

MÉTHODE DE SÉLECTION DES BANDES À BASE DE L'ANALYSE EN COMPOSANTES INDÉPENDANTES APPLIQUÉE AUX IMAGES HYPERSPECTRALES DE TÉLÉDÉTECTION

Seloua Chouaf¹, Youcef Smara¹

¹ : U.S.T.H.B, Faculté d'Electronique et d'Informatique.

Laboratoire de Traitement d'Images et Rayonnement, BP 32, El-Alia, Bab-Ezzouar, 16111 Alger, Algérie

chouaf.seloua@gmail.com

yousmara@yahoo.com

Résumé

Les données hyperspectrales se caractérisent par une importante dimension spectrale qui atteint quelques centaines de bandes étroites et contigües. Le volume occupé par ces images est non négligeable et fait que des tâches usuelles telles que : le stockage, le traitement et l'analyse soient alourdies. Pour pallier ce problème et garantir une meilleure exploitation de ces données, nous proposons par le présent article, une méthode de réduction qui vise à créer un espace de représentation de dimension moindre, informatif et libéré des redondances tout en préservant la signification physique des bandes. Vu le pouvoir de séparation de l'analyse en composantes indépendantes «ACI», reconnu dans le cas des données multidimensionnelles à grand volume, nous l'exploitons pour extraire un ensemble de composantes statistiquement indépendantes obtenues par la minimisation de la gaussiannité. Différents degrés d'importance sont affectés aux bandes spectrales, permettant leur classement. Finalement, nous appliquons une sélection à l'ensemble des bandes classées et nous retenons les plus informatives afin de construire l'espace de représentation spectral réduit.

Pour nos tests, nous avons appliqué l'ACI en considérant d'une part, une orthogonalisation à déflation (poursuite de projections, composantes ordonnées) et d'autre part, une orthogonalisation symétrique (estimation globale, composantes désordonnées). On suggère pour le cas symétrique, d'ajuster les données originales à un bruit additif afin d'obtenir des composantes indépendantes ordonnées (suivant le rapport signal à bruit). Les résultats recueillis montrent que la méthode proposée assure la réduction du cube hyperspectral avec d'acceptables rapports dimension-représentativité.

Mots clés : Imagerie Hyperspectrale, Sélection des Bandes, Analyse en Composantes Indépendantes, Poursuite de Projection, Ajustement au Bruit.

Abstract

A multitude of airborne and space-borne hyperspectral imaging systems has been designed to answer remote sensing demands. As what their name means, hyperspectral images are characterized by an important dimensionality which reaches hundreds of narrow and contiguous spectral bands. These images occupy a significant volume which burdens usual tasks like storage, processing and analysis. To solve this problem and improve data exploitation, we propose by the present article, a new method of reduction that aims to create a low dimensional and more informative space of representation, which is released from redundancy keeping the physical significance of the spectral bands.

Due to its recognized power of separating high dimensional data, we exploit the independent component analysis «ICA» to extract a set of statistically independent components minimizing the gaussianity. Different degrees are assigned to these bands and permit to file them. Finally, we apply to the set of the filed spectral bands, a selection that constructs the reduced spectral space. This operation can be seen as a thresholding applied to the number of the spectral bands to be kept or to the information quantity to be enclosed in the space to generate.

In our tests, we apply the ICA considering on the one hand a deflationary orthogonalisation (projection pursuit, sequential estimation, tidy components) and a symmetrical orthogonalisation (global estimation, untidy components) on the other hand. In the symmetric case, we suggest to adjust the original data to an additive noise, before applying the ICA, to obtain tidy independent components (organised following the signal to noise ratio). Experiments show that the proposed method reduces the hyperspectral data cube and creates acceptable lower-dimensional and representative spaces.

Keywords : *Hyperspectral imagery, Band selection, Independent Component Analysis, Projection pursuit, Noise adjustment.*

1. Introduction générale

L'imagerie hyperspectrale est une technologie d'acquisition optique, montée sur des systèmes aériens et spatiaux, et ce pour répondre à une multitude d'exigences relevant des domaines de l'agriculture de précision, de la foresterie, de la géologie, de la bathymétrie, de l'écologie et de la défense.

Le produit hyperspectral est souvent représenté par une figure 3D, dite cube hyperspectral. La hauteur et la largeur de ce cube sont de nature spatiale, renseignant sur l'étendue de la scène imagée. La troisième dimension, qui représente la profondeur du cube hyperspectral, est de nature spectrale renseignant sur le nombre de canaux spectraux couverts par l'imageur. Ce nombre atteint, en cas général, quelques centaines de bandes. Les bandes spectrales sont étroites et contiguës et assurent la couverture totale d'une région spectrale, généralement, comprise entre 0,3 et 2,5 μm (Goetz, 2009). La richesse de ces données donne naissance à un volume considérable qui rend difficile leur manipulation.

Comme l'information pertinente se concentre dans un nombre réduit de bandes spectrales et que le restant des bandes comporte beaucoup de redondances, le plus souvent, il s'avère inutile de traiter la totalité des données acquises. Pour remédier au problème de la dimensionnalité et en fonction de la nécessité d'intégration d'une connaissance a priori, en complémentarité avec l'ensemble des données à traiter, certaines méthodes de réduction peuvent être vues comme appartenant à l'une des catégories qui suivent :

- Des méthodes supervisées, principalement, basées sur des critères de similarité, proposent de retenir les bandes qui minimisent l'écart entre les données et les références.
- Des méthodes non-supervisées, principalement basées sur des critères de séparabilité, proposent de retenir les bandes qui maximisent l'écart entre les données générées.

Par ailleurs, un grand nombre de méthodes de réduction, citées dans la littérature (Chang, 1999 ; Du, 2003 ; Omam, 2010), suggère de commencer la procédure de réduction par effectuer une mesure, estimer un paramètre, ou apprécier une fonction objective à atteindre. Cette étape s'intéresse à l'extraction des caractéristiques et comporte dans certains cas, une phase d'optimisation. Les plus pertinentes des caractéristiques extraites sont exploitées dans la construction de l'espace de représentation réduit. Cette étape constitue ce que l'on appelle l'étape de sélection. L'espace retenu après sélection est inclus, soit dans l'espace des caractéristiques, soit dans l'espace hyperspectral. Dans le dernier cas cité, les composantes qui constituent l'espace réduit, préservent la signification physique des données originales.

Vue l'efficacité de l'ACI dans de nombreux domaines

(Hyvärinen, 2001), des travaux de recherche ont été entrepris pour l'orienter vers la réduction de la dimensionnalité et l'extraction des caractéristiques à partir des données multidimensionnelles de télédétection (Chiang, 2000 ; Lennon, 2001). Certains travaux proposent des méthodes de dé-mélange, qui arrivent à déterminer un nombre réduit de signatures spectrales sources à partir de l'ensemble des réponses spectrales observées. D'autres, proposent des versions ajustées de l'ACI, pour faciliter la réduction des données lorsque le nombre des matériaux présents dans la scène imagée est inconnu. Notons que les espaces réduits par les méthodes d'ACI décrites ci-dessus, sont d'une nature non-spectrale. Pour préserver, à la fois, le pouvoir de séparation de l'ACI et la signification spectrale des données, des méthodes proposent d'exploiter la matrice de dé-mélange obtenue par ACI pour le calcul d'un coefficient qui affecte différents niveaux de priorité aux bandes spectrales d'origine afin de faciliter leur sélection (Du, 2003 ; Omam, 2010).

En se basant sur les travaux réalisés jusqu'alors, nous présentons dans cet article, une méthode de sélection des bandes spectrales à base de versions dérivées de l'ACI. Pour cela, les deux sections qui suivent seront réservées à la présentation des outils théoriques exploités au cours de notre travail. La première s'intéresse à la phase d'extraction des caractéristiques (à savoir aux principes de l'ACI, de la PP, de l'algorithme FastICA, et de l'ajustement des données au bruit) et la seconde donne un aperçu sur la phase de sélection que nous nous sommes proposé d'appliquer aux bandes spectrales d'origines. La section d'après, aura pour objectifs d'identifier les données utilisées lors de la phase expérimentale, de présenter et comparer les principaux résultats obtenus puis de discuter l'apport de chacune des variantes proposées. Finalement, deux sections s'ajouteront aux précédentes afin de clôturer ce document. Elles apportent d'une part, une conclusion générale et d'autre part une liste des références bibliographiques citées.

2. Méthodes de projections à base de l'algorithme FastICA

2.1. L'analyse en composantes indépendantes par la minimisation de la gaussiannité

L'ACI est une technique d'analyse qui s'applique aux données multidimensionnelles, pour réaliser la séparation aveugle des sources. Elle cherche à projeter un ensemble de données aléatoires dans un espace de représentation où les composantes sont statistiquement indépendantes. La méthode réalise une projection linéaire et non-orthogonale des données originales, se basant sur un critère d'optimisation de l'indépendance (Hyvärinen, 2001).

L'estimation des composantes indépendantes peut être

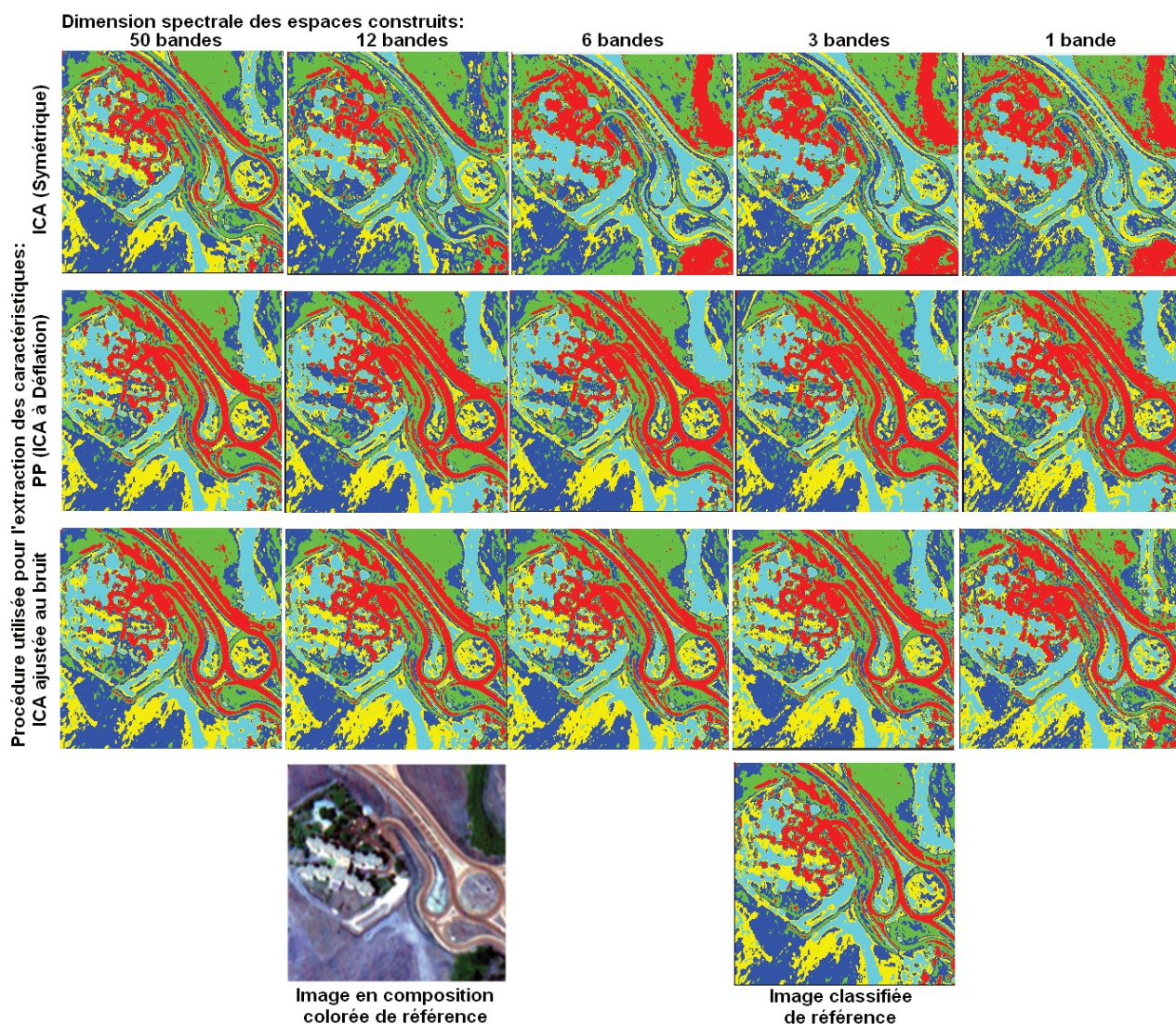


FIGURE 1 : Évaluation qualitative des résultats obtenus par la méthode proposée.

D	ICA (Symétrique)		PP (ICA à Déflation)		ICA ajustée au bruit	
	DIC (%)	TBC (%)	DIC (%)	TBC (%)	DIC (%)	TBC (%)
50	50	87,10	65,79	80,29	50	94,73
12	12	54,56	22,77	68,35	12	88,22
6	6	44,10	12,71	67,79	6	87,64
3	3	43,99	6,81	67,04	3	85,17
1	1	42,57	2,32	64,99	1	71,02

TABLEAU 1 : Évaluation quantitative des résultats obtenus par la méthode proposée.

synonyme de la recherche d'un ensemble de composantes ayant la moins de distributions gaussiennes que possible. Pour cela, le critère d'optimisation peut être formulé au travers d'une mesure de gaussiannité. Cette mesure permet d'estimer la matrice des poids, qui assure le passage de l'espace original vers l'espace d'indépendance. Le kurtosis est un cumulants d'ordre quatre, qui peut être utilisé pour mesurer la gaussiannité, vu qu'il prend une valeur nulle pour une distribution gaussienne. Cependant, cette mesure est sensible aux valeurs extrêmes. En revanche, la néguentropie, qui se lie à l'entropie peut être utilisée à cet effet. Une écriture approchée a été proposée (Hyvärinen, 2000), tel que l'indique l'équation 1 ci-dessous :

$$J(X) \approx \{[G(X)] - E[G(X_{Gauss})]\}^2, \quad (1)$$

où J est la néguentropie, $H(pX_{Gauss})$ est l'entropie d'une variable aléatoire de distribution gaussienne, et $G(\cdot)$ une fonction non-quadratique.

Ainsi, le processus de l'ACI va se charger de la recherche des composantes indépendantes qui maximisent la fonction objective énoncée par l'équation 1, ci-dessus.

2.2. La poursuite de projection

La PP est une méthode d'analyse, destinée au traitement des données multidimensionnelles à grand volume dans le but de faciliter leur visualisation. Elle consiste en la détermination d'un nombre de projections intéressantes relatives à un indice quantitatif à définir (Friedman, 1974). Il a été démontré, que les projections qui donnent naissance aux composantes les moins-gaussiennes sont les projections les plus intéressantes, ce qui rejoint le principe de transformation par l'ACI. Cependant, la PP a la particularité de déterminer les directions de projection de façon séquentielle. C'est-à-dire, qu'il suffit de trouver une seule direction qui minimise la gaussiannité, pour que l'ensemble des données soit projeté dans l'espace orthogonal à cette direction. L'aspect séquentiel de l'estimation des différentes directions, affecte à chaque projection un ordre de priorité (Lennon, 2001).

2.3. Les prétraitements de l'ACI - L'algorithme FastICA

2.3.1. Prétraitements

Des prétraitements doivent être réalisés avant d'appliquer l'ACI aux données à grand volume, afin de simplifier l'estimation des statistiques d'ordres supérieurs ainsi que pour réduire la complexité des algorithmes et accélérer la convergence. Il s'agit des prétraitements de centrage et de blanchiment. Le centrage des données originales consiste à leur imposer une moyenne nulle, ceci implique le centrage des composantes indépendantes estimées. Le blanchiment des données centrées consiste à les décorréler puis à leur imposer une variance unité. La matrice de blanchiment n'est pas unique.

Sa détermination doit vérifier que la covariance des données centrées et blanchies est une matrice identité. L'ACP est considérée comme la méthode populaire de blanchiment.

L'application de l'ACI aux données centrées et blanchies est régit par l'équation suivante :

$$Y = W \cdot X_c^b, \quad (2)$$

où Y sont les composantes indépendantes et W est la matrice des poids estimée par l'ACI.

La matrice W , dite des poids, de transformation par ACI, ou encore de dé-mélange, se constitue de vecteurs colonnes de norme unité.

2.3.2. Algorithme FastICA

L'algorithme FastICA est l'algorithme le plus favorisé pour l'application de l'ACI aux données hyperspectrales de télédétection car il représente la version la plus rapide parmi tous les algorithmes de l'ACI connus (INFOMAX, JADE, SOBI, etc.).

L'algorithme FastICA est inspiré de l'algorithme qui réalise une montée en gradient pour extraire les vecteurs colonnes de W sous la contrainte donnée par l'équation (2). Ces vecteurs sont mis à jour se basant sur le critère de minimisation de la gaussiannité par néguentropie. Une normalisation est nécessaire afin de satisfaire la contrainte de l'équation 2.

On considère qu'une valeur des vecteurs colonnes de W est un point fixe de l'algorithme du gradient. Dans ce cas la matrice W reste inchangée lors de la mise à jour et de la normalisation. Pour cela, chaque vecteur colonne doit être égale au gradient multiplié par un scalaire. Afin d'éviter l'estimation identique des vecteurs, on les décorréle par une orthogonalisation symétrique (globale) ou à déflation (séquentielle). L'application de l'ACI se fait par orthogonalisation symétrique. L'application de la PP aux données, se fait par contre, avec une orthogonalisation à déflation (Hyvärinen, 2001).

2.3.3. L'ajustement des données au bruit

En considérant que l'ensemble des données originales contienne des composantes d'un bruit additif, l'ajustement de l'ACP par le développement de la TFMB vise à produire des composantes principales rangées non seulement en décroissance de la variance mais également en fonction de la quantité du bruit qu'elles comportent. Les fractions du bruit contenues dans les données originales définissent des facteurs de qualité qui permettent d'ordonner les composantes générées (Green, 1988).

Une formulation de cet ajustement recommande d'estimer un bruit additif décorrélé par rapport aux composantes informatives et d'effectuer par la suite, une diagonalisation qui permet d'imposer à ce bruit une variance unité. Ceci restreint la recherche des directions de projections à un espace informatif dégagé du bruit. Cette

procédure d'ajustement peut être exploitée lors de la génération des composantes indépendantes (par orthogonalisation symétrique) afin de leur imposer un ordonnancement qui facilite le repérage des plus informatives entre-elles.

On considère souvent pour les données hyperspectrales de télédétection, que le bruit additif qui est à la base de l'ajustement proposé, s'estime pour chacune des bandes spectrales, par une mesure moyenne des variations radiométriques dans un voisinage spatial limité.

3. Critère de sélection des bandes spectrales

Tel qu'énoncé précédemment, la méthode que nous proposons vise à créer un nouvel espace de représentation informatif (phase d'extraction des caractéristiques), de dimension réduite (phase de sélection) mais surtout avec une signification physique (préservation de l'aspect spectral des bandes). Nous suggérons, alors, de suivre la procédure de sélection des bandes spectrales établie dans les travaux antérieurs (Du, 2003 ; Omam, 2010).

En considérant que les bandes spectrales sont les observations et que les composantes indépendantes générées sont les sources, les poids portés par les matrices de transformations à base d'ACI (obtenues au cours de la phase d'extraction des caractéristiques) renseignent sur les abondances des composantes indépendantes extraites dans les bandes hyperspectrales d'origine.

Ainsi, une mesure absolue et moyenne \bar{w}_j effectuée sur les poids w_{ij} relatifs à chacune des n bandes originales définit le coefficient de base de notre sélection. Voir l'équation 3 ci-après :

$$\bar{w}_j = \frac{1}{n} \sum_{i=1}^n |w_{ij}|, \quad (3)$$

où j est l'indice de la bande spectrale.

Plus le coefficient \bar{w}_j est important, plus la $j^{i\text{me}}$ bande spectrale est informative. Une fois que cette mesure est généralisée pour les n bandes spectrales, un ordonnancement à sa base s'impose.

La réorganisation des bandes spectrales en fonction des degrés d'importance qui leurs sont affectés permet de repérer les plus informatives d'entre-elles et qui sont favorisées d'appartenir à l'espace réduit à construire.

4. Résultats expérimentaux

La méthode proposée par cet article a été testée sur des données hyperspectrales de télédétection acquises le capteur aérien ROSIS sur une région d'Espagne (Caceres). Cette acquisition est caractérisée par une dimension spectrale qui atteint une centaine de bandes étalées entre $0,472 \mu\text{m}$ et $0,868 \mu\text{m}$ avec une très fine résolution spectrale (4 nm). La scène imagée est carrée, dimensionnée par 250 pixels de côté et avec une résolution spatiale métrique (1,2 m).

Nous avons exploité l'ACI pour l'extraction des caractéristiques utilisant l'algorithme FastICA. Dans un premier temps, nous avons considéré l'orthogonalisation symétrique tel qu'énoncé précédemment. L'expérience a montré que les composantes indépendantes produites étaient exclusivement ordonnées en fonction de la mesure du degré d'importance affecté à chaque bande spectrale. Nous avons remarqué que ces degrés étaient très proches les uns des autres et qu'ils fluctuaient autour d'une valeur moyenne. Nous avons construit des espaces avec différentes dimensionnalités et nous les avons injectés à un système de classification non-supervisée basé sur la méthode des nuées dynamiques. Les images classifiées produites nous ont permis d'évaluer la représentativité de ces espaces comparativement à l'hyperespace original. Nos tests ont révélé que les représentativités étaient acceptables pour la plupart des dimensionnalités créées.

Nous avons eu une première idée pour pondérer les bandes spectrales différemment, en imposant l'orthogonalisation à déflation à l'application de l'algorithme FastICA. Dans ce cas, la détermination des projections se fait d'une façon séquentielle permettant d'ajouter un ordre de priorité particulier aux composantes. Les expériences menées nous ont confirmé l'hypothèse et cette fois-ci, les coefficients mesurés pour les bandes spectrales étaient bien différents. Pour des dimensionnalités identiques aux précédentes, nous avons obtenu des espaces spectraux nettement plus représentatifs.

Une deuxième idée a été déduite du fait que les composantes indépendantes issues de l'orthogonalisation symétrique soient bruitées. En s'inspirant de l'idée d'ajustement de l'ACP par la TFMB, nous avons tenté d'ajuster les données originales à un bruit additif dans le but d'introduire un critère d'ordonnancement supplémentaire (en fonction de la qualité informative exprimée par le rapport signal sur bruit). Grâce à ce prétraitement, le processus de recherche des composantes indépendantes a été limité à un espace réduit libéré du bruit. Les expériences menées ont révélé que résultats obtenus par cette alternative sont très satisfaisants si comparés aux précédents.

Nous présentons ci-après, une figure qui offre une description qualitative des résultats obtenus pour différentes dimensionnalités et suite à l'application des trois procédures citées ci-dessus. Cette figure illustre également deux images que l'on peut considérer comme références. La première image représente une composition colorée qui offre une image proche de celle observée dans le cas réel. La seconde représente l'image classifiée obtenue en considérant la totalité des bandes spectrales d'origine.

Le tableau 1 ci-dessous, quant à lui, quantifie ces résultats et met en évidence les rapports qui existent entre la dimension spectrale D , les degrés d'importance cumulés DIC (associés aux bandes spectrales retenues)

et les taux de bonne classification TBC estimés sur la totalité de la scène pour évaluer la représentativité des espaces construits en prenant comme référence l'image classifiée illustrée précédemment.

5. Conclusion générale

Le présent article a eu pour objectif, la présentation d'une méthode de réduction des données hyperspectrales par la sélection des bandes. Contrairement aux méthodes classiques, cette méthode présente l'avantage de préserver la signification spectrale des données par la construction d'espaces de représentation inclus dans l'hyperespace d'origine. Elle exploite l'ACI pour évaluer l'importance des bandes spectrales, vu son grand pouvoir de séparation reconnu dans le cas des données multidimensionnelles à grand volume. La sélection des bandes spectrales a été réalisée suite à un ordonnancement introduit par la définition de trois procédures différentes.

L'expérience a montré que l'ACI peut être exploitée avec succès dans la réduction de la dimensionnalité des données et cela sans changer leur sens physique. Bien que l'application de cette méthode avec une orthogonalisation symétrique génère des espaces d'une représentativité acceptable, de meilleurs résultats sont obtenus si les données sont ajustées au préalable par rapport à un bruit additif blanc. L'orthogonalisation à déflation, quant à elle, permet de générer des espaces réduits avec une représentativité acceptable sans nécessiter le prétraitement à priori des données.

Références

Bruce, L., 2002, Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction. *IEEE Transactions on Geoscience and Remote Sensing* 40(10) :2331–2338.

Chang, C.-I., 1999, A noise subspace projection approach to determination of intrinsic dimensionality for hyperspectral imagery. *EOS/SPIE Symposium on Remote Sensing, Conference on Image and Signal Processing for Remote Sensing*, Vol. 3871, pp. 34–44, Florence, Italy.

Chiang, S., 2000, Unsupervised hyperspectral image analysis using independent component analysis. In : *IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 4, pp. 3136–3138, Honolulu, HI, USA.

Du, H., 2003. *Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop* 0-7695-2029-4/03.

Friedman, J.H., 1974. A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers* c-23(9) :881–889.

Goetz, A.F.H., 2009. Three decades of hyperspectral remote sensing of the Earth : A personal view. *Remote Sensing of Environment* 113 :S5–S16.

Green, A.A., 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing* 26(1) : 65–74.

Hyvärinen, A., 2001. *Independent Component Analysis*. John Wiley and Sons, New-York USA.

Jiemenes, L., 1995. Projection Pursuit For High Dimensional Feature Reduction : Parallel And Sequential Approaches. In : *IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, Florence, Italy.

Keshava, N., 2001. Best bands selection for detection in hyperspectral processing. In : *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. 5, pp. 3149-3152.

Landgrebe, D., 1999. *Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data*. Chapter 1 : Information, Processing for Remote Sensing. World Scientific Publishing Co., River Edge NJ, C.H. Chen edition.

Lee, J. B., 1990. Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal component transform. *IEEE Transactions on Geoscience and Remote Sensing* 28(3) : 295–304.

Lennon, M., 2001. Independent component analysis as a tool for the dimensionality reduction and the representation of hyperspectral images. In : *IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, Sydney, Australia.

Liu, X., 2009. A maximum noise fraction transform with improved noise estimation for hyperspectral images. *Sciences China Series F-Information Sciences*.

Omam, M. A., 2010, Band Selection of Hyperspectral-Image Based Weighted Independent Component Analysis *Optical Review*, Vol. 17, no. 4, 367–370.

Stein, D., 1999. Band selection for viewing underwater objects using hyperspectral sensors. In *SPIE : Conference on Airborne and In-Water Underwater Imaging*, Vol. 3761, pp. 50–61, Denver, Colorado.