

1. Introduction à notre problème

Données d'Observation de la Terre = images hyperspectrales issues d'acquisition satellitaire. Chaque pixel de ces images = **mélange de spectres de réflectance** de matériaux purs.

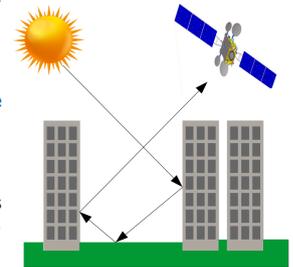
→ Nécessité d'une étape de Séparation Aveugle de Sources (SAS) pour retrouver ces spectres

Dans le cas de **scènes urbaines** : le rayon lumineux peut se refléter sur deux matériaux avant d'atteindre le satellite = **un modèle linéaire quadratique**.

→ Nécessité de construire une méthode de SAS adaptée au modèle

Dans la **littérature** : méthode de SAS repose souvent sur les **propriétés des sources** (Non négativité, indépendance statistique des sources, parcimonie...). Dans notre cas, nous ne pouvons utiliser que la non négativité, mais cette méthode ne garantit pas l'unicité de la solution...

→ Nécessité d'explorer d'autres approches pour améliorer les performances : une méthode de SAS adaptée aux type de données



2. Modèle de mélange bilinéaire

Modèle des données = non linéaire type **bilinéaire**

$$x_i(\lambda) = \sum_{j=1}^N a_{ij} s_j(\lambda) + \sum_{k=1}^{N-1} \sum_{l=k+1}^N b_{i,k,l} s_k(\lambda) s_l(\lambda) \quad (1)$$

- x les données mélangées, $i \in \{1, \dots, P\}$
- N le nombre de sources et P le nombre de pixels
- a et b les coefficients de mélange
- s les spectres sources de matériaux purs
- $s_k s_l$ **terme bilinéaire**

Cas étudié $N = 2$ sources :

$$\begin{bmatrix} x_1(\lambda) \\ x_2(\lambda) \\ x_3(\lambda) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & b_{1,1,2} \\ a_{21} & a_{22} & b_{2,1,2} \\ a_{31} & a_{32} & b_{3,1,2} \end{bmatrix} \begin{bmatrix} s_1(\lambda) \\ s_2(\lambda) \\ s_1(\lambda) s_2(\lambda) \end{bmatrix} \quad (2)$$

Volonté de créer une méthode de séparation générique nécessitant **le moins de contraintes possibles** sur les données.

→ méthode de SAS reposant sur le type de modèle et sur l'indépendance linéaire entre les sources et certains de leurs produits

3. Méthode de SAS (ou Démélange)

Création d'un système de séparation pour retrouver les signaux sources :

$$\begin{bmatrix} y_1(\lambda) \\ y_2(\lambda) \\ y_3(\lambda) \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} x_1(\lambda) \\ x_2(\lambda) \\ x_3(\lambda) \end{bmatrix} \quad (3)$$

- y les sources estimées appelées "sorties"
- c les coefficients de séparation
- x les données (mélanges de spectres)

But : estimer s_1 (ou resp. s_2) dans y_1 , s_2 (ou resp. s_1) dans y_2 et $s_1 \cdot s_2$ dans y_3

Méthode **BOCSS** (Bilinear Output Correlation-based Source Separation) basée sur la corrélation entre la multiplication terme à terme des deux premières sorties et la troisième sortie, décrite dans le critère ci-dessous.

$$J = \rho^2(y_1 \odot y_2, y_3) \quad (4)$$

avec ρ le coefficient de corrélation défini comme suit : $\rho(\alpha, \beta) = \frac{\text{cov}(\alpha, \beta)}{\sigma(\alpha)\sigma(\beta)}$, avec cov la covariance et σ l'écart-type.

Optimisation de la matrice C de séparation afin de maximiser le critère J

Une preuve de séparabilité a été réalisé pour cette méthode, on a l'unicité de la solution aux indéterminations de permutation et facteur d'échelle près, avec des contraintes d'indépendance linéaire

4. Algorithme

entrée : x les observations
sortie : y les sources estimées

- 1 Pour $t = 1$ à T :
- 2 Initialiser la matrice de mélange \tilde{A}_t^{init} ;
- 3 $C_t = \text{inv}(\tilde{A}_t^{init})$;
- 4 Adapter C_t pour maximiser J dans (4);
- 5 Si $\text{inv}(C_t) \geq 0$:
- 6 Enregistrer C_t dans C_{global} ;
- 7 Fin
- 8 Fin
- 9 Clusteriser C_{global} ;

Algorithme 1 : Méthode de séparation

Détails :

- Maximisation du critère J avec l'opposé de $fminsearch$ (fonction MATLAB®)
- Plusieurs tests = plusieurs initialisations car maxima locaux dus à $fminsearch$
- Ajout contrainte de positivité sur la matrice de mélange → propriété en Observation de la Terre
- Cluster = k-medians sur les lignes de C_{global}

2 variantes de BOCSS : avec et sans contrainte sur la diagonale de C

variante sans : 9 coefficients à optimiser (9 éléments de C) et indétermination d'échelle

variante avec : contraindre la diagonale à sa valeur initiale = 6 coefficients à optimiser et pas d'indétermination d'échelle.

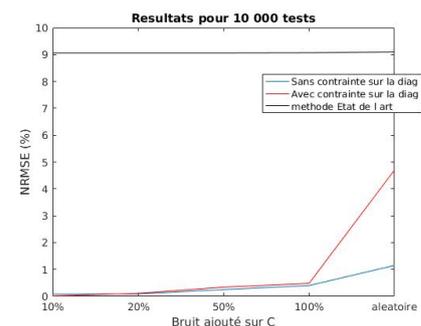
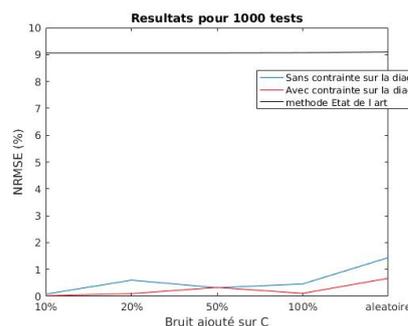
5. Résultats

Données : mélange de 2 sources "artificielles" de 10 échantillons.

1 échantillon = moyenne de 200 échantillons issus de spectres réels provenant de la librairie USGS [1]

Évaluation :

- comparaison avec une méthode de l'état de l'art [2] fondée sur la **NMF** (Non-negative Matrix Factorization) restreinte au cas bilinéaire.
- critère d'évaluation NRMSE (Normalized Root Mean Square Error)



→ **autour de 1% d'erreur d'estimation des sources pour notre méthode contre 9% pour la méthode issue de l'état de l'art !**

6. References

- [1] R. Kokaly, R. Clark, G. Swayze, K. Livo, T. Hoefen, N. Pearson, R. Wise, W. Benzel, H. Lowers, R. Driscoll, and A. Klein. USGS Spectral Library version 7: U.S. Geological Survey Data Series 1035. <https://doi.org/10.3133/ds1035>, 2017.
- [2] I. Meganem, Y. Deville, S. Hosseini, P. Déliot, and X. Briottet. Linear-quadratic blind source separation using NMF to unmix urban hyperspectral images. *IEEE Transactions on Signal Processing*, 62(7):1822–1833, April 2014.