent des corrélations fortes, aussi bien dans le domaine spectral (bandes spectrales adjacentes) que dans le domaine spatial (positions spatiales adjacentes), au sein d'un cube. On peut organiser les algorithmes de reconstruction en trois grandes familles : les algorithmes conventionnels de régularisation des problèmes inverses, ceux à base d'apprentissage (réseau de neurones profond et apprentissage), et les architectures hybrides mélangeant apprentissage et méthodes conventionnelles, ce qui est schématisé figure 3.

Les premiers travaux exploitant la régularisation se sont appuyés sur des informations *a priori* de parcimonie, s'inspirant de la théorie de l'acquisition comprimée (*compressed sensing*) [3], introduite au début des années 2000, car ils s'appuient sur des configurations aléatoires de masques utilisés. Le principe de parcimonie (*sparsity*) repose sur l'hypothèse que le signal observé peut se décrire comme une combinaison linéaire d'un faible nombre de composantes d'une famille redondante

**Figure 2 :** Architecture optique d'un imageur hyperspectral à ouverture codée Double-Disperser (DD-CASSI).





**Figure 3 :** Différentes approches de reconstruction: (a) algorithmes conventionnels de régularisation; (b) réseaux de neurones pré-entraînés; (c) et (d) structures hybrides (d'après [2]).

(appelée dictionnaire) de signaux élémentaires (appelés atomes). Ce dictionnaire peut être fixé a priori, construit par apprentissage sur un ensemble de données, voire construit par apprentissage simultanément à la reconstruction. Sous cette condition de parcimonie et moyennant une hypothèse d'incohérence de la matrice de mesure (ie. le filtrage imposé par le système d'acquisition), il est possible de reconstruire le signal avec moins de données que nécessaire selon le théorème d'échantillonnage de Nyquist-Shannon. Dans les versions les plus simples, le calcul des solutions parcimonieuses exploite des algorithmes itératifs d'optimisation minimisant un terme de fidélité aux données pénalisé par la valeur absolue des amplitudes des coefficients de chaque atome. Bien que les garanties théoriques de reconstruction du compressed sensing ne s'appliquent pas au sens strict, les matrices de mesure des imageurs CASSI n'étant que partiellement incohérentes, les résultats expérimentaux ont montré des reconstructions fidèles de cubes hyperspectraux à partir d'un nombre réduit d'acquisitions.

D'autres méthodes de régularisation classiques ont également été exploitées, consistant à pénaliser le critère de fidélité aux données par un terme favorisant des propriétés supposées pour le cube reconstruit. Deux exemples courants sont la régularisation quadratique (dite de Tikhonov), très peu coûteuse algorithmiquement qui impose un lissage des données (dans le domaine spatial et/ou spectral) et tend à flouter les contours et la régularisation par variation totale (ou *Total Variation TV*), nettement plus coûteuse, lissant les données tout en préservant les contours sans connaître a priori

leur position. Ces algorithmes itératifs d'optimisation conventionnels ont un temps de calcul élevé qui pose un véritable problème pour leur intégration dans des systèmes temps réels. Il est néanmoins possible d'exploiter dans ces algorithmes les propriétés du dispositif DD-CASSI de préservation de la structure spatiale de la scène, afin d'effectuer des reconstructions à faible coût de calcul, sur des régions homogènes de l'image de façon indépendante [4]. Un exemple (issu de [4]) est donné sur la figure 4 pour des données réelles acquises sur une scène simple, où quelques masques permettent l'acquisition d'images codées à partir desquelles la scène hyperspectrale est reconstruite, région par région (contours blancs dans la scène reconstruite). La préservation de la structure spatiale permet d'isoler et de paralléliser la reconstruction de chaque région, réduisant très fortement le temps de reconstruction.

L'utilisation de réseaux de neurones convolutionnels profonds a révolutionné le traitement d'image à base d'apprentissage (*machine learning*). Appliquées à la reconstruction de cubes hyperspectraux à partir d'acquisitions codées, de telles techniques ont l'avantage d'offrir des temps de calcul très courts pour la reconstruction. Dans leur forme générale, les modèles neuronaux modélisent à la fois la corrélation présente au sein du cube hyperspectral et le lien entre le cube hyperspectral et les acquisitions codées. Les paramètres de ces modèles sont estimés durant une phase d'apprentissage exploitant des images pour lesquelles on dispose à la fois d'acquisitions codées et les cubes hyperspectraux correspondants. Cette approche présente néanmoins des inconvénients, ●●●

# L'OPTIQUE EST NOTRE AVENIR



NOUVEAU **TECHSPEC®**

## Objectifs Corrigés à l'Infini 120i

Conçus pour réduire le poids et la taille d'un système d'imagerie tout en maintenant les performances optiques :

- Réduction de la longueur du système jusqu'à 42% par rapport aux systèmes de microscopie conventionnels
- Intégration facile dans de nombreux systèmes de vision industrielle
- Conçus pour être utilisés avec les capteurs d'imagerie de la prochaine génération

Plus d'info sur:

www.edmundoptics.fr/  
imaging



Contactez-nous:  
+33 (0) 820 207 555  
sales@edmundoptics.fr

notamment la nécessité de disposer d'une grande base d'apprentissage.

Les approches d'optimisation itérative pour la régularisation et celles basées sur les réseaux de neurones ne sont pas antagonistes. Des architectures hybrides ont été proposées telles que le *deep unfolding* et le *plug and play* (voir figures 3 (c) et (d)) s'appuyant sur des structures neuronales s'inspirant du fonctionnement des algorithmes des techniques de régularisation. Ainsi, dans l'exemple de cette dernière structure, les paramètres du réseau de neurones sont appris pour une tâche de débruitage d'images hyperspectrales. Le réseau se focalise ainsi sur la modélisation de la corrélation au sein des images, indépendamment du dispositif d'acquisition. Ce réseau de neurones et le modèle d'acquisition des données sont ensuite exploités itérativement dans l'algorithme de reconstruction.

### Extraction directe d'informations sémantiques

L'obtention d'un cube hyperspectral est rarement une fin en soi pour l'analyse d'une scène. Par exemple, les images hyperspectrales sont bien souvent exploitées pour des tâches de classification, de segmentation ou de détection d'anomalies. Aussi, si l'attention de la communauté s'est d'abord portée sur la reconstruction d'images hyperspectrales, quelques équipes de recherche ont exploré des méthodes pour extraire directement les informations sémantiques d'intérêt depuis les acquisitions codées [5], en s'affranchissant d'une étape de reconstruction du cube.

Cette approche d'inférence directe présente des avantages évidents : éviter les calculs et la propagation d'erreurs liées à l'étape de reconstruction, et limiter la quantité de données à sauvegarder et à transmettre, les informations sémantiques étant généralement beaucoup moins volumineuses que les informations spectrales brutes. Les travaux sont encore peu nombreux dans ce domaine et difficiles à évaluer de façon systématique car les algorithmes sont dimensionnés de façon à répondre à des

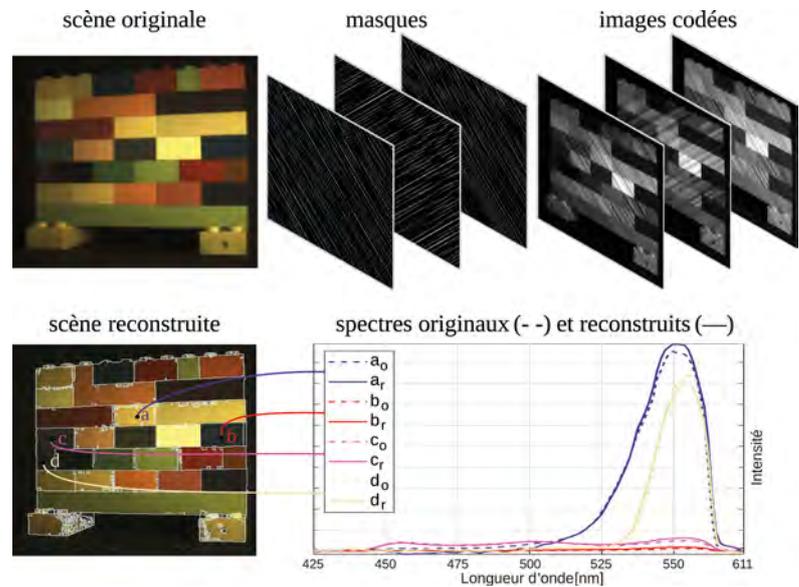


Figure 4 : Exemple de reconstruction avec l'algorithme SA [4]: image reconstruite sur 110 longueurs d'ondes à partir de 6 acquisitions codées.

besoins précis. L'arrivée des approches basées sur l'apprentissage profond, qui excelle à des tâches spécifiques, a récemment généré un regain d'intérêt. Ces méthodes sont prometteuses car exploitant pleinement l'avantage de réduction de la quantité de données inhérent aux systèmes CASSI (pas de reconstruction inutile), et s'attachant à diminuer les temps de calcul, qui est une des premières limitations à l'utilisation de systèmes CASSI.

### UNE CO-CONCEPTION NÉCESSAIRE

En imagerie traditionnelle, il existe un grand nombre de métriques reconnues et bien établies pour estimer la qualité des images mesurées (pouvoir de séparation avec différents critères, PSF, MTF, etc.). Cependant, ces métriques ne sont pas adaptées aux images issues d'instruments à ouverture codée.

Par exemple, les images hyperspectrales codées d'un SD-CASSI mélangent des composantes spectrales provenant de différentes positions dans la scène, ce qui correspond à des faibles résolutions spatiales et spectrales si l'on s'en tient aux métriques traditionnelles. Pourtant, après traitement par des algorithmes adaptés, ces images donnent accès à une scène reconstruite avec de bonnes résolutions.

Ne pas disposer de métrique faisant l'unanimité pour juger de la qualité d'acquisitions codées est donc une difficulté supplémentaire pour ce type d'imagerie. Ainsi, les performances d'un imageur hyperspectral codé doivent être validées dans un environnement de simulation associant optique et algorithmes d'exploitation des données, afin d'accéder aux cubes reconstruits ou aux informations sémantiques extraites, informations sur lesquelles des métriques claires sont établies. Si des modèles optiques simplifiés faisant fi des distorsions peuvent être aisément mis en œuvre, il n'existe pas à ce jour de moyens de simuler de façon réaliste l'acquisition d'images hyperspectrales codées. Les logiciels de conception optique couramment utilisés dans l'industrie ne permettent pas la génération de telles images, et de plus s'interface difficilement avec les environnements de traitement des données. Définir des modèles précis du processus d'acquisition pour pouvoir les implémenter dans un simulateur est le moyen de pallier ces difficultés de développement intégré. Cela permettrait de plus de quantifier explicitement l'effet des différents paramètres physiques qui caractérisent le système, offrant ainsi la capacité de dimensionner les systèmes.

**DÉFIS**

En mesurant des mélanges plutôt que des composantes individuelles d'une scène hyperspectrale, les imageurs hyperspectraux à ouverture codée réduisent le nombre d'acquisitions nécessaires et donc la quantité de données produites. En multiplexant l'information contenue dans une scène hyperspectrale, ils la mesurent sous une forme réduite que les algorithmes peuvent interpréter. Dans certains cas, ils augmentent la quantité de signal détectée en chaque pixel du détecteur et améliorent le rapport signal sur bruit par rapport à imageur à fente classique.

Bien évidemment, la réduction de la quantité de données acquise s'accompagne d'une perte d'information par rapport au contenu de la scène hyperspectrale. Cette perte est compensée par les algorithmes utilisés qui s'appuient sur des informations *a priori* sur la scène observée (présence de redondance, parcimonie, spectres présents éventuellement connus, ...). Plus ces *a priori* sont nombreux et fidèles à la réalité de la scène observée, plus les données acquises peuvent être réduites tout en conservant une bonne qualité de l'information reconstruite.

Les démonstrations expérimentales actuelles et les algorithmes associés se sont essentiellement concentrés sur la reconstruction du cube hyperspectral, avec un nombre de longueurs d'onde limité (quelques dizaines). Elles ont permis de démontrer la possibilité de réduire de plus d'un ordre de grandeur

la quantité de données à acquérir [4], transmettre et stocker. Un premier défi consiste à exploiter de telles techniques et algorithmes en augmentant la résolution spectrale, avec un nombre plus important de longueurs d'onde.

La reconstruction de l'ensemble du cube hyperspectral par de tels algorithmes n'est généralement qu'une première étape pour l'analyse de scène. Des traitements classiques sont généralement utilisés par la suite pour remonter à l'information sémantique pertinente pour une application donnée. Quelques études actuelles visent à extraire directement l'information pertinente (classification spectrale, détection d'anomalie, etc ...) des acquisitions codées, mais des recherches supplémentaires sont nécessaires.

Les algorithmes actuels s'appuient généralement sur des motifs de masques aléatoires (les masques à bruit bleu ou hyper-uniformes donnant le plus souvent les meilleurs résultats). Quelques travaux ont été menés sur l'optimisation de la configuration des masques pour un objectif donné, avec un dispositif fixé et un algorithme associé mais cela reste encore de l'ordre du défi.

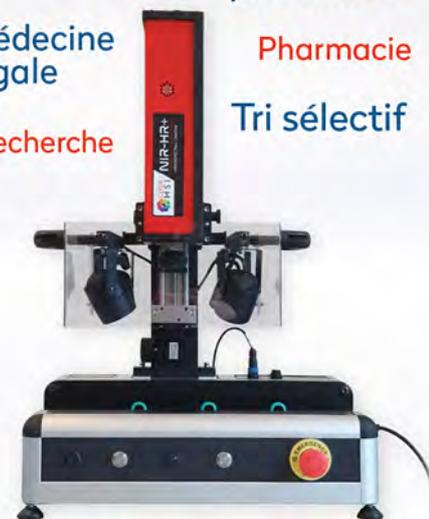
Enfin, un objectif à plus long terme est de tendre vers des schémas d'acquisition actifs, c'est-à-dire au cours desquels l'instrument adapte la configuration des masques à la scène observée au fur et à mesure des acquisitions, en fonction des acquisitions précédentes et des objectifs de mesure prédéterminés. ●

**ANALYSE PAR IMAGERIE  
HYPERSPÉCTRALE**


- Solution sur mesure dédiée à votre application
- Large choix de caméras: Visible, NIR, Raman de 400 à 2500 nm
- Large gamme de scanners motorisés

**Agroalimentaire**

- |                 |                       |
|-----------------|-----------------------|
| Tri sélectif    | Recherche             |
| Pharmacie       | Médecine légale       |
| Agroalimentaire | Science du patrimoine |
| Médecine légale | Pharmacie             |
| Recherche       | Tri sélectif          |


**RÉFÉRENCES**

- [1] M. E. Gehm, R. John, D. J. Brady, R. M. Willett, and T. J. Schulz, *Opt. Express* **15**, 14013 (2007), doi: 10.1364/oe.15.014013
- [2] X. Yuan, D. J. Brady, and A. K. Katsaggelos, *IEEE Signal Processing Magazine* **38**, 65–88 (2021), doi: 10.1109/msp.2020.3023869
- [3] E. J. Candes and M. B. Wakin, *IEEE Signal Processing Magazine* **25**, 21–30 (2008), doi: 10.1109/msp.2007.914731
- [4] E. Hemsley, I. Ardi, T. Rouvier, S. Lacroix, H. Carfantan, and A. Monmayrant, *Opt. Express* **30**, 8174 (2022), doi: 10.1364/oe.448893
- [5] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, *IEEE Journal of Selected Topics in Signal Processing* **4**, 445–460 (2010), doi: 10.1109/jstsp.2009.2039178