

UNIVERSITÉ BRETAGNE LOIRE

THÈSE / IMT Atlantique

sous le sceau de l'Université Bretagne Loire

pour obtenir le grade de

DOCTEUR D'IMT Atlantique

Spécialité : Signal, Image, Vision

École Doctorale Mathématiques et STIC

**Et en cotutelle avec l'Ecole Nationale des Sciences
de l'Informatique**

Présentée par

Akrem Sellami

Préparée dans le département Image & traitement de
l'information

Interprétation sémantique d'images hyperspectrales basée sur la réduction adaptative de dimensionnalité

Thèse soutenue le 11 décembre 2017

devant le jury composé de :

Sami Faiz

Professeur, Inria – Rennes / président

Ali Khenchaf

Professeur, ENSTA Bretagne / rapporteur

Ouajdi Korbaa

Professeur, Inst. Sup. d'Informatique et des Tech. de Communication / rapporteur

Touria Bajjouk

Cadre de recherche, Ifremer - Brest / examinateur

Imed Riadh Farah

Professeur, Institut sup. des arts multimédia - La Manouba / co-directeur de thèse

Basel Solaiman

Professeur, IMT Atlantique / directeur de thèse

Interprétation sémantique d'images hyperspectrales basée sur la réduction adaptative de dimensionnalité

Résumé :

L'imagerie hyperspectrale permet d'acquérir des informations spectrales riches d'une scène dans plusieurs centaines, voire milliers de bandes spectrales étroites et contiguës. Cependant, avec le nombre élevé de bandes spectrales, la forte corrélation inter-bandes spectrales et la redondance de l'information spectro-spatiale, l'interprétation de ces données hyperspectrales massives est l'un des défis majeurs pour la communauté scientifique de la télédétection. Dans ce contexte, le grand défi posé est la réduction du nombre de bandes spectrales inutiles, c'est-à-dire de réduire la redondance et la forte corrélation de bandes spectrales tout en préservant l'information pertinente. Par conséquent, des approches de projection visent à transformer les données hyperspectrales dans un sous-espace réduit en combinant toutes les bandes spectrales originales. En outre, des approches de sélection de bandes tentent à chercher un sous-ensemble de bandes spectrales pertinentes. Dans cette thèse, nous nous intéressons d'abord à la classification d'imagerie hyperspectrale en essayant d'intégrer l'information spectro-spatiale dans la réduction de dimensions pour améliorer la performance de la classification et s'affranchir de la perte de l'information spatiale dans les approches de projection. De ce fait, nous proposons un modèle hybride permettant de préserver l'information spectro-spatiale en exploitant les tenseurs dans l'approche de projection préservant la localité (TLPP) et d'utiliser l'approche de sélection non supervisée de bandes spectrales discriminantes à base de contraintes (CBS). Pour modéliser l'incertitude et l'imperfection entachant ces approches de réduction et les classifieurs, nous proposons une approche évidentielle basée sur la théorie de Dempster-Shafer (DST). Dans un second temps, nous essayons d'étendre le modèle hybride en exploitant des connaissances sémantiques extraites à travers les caractéristiques obtenues par l'approche proposée auparavant TLPP pour enrichir la sélection non supervisée CBS. En effet, l'approche proposée permet de sélectionner des bandes spectrales pertinentes qui sont à la fois informatives, discriminantes, distinctives et peu redondantes. En outre, cette approche sélectionne les bandes discriminantes et distinctives en utilisant la technique de CBS en injectant la sémantique extraite par les techniques d'extraction de connaissances afin de sélectionner d'une manière automatique et adaptative le sous-ensemble optimal de bandes spectrales pertinentes. La performance de notre approche est évaluée en utilisant plusieurs jeux de données hyperspectrales réelles.

Mots clés :

Réduction de dimension, apprentissage automatique, analyse des données, imagerie hyperspectrale, algèbre multi-linéaire, sélection de bandes, extraction des caractéristiques, interprétation sémantique.

Semantic Interpretation of Hyperspectral Images based on Adaptive Reduction of Dimensionality

Abstract :

Hyperspectral imagery allows to acquire a rich spectral information of a scene in several hundred or even thousands of narrow and contiguous spectral bands. However, with the high number of spectral bands, the strong inter-bands spectral correlation and the redundancy of spectro-spatial information, the interpretation of these massive hyperspectral data is one of the major challenges for the remote sensing scientific community. In this context, the major challenge is to reduce the number of unnecessary spectral bands, that is, to reduce the redundancy and high correlation of spectral bands while preserving the relevant information. Therefore, projection approaches aim to transform the hyperspectral data into a reduced subspace by combining all original spectral bands. In addition, band selection approaches attempt to find a subset of relevant spectral bands. In this thesis, firstly we focus on hyperspectral images classification attempting to integrate the spectro-spatial information into dimension reduction in order to improve the classification performance and to overcome the loss of spatial information in projection approaches. Therefore, we propose a hybrid model to preserve the spectro-spatial information exploiting the tensor model in the locality preserving projection approach (TLPP) and to use the constraint band selection (CBS) as unsupervised approach to select the discriminant spectral bands. To model the uncertainty and imperfection of these reduction approaches and classifiers, we propose an evidential approach based on the Dempster-Shafer Theory (DST). In the second step, we try to extend the hybrid model by exploiting the semantic knowledge extracted through the features obtained by the previously proposed approach TLPP to enrich the CBS technique. Indeed, the proposed approach makes it possible to select a relevant spectral bands which are at the same time informative, discriminant, distinctive and not very redundant. In fact, this approach selects the discriminant and distinctive spectral bands using the CBS technique injecting the extracted rules obtained with knowledge extraction techniques to automatically and adaptively select the optimal subset of relevant spectral bands. The performance of our approach is evaluated using several real hyperspectral data.

Keywords :

Dimensionality reduction, machine learning, data analytics, hyperspectral imagery, multi-linear algebra, band selection, feature extraction, semantic interpretation.

DÉDICACES

À mon cher papa *Habib* et ma chère mère *Zoubeida*,

Tous les mots du monde ne sauraient traduire l'immense amour que je vous porte, ni la profonde gratitude que je vous témoigne pour tous les efforts et les sacrifices que vous n'avez jamais cessé de consentir pour mon instruction et mon bien-être.

Merci pour votre soutien durant les moments difficiles.

À ma chère fiancée *Fatma*. Je te dis MERCI pour tout. Les bons moments restent gravés dans mon cœur éternellement.

À mes chers frères *Bassem* et *Oussama* (*Aloulou*), ma chère sœur *Asma* et son époux et mon cher cousin *Khaled*. Mes chers un grand MERCI et j'espère que nos relations restent solides tout au long de notre vie.

À toutes mes grandes familles *Sellami & Inoubli*.

Aux âmes de mes grands-parents, *Abbes* (*Boucharfa*), *Habiba* et *Mariam*. Je n'oublierai jamais l'encouragement de mon cher grand-père *Boucharfa* et son soutien moral. Je lui souhaite passer un grand témoignage et un grand merci.

À mon cher grand-père *Mustapaha*. Je te remercie également pour tes encouragements.

À tous mes amis.

Et à tous ceux qui m'ont aidé de près ou de loin...

Akrem

REMERCIEMENTS

*Soyons reconnaissants aux personnes qui nous donnent
du bonheur ; elles sont les charmants jardiniers
par qui nos âmes sont fleuries.*

Marcel Proust

*Le temps met tout
en lumière.*

Thalès

Le seul moyen de se délivrer d'une tentation, c'est d'y céder paraît-il ! Alors j'y cède en disant en grand Merci aux personnes qui ont cru en moi et qui m'ont permis d'arriver au bout de cette thèse.

Au terme de cette expérience riche et passionnante, je tiens à exprimer mes plus vifs sentiments et ma haute reconnaissance à mes directeurs de thèse : Monsieur Basel SOLAIMAN, Professeur à l'IMT Atlantique, Brest et Directeur du département Image et Traitement de l'Information et Monsieur Imed Riadh FARAH, Professeur à l'Institut Supérieur des Arts Multimédias de l'université de la Manouba, pour la qualité de leurs encadrements et orientations, qui ont su conduire mes travaux de recherches et m'accompagner tout au long de ces années en combinant, de façon subtile, les critiques et, dans les moments difficiles, les encouragements. Merci donc pour votre soutien sans faille, votre confiance et votre disponibilité.

J'adresse également mes remerciements à Monsieur Sami FAIZ, Professeur à l'Institut Supérieur des Arts Multimédias, pour l'honneur qu'il m'a accordé en présidant mon jury de thèse.

Je tiens à exprimer ma profonde reconnaissance à Monsieur Ali KHENCHAF, Professeur à l'ENSTA Bretagne, et Monsieur Ouajdi KORBAA, Professeur à l'Institut Supérieur d'Informatique et des Techniques de Communication de l'université de Sousse, d'avoir accepté la fastidieuse tâche de rapporter ce travail de recherche. Je les remercie pour les remarques et les critiques qu'ils m'ont prodiguées et qui ont permis de perfectionner le présent manuscrit.

Mes remerciements vont également à Madame Touria BAJJOUK, Cadre de recherche à L'Institut Français de Recherche pour l'Exploitation de la Mer, pour l'honneur qu'elle me fait d'être l'examineur de ce travail de recherche. Je la remercie également pour ses critiques et ses remarques pertinentes.

Je remercie également Monsieur Eloi BOSSÉ, Chercheur à l'École polytechnique de Montréal, qui m'a fait plaisir de participer à ce jury.

Mes remerciements vont également à tous les membres du du département ITI et du laboratoire RIADI-GDL et à tous ceux qui m'ont aidé et encouragé tout au long de cette thèse.

Table des matières

Liste des figures	x
Liste des tableaux	xiv
Introduction générale	1
1 Contexte général	1
2 Positionnement du problème	3
3 Objectifs et contributions escomptées	5
4 Organisation du manuscrit	7
1 Imagerie satellitale hyperspectrale	11
1 Introduction	12
2 Généralités sur la télédétection	12
2.1 Concepts de base	12
2.2 Types d'images satellitales optiques	12
2.3 Résolution	16
3 Propriétés et concepts de base de l'IHS	17
3.1 Interaction du rayonnement solaire et des matériaux	18
3.2 La spectroscopie	18
3.3 Signature spectrale	19
3.4 Bibliothèques spectrales	19
3.5 Modes de représentation	20
4 Processus général d'analyse et d'interprétation d'IHS	22
5 Prétraitements	24
5.1 Corrections radiométriques	24
5.2 Corrections géométriques	24
5.3 Corrections atmosphériques	24
6 La classification	25
6.1 Principe	25
6.2 Classification supervisée et non supervisée	26
6.3 Classification paramétrique et non paramétrique	28
6.4 Classification spectro-spatiale	29
6.5 Progrès récents	30
7 De l'analyse aux défis de la grande dimensionnalité d'IHS	30
8 Conclusion	33
2 La réduction de dimensionnalité en imagerie hyperspectrale : un état de l'art	35
1 Introduction	36
2 Représentation des données hyperspectrales pour la RD	36
3 La RD en IHS	36
4 La RD basée sur les approches de projection	37

4.1	Approches par projection linéaire	38
4.2	Approches par projection non linéaire	40
5	La sélection de bandes spectrales	41
5.1	Approches par sélection supervisée	43
5.2	Approches par sélection non supervisée	45
5.3	Approches par sélection semi-supervisée	46
6	Estimation du nombre de bandes/composantes à conserver	48
6.1	Le critère de Kaiser	49
6.2	Pourcentage de la variance	49
6.3	Scree-test de Cattell	49
6.4	Rapport de croissance	51
6.5	La dimensionnalité virtuelle	51
7	Travaux existants : étude comparative	51
7.1	Tableau comparatif	51
7.2	Synthèse et discussions	54
8	Conclusion	54
3	Réduction adaptative de dimensions pour l'interprétation sémantique d'IHS	57
1	Introduction	58
2	Motivations et ébauche	58
3	De la RD vers l'extraction de la sémantique bas-niveau / haut-niveau	59
3.1	Sémantique bas-niveau	59
3.2	Sémantique haut-niveau	60
3.3	Le passage de la sémantique bas-niveau vers la sémantique haut-niveau	60
4	Représentation et approximation tensorielle d'IHS	61
4.1	Représentation tensorielle d'IHS	61
4.2	Aplatissement tensoriel des vecteurs et matrices en mode-n	61
5	Édification de l'approche proposée de RD spectro-spatiale	63
5.1	Architecture générale	63
5.2	La projection par TLPP	64
5.3	La sélection de bandes spectrales avec CBS	65
5.4	La classification spectro-spatiale avec SVM	66
5.5	La fusion de classification avec la théorie de Dempster-Shafer	67
5.6	Synthèse et discussion	68
6	Approche de RD adaptative pour l'interprétation sémantique d'IHS	69
6.1	Motivations	69
6.2	Architecture générale	70
6.3	Extraction des connaissances à partir des caractéristiques	70
6.4	Sélection adaptative des bandes spectrales pertinentes sous contraintes sémantiques	73
6.5	Discussion	75
7	Conclusion	76

4	Résultats expérimentaux	77
1	Introduction	78
2	Jeux de données	78
2.1	Indian Pines	78
2.2	Pavia University	79
3	Approche de RD spectro-spatiale	81
3.1	La RD spectro-spatiale : modèle de tenseur	81
3.2	Classification avec SVM	82
3.3	Fusion de décision avec DST	84
3.4	Évaluation de temps de calcul	86
3.5	Synthèse et discussions	87
4	Approche de RD adaptative sémantique d'IHS	88
4.1	Techniques de comparaison	89
4.2	Ajustement des paramètres	89
4.3	Extraction des règles	89
4.4	Résultats de classification et analyse de performance de la sélection de bandes spectrales	91
4.5	Impact de la sélection aléatoire des échantillons d'apprentis- sage sur la sélection de bandes spectrales	95
4.6	Interprétation des bandes spectrales sélectionnées	98
4.7	Etude de l'impact de la variation du paramètre K dans l'al- gorithme K-means	99
4.8	Synthèse et discussions	99
5	Conclusion	100
	Conclusions générales et perspectives	101
1	Synthèse	102
2	Perspectives	103
	Bibliographie	107

Liste des figures

1	Principe de l'imagerie hyperspectrale	2
1.1	Spectre du rayonnement et transparence de l'atmosphère.	13
1.2	De la vision numérique panchromatique à l'imagerie hyperspectrale, (a) : image panchromatique, (b) : image multipsectrale, (c) : image hyperspectrale	14
1.3	Imagerie hyperspectrale. (a) : Exemple d'une image panchromatique, (b) : représentation avec histogramme	14
1.4	Affichage d'une image en composition colorée	15
1.5	(a) : Exemple d'un cube hyperspectral, (b) : Concept de la signature spectrale	16
1.6	Comparaison de signatures spectrales (en réflectance) correspondant à différents matériaux : en vert la végétation, en bleu l'eau et en rouge le sol.	20
1.7	La représentation image, la plus intuitive. Un exemple d'une com- position colorée faite à partir de trois bandes d'une IHS (bande 50, bande 27 et bande 17).	21
1.8	Regroupement des pixels similaires dans un espace vecteur	22
1.9	Illustration d'un processus typique d'analyse et d'interprétation d'IHS	23
1.10	Catégorisation des techniques de classification	25
1.11	Classification ascendante hiérarchique basée sur la succession de fusion	27
1.12	Architecture classique d'un réseau mult-couches	30
2.1	Représentation d'une IHS : $I_{(3D)}$ en matrice $X_{(2D)}$	37
2.2	Les différentes catégories des techniques de RD d'IHS	37
2.3	Exemple de projection de deux bandes spectrales : l'axe en rouge permet de séparer les nuages inter-pixels pour discriminer deux classes différentes	38
2.4	Exemple de projection linéaire avec ACP	39
2.5	Exemple de projection non linéaire : (A) : Données Swiss-roll, (B) : projection linéaire, (C) : projection non linéaire	40
2.6	Exemple de matrice de corrélation spectrale inter-bandes spectrales	42
2.7	Processus général de la sélection de bandes spectrales	42
2.8	Principe général de MI : (a) : entropie $H(X_1)$ (b) MI $T(X_1, X_2)$. . .	44
2.9	Exemple d'un seuil de 95% : cas de pourcentage des valeurs propres cumulées	50
2.10	Exemple de la technique du coude avec $\alpha = 10\%$	50
3.1	Principe de projection de bandes spectrales	60
3.2	Illustration de la représentation tensorielle d'une IHS : $R \in \mathbb{R}^{M \times I_2 \times I_1}$	62

3.3	Illustration de l'aplatissement avec les tenseurs en mode-3	62
3.4	Représentation tensorielle en mode-3 d'IHS	63
3.5	Scénario générale de l'approche de RD spectro-spatiale	64
3.6	Approche proposée d'interprétation sémantique basée sur une RD adaptative	71
3.7	Les règles générées par SVM	72
4.1	Image Indian Pines. (a) : image couleur composite, (b) image de vérité terrain, (c) référence des classes	78
4.2	Les différentes signatures spectrales de 16 classes de l'Indiana Pines .	79
4.3	Image Pavia University. (a) : image couleur composite, (b) image de vérité terrain, (c) référence des classes	80
4.4	Les différentes signatures spectrales de 9 classes de Pavia University	81
4.5	Composantes extraites avec TPCA et TICA	82
4.6	Composantes extraites avec la technique de TLPP	83
4.7	Taux de classification avec TLPP, TPCA, TICA, CBS, AP, MI et BandClust	84
4.8	Les images classifiées avec SVM : (a) : SVM/CBS, (b) : SVM/TLPP	85
4.9	Fusion d'images classifiées avec DST	85
4.10	Résultats de la classification (CBS, TLPP et image fusionnée)	86
4.11	Temps de calcul pour les différentes techniques de RD	87
4.12	Temps de calcul pour la classification des différentes images réduites avec SVM	88
4.13	Comparaison des taux de classifications OA par rapport au nombre des caractéristiques p en utilisant TPP, TPCA, TLPP et TICA : (a) Indian Pines (b) Pavia University.	91
4.14	Comparaison des taux de classifications OA par rapport au nombre des bandes spectrales p en utilisant WaLuDi/TLPP, WaLuMI/TLPP, PA/TLPP et ACBS : (a) AVIRIS Indian Pines (b) ROSIS Pavia University.	92
4.15	Comparaison des taux de classifications OA par rapport au nombre des bandes spectrales p en utilisant TLPP/ACBS, SSHM/ACBS, et SSDR/ACBS : (a) AVIRIS Indian Pines (b) ROSIS Pavia University.	92
4.16	Comparison des taux de classifications OA par rapport au nombre des caractéristiques : (a) the AVIRIS Indian Pines (b) ROSIS Pavia University.	93
4.17	Les résultats de classifications obtenues par SVM de l'Indiana Pines en utilisant les différentes approches de RD : (a) ORIGIN, (b) SSDR, (c) SSHM, (d) BSICA, (e) SESVM, (f) FEWT, (g) ACBS	96
4.18	Les résultats de classifications obtenues par SVM de Pavia University en utilisant les différentes approches de RD : (a) ORIGIN, (b) SSDR, (c) SSHM, (d) BSICA, (e) SESVM, (f) FEWT, (g) ACBS	97

-
- 4.19 Bande spectrales sélectionnées avec les différentes méthodes de sélection de bandes. (a) Indian Pines ($p = 15$), (b) Pavia University ($p = 15$). La barre de couleur indique l'indice d'une bande spectrale. 98
- 4.20 L'impact de la variation du paramètre K dans l'algorithme k-means pour 15 bandes spectrales sélectionnées. (a) Indian Pines ($K = 6, 9, 14, 16, 17$ et 18). (b) Pavia University ($K = 4, 6, 8, 9, 10$ et 11) . . 99

Liste des tableaux

1.1	Les spectro-imageurs commerciales et de recherche [Pons, 2014] . . .	19
2.1	Estimation du nombre de bandes/composantes à retenir	48
2.2	Quelques travaux de projection linéaire/non linéaire en IHS	52
2.3	Quelques travaux de sélection de bandes spectrales	53
4.1	Classes de vérité terrain et leurs distributions (Indian Pines)	79
4.2	Classes de vérité terrain et leurs distributions (Pavia University) . . .	80
4.3	bandes sélectionnées en utilisant CBS, AP, MI et BandClust	83
4.4	classes de vérité terrain et leurs distributions entre l'ensemble d'apprentissage et de test	84
4.5	Matrice de confusion de l'image classifiée avec SVM ($p = 9$)	86
4.6	Règles obtenues pour Indiana Pines	90
4.7	Règles obtenues pour Pavia University	90
4.8	Taux de classification pour l'IHS de Indian Pines AVIRIS ($p = 15$) . .	94
4.9	Taux de classification pour l'IHS de Pavia University ROSIS ($p = 15$)	95
4.10	Les bandes sélectionnées avec ACBS. Classes d'Indian Pines : Alfalfa (C1) et Corn-N (C5). Classes de Pavia University : Asphalt (C1) et Meadows(C5)	95
4.11	Mesures de stabilité de bandes sur dix sous-ensembles de bandes sélectionnées pour différentes méthodes de sélection de bandes ($p = 15$). .	98

Liste des abréviations et des acronymes

PAN	Image PANchromatique
IMS	Image MultiSpectrale
IHS	Image HyperSpectrale
IUS	Image UltraSpectrale
RVB	Rouge, Vert, Bleu
SVM	Séparateur à vastes marges
KPP	K plus proches voisins
LPP	Projection préservant la localité
TLPP	Projection préservant la localité basée sur les tenseurs
CBS	Sélection de bandes basée sur les contraintes
LLE	Collecteur linéaire local
MI	Information mutuelle
VD	Dimensionnalité virtuelle
PBS	Sélection progressive de bandes
ACP	Analyse en composantes principales
ACI	Analyse en composante indépendantes
LDA	Analyse linéaire discriminante
AP	Propagation d'affinité

Introduction générale

1 Contexte général

Depuis quelques décennies, la technologie d'observation de la terre joue un rôle très important dans plusieurs domaines d'applications, à savoir le suivi de la croissance urbaine, l'analyse des zones minérales, le suivi des phénomènes naturels, etc..., en particulier l'imagerie de télédétection multi-bandes qui a connue un progrès agile suite au développement des nouveaux capteurs hyperspectraux. Au départ, l'acquisition des images de télédétection à partir des capteurs aéroportés ou satellitaires a été effectuée en mode panchromatique, où chaque image contient une seule bande spectrale en niveau de gris. Ensuite, pour une meilleure analyse de données satellitaires, des capteurs ont permis d'acquérir une même scène dans différentes longueurs d'ondes. En d'autre terme, chaque image satellitaire acquise contient plusieurs bandes spectrales. Aujourd'hui, les récents capteurs ont conduit à l'émergence de l'imagerie hyperspectrale : la scène observée est acquise dans plusieurs centaines, voire milliers de bandes spectrales contiguës et étroites (Figure 1)¹. Grâce à l'essor rapide des capteurs satellitaires qui sont de plus en plus performants en termes de résolution spectrale et spatiale, le traitement d'images hyperspectrales est devenu depuis quelques années un domaine de recherche extrêmement porteur, menant à de nombreux domaines d'applications, tels que la classification, la détection de cibles, etc. En effet, la télédétection hyperspectrale est une technologie où sa croissance est de plus en plus rapide. Elle augmente la perception et la connaissance de la surface de la terre [Ahmad et al., 2012]. Les capteurs hyperspectraux fournissent de grandes quantités d'information précise des propriétés spectrales d'une scène. Grâce à la richesse et la diversité des données hyperspectrales, ces images sont favorisées par rapport aux autres types d'images satellitaires dans certaines applications telles que l'exploration minérale [Sabins, 1999, Sarahoğlu et al., 2016], foresterie [Smith et al., 2002, Özdemir et al., 2015] et l'analyse des cultures [Lelong et al., 1998, Clark, 2016]. Par exemple, le spectre de la végétation est très différent de celui des objets artificiels (environ 1100-1600 nm). Dans ce cas, une photographie simple ne serait pas capable de distinguer les objets artificiels qu'une image hyperspectrale [Mayer et al., 2000, Chaudhary et al., 2016]. L'interprétation d'une image hyperspectrale a comme téléologie l'extraction des objets d'une scène en utilisant des signatures spectrales étiquetées. En effet, les bandes spectrales d'une image hyperspectrale sont très étroites, principalement dans la plage de $5\text{ }\mu\text{m} - 20\text{ }\mu\text{m}$ en fonction du capteur d'imagerie. Grâce au nombre important de bandes spectrales, l'imagerie hyperspectrale a une résolution spectrale très élevée, ce

1. <http://mia.univ-larochelle.fr/content/view/202/146/lang,french/>

qui permet de discriminer et identifier de façon précise les différents matériaux d'une scène hyperspectrale. Cependant, avec le nombre limité des échantillons d'apprentissage et le nombre élevé de bandes spectrales, la performance de taux de classification se détériore lorsque la dimensionnalité augmente [Plaza et al., 2009] (malédiction de dimensionnalité ou phénomène de Hughes [Hughes, 1968]). En outre, le traitement de ces données hyperspectrales massives entraîne des coûts de calcul élevés dû à la taille énorme de l'image hyperspectrale acquise (150 -300 bandes spectrales). La redondance de données hyperspectrales est un grand défi aussi dans la classification d'imagerie hyperspectrale [Burgers et al., 2009]. En effet, les bandes spectrales sont trop corrélées, c'est-à-dire le nombre élevé de bandes spectrales n'augmente pas toujours la capacité de discrimination des objets d'une scène hyperspectrale pour les classifieurs. Durant la dernière décennie, pour répondre aux défis de la grande dimen-

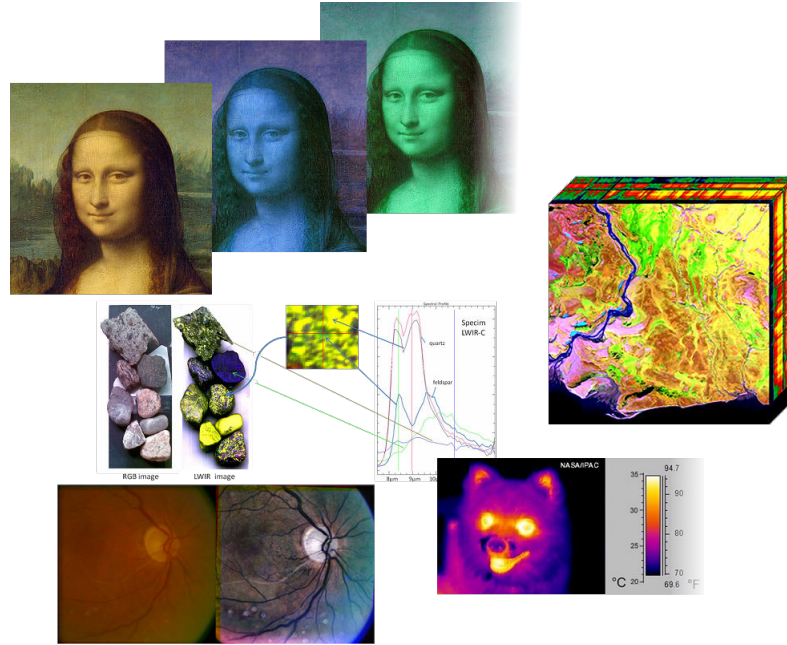


FIGURE 1 – Principe de l'imagerie hyperspectrale

sionnalité résultante dans le traitement de l'imagerie hyperspectrale et pour obtenir une meilleure précision d'interprétation, différentes approches ont été développées pour réduire le nombre de bandes spectrales tout en essayant de préserver l'information pertinente. Deux familles de réduction de dimensionnalité en imagerie hyperspectrale se sont distinguées telles que des approches par sélection de bandes spectrales et des approches par projection de données hyperspectrales. En fait, la sélection de bandes spectrales vise à sélectionner un sous-ensemble de bandes pertinentes soit d'une manière supervisée, non supervisée ou semi-supervisée selon la disponibilité des échantillons étiquetés. Parmi ces approches de sélection de bandes, nous pouvons citer la sélection basée sur l'information mutuelle (MI) [Peng et al., 2005], propagation d'affinité (AP) [Qian et al., 2009], sélection de bandes spectrales basée

sur les contraintes (CBS) [Chang and Wang, 2006], etc... Cependant, la projection consiste à transformer les bandes spectrales originales dans un nouvel sous espace réduit contenant des informations utiles. Dans ce cadre, plusieurs techniques ont été développé, notamment l'analyse en composante principale (ACP), [Jolliffe, 2002], l'analyse discriminante linéaire de Fisher (LFDA) [Teke and Sakarya, 2013], collecteur linéaire local [Fang et al., 2014], etc... D'une manière générale, la réduction de dimension est une phase très importante dans l'interprétation d'imagerie hyperspectrale (segmentation, classification, visualisation, etc.) permettant d'extraire des caractéristiques significatives et des informations pertinentes et utiles pour l'interprétation d'une scène hyperspectrale. En outre, plusieurs travaux ont prouvé que ces approches ne sont pas suffisantes pour améliorer l'interprétation d'imagerie hyperspectrale. Dans ce cadre de recherche s'inscrit le travail de cette thèse. Le contexte général de cette thèse est celui de la réduction de dimensions et l'interprétation sémantique d'imagerie hyperspectrale. Dans un cadre général, la réduction de dimensions constitue une phase essentielle dans l'interprétation d'imagerie hyperspectrale, puisqu'elle consiste à réduire la masse de données hyperspectrales en fournissant des informations riches et pertinentes qui améliorent l'interprétation sémantique en imagerie hyperspectrale. La réduction de dimensions comporte un aspect important relatif aux choix de l'approche de réduction de dimensions, soit d'une approche de sélection de bandes spectrales ou une approche de transformation de toutes les données hyperspectrales originales. Les travaux de cette thèse tentent à introduire un parangon sous la forme d'un compromis entre la réduction des données hyperspectrales massives et la préservation de l'information pertinente pour l'amélioration de l'interprétation sémantique en imagerie hyperspectrale.

2 Positionnement du problème

Le développement de approches efficaces pour la réduction de dimensions des données hyperspectrales est l'un des enjeux majeurs pour la communauté de la télédétection. Cette question ne cesse de gagner d'intérêt suite à l'augmentation incessante du volume des données hyperspectrales fournies par les nouvelles générations de capteurs placés sur les nouveaux satellites navigant dans l'atmosphère. Pour poser les problèmes traités dans ce travail de thèse, une étude est consacrée à la performance et l'efficacité des approches de réduction de dimensions en imagerie hyperspectrale, à travers un survol des travaux existants, a été préalablement réalisé.

Le questionnement principal concerne, alors, l'approche permettant de réduire les dimensions des images hyperspectrales en préservant l'information pertinente pour l'interprétation sémantique. De ce fait, ce travail de thèse vise à proposer modèle hybride et coopératif de réduction adaptative de dimensionnalité permettant de réduire les données hyperspectrales de façon automatique et adaptative en préservant l'information pertinente et le sens physique de données hyperspectrales originales pour l'amélioration de l'interprétation sémantique, notamment la classification. Par

conséquent, la thèse tente à répondre aux questions suivantes :

- *Quelle démarche à suivre pour réduire les grandes dimensions tout en améliorant la classification d'imagerie hyperspectrale ?*

La classification d'imagerie hyperspectrale est un grand défi, dû à la grande dimensionnalité des données hyperspectrales et la nécessité du nombre énorme d'échantillons d'apprentissage. En effet, la réduction de dimensions en imagerie hyperspectrale peut être réalisée de deux manières différentes : approches par projection et approches par sélection de bandes. Les approches de projection visent à produire des nouvelles caractéristiques de l'image hyperspectrale originale en transformant toutes les bandes spectrales moyennant une combinaison linéaire ou non linéaire. De ce fait, ces approches permettent d'avoir des informations significatives et nouvelles qui peuvent être utiles dans l'interprétation sémantique. Cependant, dans ces approches de projection la signification physique de données hyperspectrales originale est manquante et déformée, dû au changement de l'espace original vers un autre espace réduit et à la transformation des données originales d'entrée. Ceci entraîne une perte des informations spatiales et spectrales qui peuvent être pertinentes et significatives pour la classification et généralement à l'interprétation sémantique. Dans une approche de sélection de bandes spectrale, la signification physique est préservée, puisqu'elle permet de sélectionner les bandes spectrales de données hyperspectrales originales. Dans le cas supervisé, ces approches de sélection se focalisent essentiellement sur la d'une base d'apprentissage étiquetée pour la sélection de bandes informatives. Ce qui rend cette approche difficile à exploiter vu la non disponibilité des échantillons d'apprentissage étiquetés dans les données hyperspectrales. En outre, dans la sélection non supervisée des critères statistiques pour la sélection de bandes discriminantes. Cependant, les bandes sélectionnées par ces dernières sont redondantes, ce qui entraîne des problèmes liés à la classification. Une partie de ce travail de thèse s'activera ainsi sur la problématique de préservation de l'information spectro-spatiale dans la réduction de dimensions et l'amélioration de la classification d'imagerie hyperspectrale.

- *Quelles améliorations peut-on déduire à partir du processus de réduction de dimensions pour sélectionner l'information pertinente et extraire des connaissances pour l'interprétation sémantique ?*

La définition d'une approche adaptative de réduction de dimensions ouvre, évidemment les portes vers l'exploitation des connaissances. Bien que les approches de projection changent l'espace original en transformant les données hyperspectrales d'entrée en fournissant des nouvelles informations pertinentes, les approches de sélection non supervisée de bandes spectrales cherchent l'ensemble de bandes spectrales distinctives et discriminantes. Ceci donne lieu à une sélection des bandes redondantes et corrélées qui ne sont pas toujours utiles par exemple pour une tâche de classification et même elles détériorent les performances des taux de classification. Il est souhaitable alors de chercher des bandes spectrales à la fois discriminantes, distinc-

tives et informatives selon le contexte étudié. En d'autres termes, il s'agit de chercher une approche adaptative permettant de réduire les données massives de données hyperspectrales tout en préservant les informations pertinentes et le sens physique de données originales. De ce fait, exploitant les deux familles de réduction de dimensions serait un compromis entre la préservation de la signification physique de données originales et la sélection des informations pertinentes tout en profitant des techniques d'extraction de connaissances. En effet, l'extraction de connaissances dans le processus de réduction de dimensions permettant d'extraire les informations sémantiques qui peuvent être utiles pour la sélection de bandes spectrales pertinentes. Par conséquent, le terme sémantique signifie le passage de l'extraction des informations de bas-niveau à la découverte des informations de haut-niveau. Usuellement, l'extraction des informations bas-niveau permettant d'extraire des caractéristiques à partir de l'image originale tout en exploitant toutes les bandes spectrales d'origine. Ces caractéristiques contiennent les informations significatives extraites par le biais des approches de projection. Les informations haut-niveau qui représentent la sémantique extraite par les techniques d'extraction de connaissances à partir des informations bas-niveau qui sont fournies par les caractéristiques de l'image originale.

3 Objectifs et contributions escomptées

L'objectif général de cette thèse est de contribuer à la résolution du problème de réduction de dimensionnalité pour l'interprétation sémantique d'imagerie hyperspectrale. Pour cela, nous proposons de développer une approche de réduction adaptative de dimensionnalité générique permettant de sélectionner les bandes spectrales pertinentes : discriminantes, distinctives, peu redondantes et informatives selon le contexte étudié sans aucune nécessité de connaissances *a priori*. Généralement, l'approche de RD adaptative proposée vise à exploiter des connaissances sémantiques extraites par le biais des techniques de projection. En effet, les techniques de projection permettant d'extraire des caractéristiques pertinentes de bas-niveau, c'est à dire au niveau pixel. Physiquement, elles consistent à transformer les pixels originales de l'image afin de trouver des nouveaux points de données représentatifs de l'image d'entrée. Nous visons à partir de ces données projetées d'extraire des connaissances sémantiques de haut-niveau qui seront utiles pour améliorer la sélection de bandes spectrales dont l'objectif d'avoir une approche de RD adaptative permettant de préserver l'information pertinente en sélectionnant les bandes spectrales informatives et discriminantes. Ceci permet d'améliorer le processus de RD d'IHS ainsi l'interprétation d'une scène hyperspectrale.

Le travail de cette thèse a été divisé en deux grandes parties :

- La première partie concerne le développement d'un modèle hybride basé sur deux approches de réduction de dimensions pour l'amélioration de la classification en imagerie hyperspectrale. Après un survol de l'état de l'art

et une étude comparative des approches de réduction de dimensionnalité, nous présentons notre première contribution qui porte sur la proposition d’une approche de projection linéaire TLPP (*Tensor Locality Preserving Projection*) basée sur la technique de projection préservant la localité LPP [Wang and He, 2011] et sur le modèle de tenseur. En effet, la représentation tensorielle a été introduite dans l’approche de projection LPP afin de préserver la structure spatiale de l’image originale, notamment l’arrangement spatial et l’information de voisinage de pixels. De ce fait, le problème qui a été traité dans cette contribution est la déformation de l’information spatiale, puisque les approches de projection linéaire visent à vectoriser les données hyperspectrales d’entrée. La technique de sélection non supervisée de bandes spectrales CBS [Chang and Wang, 2006] a été utilisée dans cette partie pour sélectionner les bandes spectrales discriminantes tout en minimisant la corrélation et la dépendance entre les bandes spectrales sélectionnées. Deux images classifiées sont obtenues par la suite en utilisant le classifieur SVM. Afin de modéliser l’incertitude et l’imprécision entachées par les deux images classifiées, nous présentons notre deuxième contribution qui porte sur la fusion de deux images classifiées en utilisant la théorie de l’évidence de Dempster-Shafer DST (Dempster Shafer Theory). L’originalité, et le point fort, de notre étude porte sur la proposition d’un modèle de réduction de dimensionnalité hybride entre les approches de sélection et de projection pour l’amélioration de la classification en intégrant à la fois l’information spatiale et spectrale par les biais des tenseurs et l’algèbre multi-linéaire tout en réduisant l’incertitude et l’imperfection entachant les classifieurs par une approche de fusion évidentielle.

- La deuxième partie consiste à étendre le modèle proposé dans la première partie en intégrant des connaissances sémantiques entre les deux techniques de réduction de dimensionnalité CBS et TLPP afin de rendre l’approche adaptative permettant de sélectionner les bandes spectrales discriminantes, distinctives en informatives selon les classes de l’image hyperspectrale. Il s’agit d’une approche générique permettant de projeter les données hyperspectrales originales tout en exploitant la technique proposée TLPP dans la première partie. À partir de ces caractéristiques extraites de bas-niveau, des règles ont été générées en utilisant les deux techniques de classification k-means et SVM pour avoir des connaissances sémantiques de haut-niveau. De ce fait, la troisième contribution de cette thèse tente à extraire des connaissances sémantiques de haut-niveau à partir des caractéristiques issues des approches de projection afin de les injecter dans les approches de sélection pour aboutir à un modèle de réduction adaptative et générique permettant de préserver les informations pertinentes et le sens physique des données originales en résolvant les problèmes de corrélation et de redondance. La sémantique extraite à partir des caractéristiques de bas-niveau obtenues par la technique TLPP, ont été présentées comme étant des règles décrivant la contribution de chaque caractéristique pour la discrimination des classes de l’image origi-

nale. En se basant, sur ces connaissances sémantiques, les bandes spectrales informatives peuvent être découvertes et sélectionnées pour l'amélioration de l'interprétation sémantique d'IHS. Par conséquent, en utilisant l'approche de sélection de bandes spectrales CBS et les règles issues à partir des caractéristiques extraites moyennant la technique TLPP, le modèle sélectionne de façon adaptative et automatique l'ensemble de bandes spectrales pertinentes selon le contexte étudié.

Ce travail a été appliqué et testé sur des données hyperspectrales réelles. De ce fait, nous avons utilisé deux images hyperspectrales. Une image qui représente la zone d'Indiana Pines issue d'un capteur AVIRIS avec 220 bandes spectrales. La deuxième image représente l'université de Pavia en Italie issue d'un capteur ROSIS avec 115 bandes spectrales. Ces deux images ont été utilisées dans une application de classification basée sur l'approche proposée de réduction adaptative de dimensionnalité.

4 Organisation du manuscrit

Le travail de cette thèse est organisé en 4 chapitres. Dans un premier lieu, nous avons présenté le contexte général et la problématique liée à la réduction de dimensions d'imagerie hyperspectrale. De ce fait, les motivations du travail par rapport à la littérature ont été aussi étudiées ainsi que les objectifs et les contributions escomptées de ce travail de thèse.

Les concepts de base de la technologie hyperspectrale et le processus d'analyse d'imagerie hyperspectrale sont présentés dans le chapitre 1. Celui-ci introduit d'abord les notions fondamentales de télédétection et vise à présenter les différents types d'images satellitaires. Il contient également un tutoriel des propriétés et notions physiques sur la technologie hyperspectrale. Ensuite, nous présentons les différentes phases d'interprétation d'imagerie hyperspectrale. En effet, les différentes opérations de pré-traitement de l'imagerie hyperspectrale, notamment des corrections géométriques, atmosphériques et radiométriques sont abordées. En outre, nous présentons une revue de la littérature représentative et ciblée des travaux liés à la classification d'imagerie hyperspectrale, ainsi nous introduisons quelques travaux de progrès récents relatifs au même contexte d'étude. Cet état de l'art nous amènera à cerner les problèmes liés à la classification d'imagerie hyperspectrale. Enfin, nous présentons à la fin du chapitre les différents problèmes liés à l'analyse et l'exploitation de l'imagerie hyperspectrale.

Dans le chapitre 2, nous présentons un état de l'art concernant les études menées sur la réduction de dimensionnalité en imagerie hyperspectrale. Dans une première section, nous mettons en relation la méthode de représentation d'imagerie hyperspectrale et la réduction de dimensionnalité. Cela consiste à vectoriser l'image originale de l'espace $3D$ vers un espace $2D$. Nous introduisons aussi le formalisme mathématique de la réduction de dimensionnalité dans un contexte hyperspectral. Ensuite, nous présentons une revue bibliographique des études réalisées suivie d'une étude comparative entre les différentes approches de réduction de dimensionnalité afin de

mieux cerner les enjeux et les objectifs du travail de thèse. En outre, quelques techniques d'estimation de nombre de bandes spectrales et de composantes à conserver lors de la réduction de dimensions sont présentées. Cette étude a pour objectif de préciser l'estimation de la dimensionnalité intrinsèque d'une imagerie hyperspectrale.

Dans le chapitre 3, nous présentons les contributions de cette thèse. Dans une première partie, nous présentons l'approche proposée de la réduction spectro-spatiale pour la classification haut-niveau d'imagerie hyperspectrale. En effet, nous introduisons le modèle de tenseur ainsi que la représentation tensorielle d'imagerie hyperspectrale. Ensuite, nous développons notre approche proposée TLPP en présentant l'apport de tenseur d'intégrer l'information spectro-spatiale dans la technique LPP. Nous présentons aussi le modèle de fusion basé sur la théorie de l'évidence DST pour l'amélioration de la classification d'imagerie hyperspectrale. Nous finissons cette première partie par la présentation de l'algorithme adopté pour notre approche proposée pour la réduction spectro-spatiale. Dans la deuxième partie, nous présentons l'approche de réduction adaptative de dimensionnalité. En effet, nous focalisons sur l'extraction des caractéristiques à partir des caractéristiques extraites issues du TLPP. De ce fait, l'algorithme d'extraction des connaissances sémantiques à partir des données réduites en utilisant le k-means et le SVM est présenté. Dans ce cadre, les contributions de bandes spectrales et des composantes pour la discrimination des classes d'imagerie hyperspectrales ont été démontrés. Enfin, nous achevons ce chapitre par la présentation de l'algorithme de sélection adaptative de bandes spectrales basée sur la base de connaissance.

Le chapitre 4, vise à évaluer et valider l'approche proposée. Pour cela, nous présentons les données hyperspectrales utilisées dans l'expérimentation. Nous présentons dans un premier temps les résultats obtenu en termes de réduction et classification. Ensuite, les résultats de l'approche proposée de réduction spectro-spatiale sont comparés avec d'autres approches de réduction de dimension. Dans la deuxième expérimentation, nous utilisons les même jeux de données pour valider la deuxième approche proposée. Dans ce cadre, une évaluation quantitative des résultats obtenus avec les différentes approches est fournie.

Les différentes contributions de cette thèse ont été démontrées sur un ensemble d'expérimentations qui sont présentées dans le quatrième chapitre. En effet, nous avons essayé de valider le premier modèle hybride proposé de réduction de dimensionnalité spectro-spatiale sur des données réelles, à savoir l'image hyperspectrale AVIRIS de l'Indian Pines. Dans un premier lieu, nous avons effectuée une réduction de dimensions de cette image en utilisant les deux techniques TLPP et CBS. Le classifieur SVM a été implémenté par la suite pour classer ces deux images réduites. Ensuite, une étude comparative a été réalisée pour démontrer l'efficacité de l'approche proposée. Les résultats obtenus par notre approche par comparaison à d'autres approches de réduction de dimensionnalité montrent son efficacité. En utilisant les même jeux de données, nous avons essayé de valider la deuxième approche de sélection adaptative de bandes spectrales basée sur les connaissances sémantiques en essayant de déterminer les taux de classification obtenue pour chaque classe des

données hyperspectrales. Les résultats obtenus de l'approche de sélection adaptative donnent des meilleurs résultats en comparant par la première approche de réduction de dimensionnalité spectro-spatiale et d'autres approches de sélection de bandes spectrales.

Nous achevons ce manuscrit de thèse par une synthèse des contributions, des conclusions et quelques perspectives à court et à moyen termes.

Imagerie satellitale hyperspectrale

Sommaire

1	Introduction	12
2	Généralités sur la télédétection	12
2.1	Concepts de base	12
2.2	Types d'images satellitales optiques	12
2.3	Résolution	16
3	Propriétés et concepts de base de l'IHS	17
3.1	Interaction du rayonnement solaire et des matériaux	18
3.2	La spectroscopie	18
3.3	Signature spectrale	19
3.4	Bibliothèques spectrales	19
3.5	Modes de représentation	20
4	Processus général d'analyse et d'interprétation d'IHS	22
5	Prétraitements	24
5.1	Corrections radiométriques	24
5.2	Corrections géométriques	24
5.3	Corrections atmosphériques	24
6	La classification	25
6.1	Principe	25
6.2	Classification supervisée et non supervisée	26
6.3	Classification paramétrique et non paramétrique	28
6.4	Classification spectro-spatiale	29
6.5	Progrès récents	30
7	De l'analyse aux défis de la grande dimensionnalité d'IHS	30
8	Conclusion	33

1 Introduction

Dans ce chapitre, nous évoquons en premier lieu un aperçu général sur les techniques de la télédétection en présentant ses principaux aspects en termes d'acquisition et d'analyse. Nous nous intéressons particulièrement à la technologie hyperspectrales. Après avoir évoqué les techniques de classification d'IHS, nous dressons un bilan des problèmes liés à son exploitation.

2 Généralités sur la télédétection

2.1 Concepts de base

La télédétection, aussi appelée l'observation de la terre, désigne : "l'ensemble des connaissances et techniques utilisées pour déterminer des caractéristiques physiques et biologiques d'objets par des mesures effectuées à distance, sans contact matériel avec ceux-ci", (*Commission interministérielle de terminologie de la télédétection aérospatiale*, 1988). La télédétection peut détecter la cible de l'objet terrestre et le phénomène naturel en fonction de l'onde électromagnétique collectée, car tous les objets possèdent des caractéristiques de rayonnement électromagnétique entièrement différents en fonction de différentes catégories, caractéristiques et conditions environnementales. La technique de télédétection se fonde principalement sur le principe de la réflexion de l'objet ou de l'émission d'ondes électromagnétiques. Les capteurs de télédétection peuvent être divisés en deux grandes catégories : *passifs* et *actifs*. Les capteurs passifs mesurent les niveaux ambiants des sources d'énergie existantes, tandis que les actifs fournissent leur propre source d'énergie. La majorité de la télédétection se fait avec des capteurs passifs, pour lesquels le soleil est la principale source d'énergie. La photographie aérienne est une forme majeure de télédétection, les nouvelles technologies à l'état solide ont des capacités étendues pour la visualisation dans les longueurs d'onde visible et proche de l'infrarouge pour inclure un rayonnement solaire à longueurs d'ondes plus longues. Cependant, tous les capteurs passifs n'utilisent pas l'énergie du soleil. De ce fait, ils ne fournissent pas eux-même l'énergie détectée. Par contre, les capteurs actifs fournissent leur propre source d'énergie. Ces capteurs émettent de l'énergie dans la zone hyperfréquence du spectre électromagnétique (Figure 1.1). Dans ce cadre, nous évoquons les différents concepts de bases liés à l'imagerie de télédétection tels que : les différents types d'images satellitaires et leurs résolutions.

2.2 Types d'images satellitaires optiques

La télédétection optique utilise des capteurs visibles, à infrarouge proche et à infrarouge à ondes courtes pour produire des images de la surface terrestre en détectant le rayonnement solaire réfléchi par les objets du sol. En effet, différents matériaux se reflètent et s'absorbent différemment à différentes longueurs d'onde. Ainsi, les objets peuvent être différenciés par leurs signatures de réflectance spectrale dans les

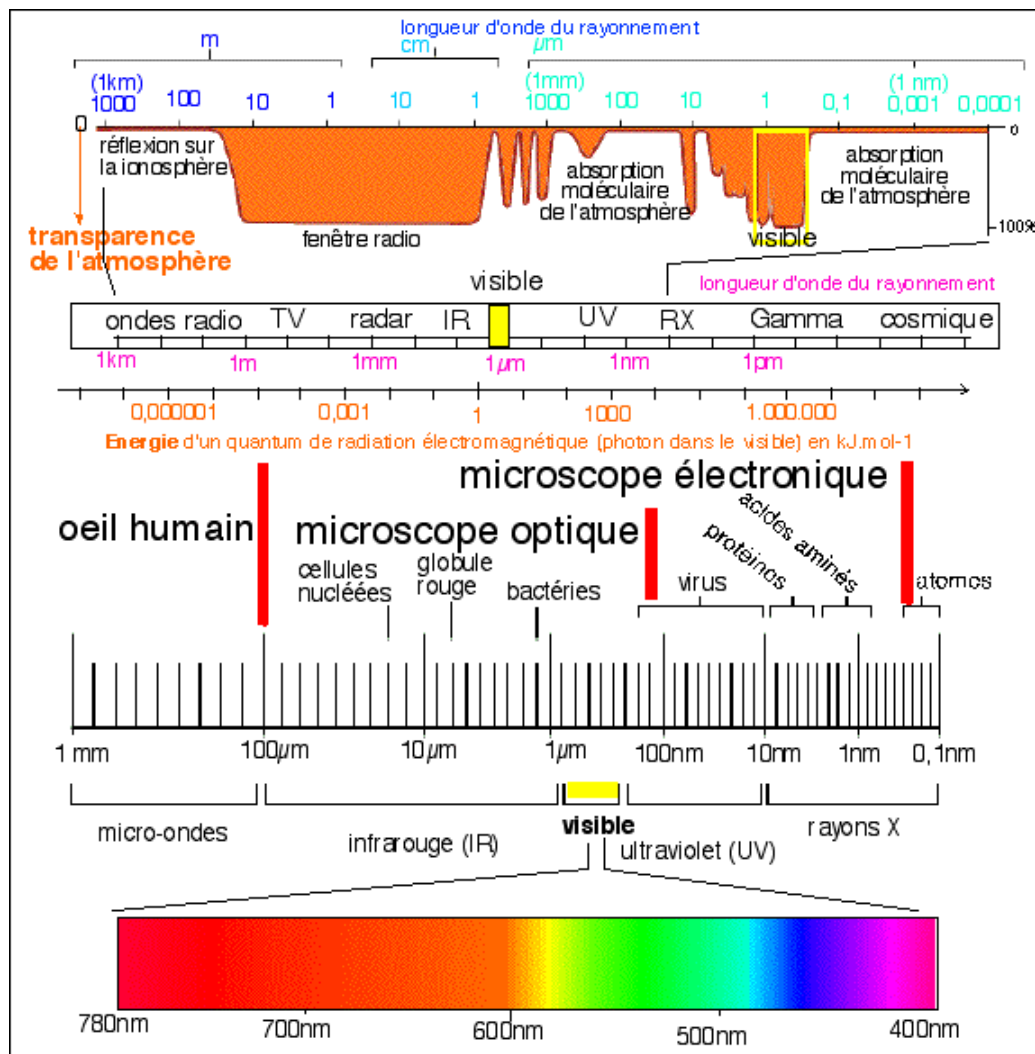


FIGURE 1.1 – Spectre du rayonnement et transparence de l'atmosphère.

images de télédétection. Actuellement, l'évolution des capteurs d'images satellitales optiques conduit à l'acquisition de centaines de bandes spectrales, ce qui augmente la capacité de discriminer chaque type d'occupation du sol (Figure 1.2). Ceci donne lieu à classer ces images selon le nombre de bandes spectrales utilisées dans le processus d'imagerie. Trois types d'images satellitales se sont distinguées tels que : des images panchromatiques, multispectrales, hyperspectrales et ultraspectrales.

2.2.1 Les images panchromatiques

Les images PANchromatiques (PAN) sont acquises à partir de l'enregistrement du rayonnement électromagnétique dans un unique intervalle du spectre visible avec la longueur d'onde entre 0.4 et 0.7 μm . Elles constituent d'une seule bande (matrice $2D$), c'est à dire elles sont affichées sous la forme d'une image en 255 niveau de gris

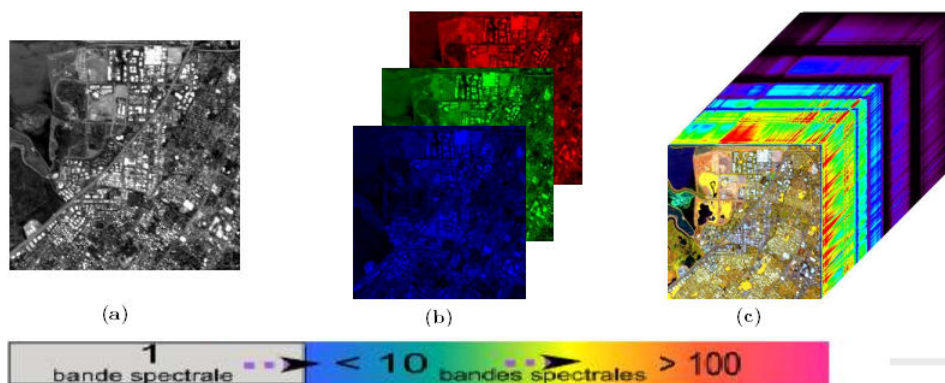


FIGURE 1.2 – De la vision numérique panchromatique à l'imagerie hyperspectrale, (a) : image panchromatique, (b) : image multispectrale, (c) : image hyperspectrale

(Figure 1.3). Bien que ces images ont une faible résolution spectrale, elles sont plus riches du point de vue résolution spatiale.

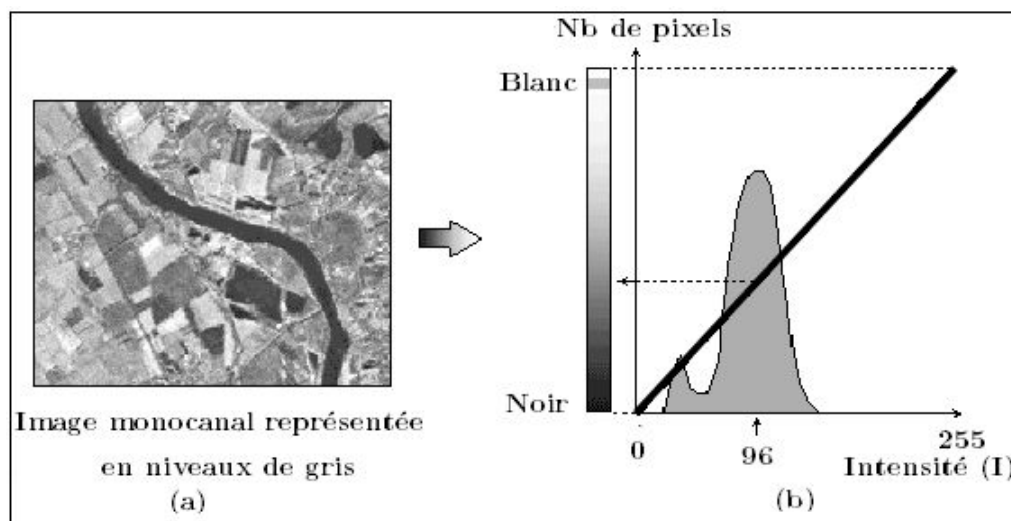


FIGURE 1.3 – Imagerie hyperspectrale. (a) : Exemple d'une image panchromatique, (b) : représentation avec histogramme

2.2.2 Les images multispectrales ou multibandes

Les Images MultiSpectrales (IMS) sont obtenues à partir des enregistrements simultanés d'une même scène avec un ensemble de bandes spectrales de longueurs d'ondes différentes. Le nombre maximum de bandes utilisées ne dépasse pas 10. La combinaison de ces bandes en utilisant les trois couleurs primaires RVB (Rouge, Vert, Bleu) permet d'obtenir des images en couleurs. Selon les valeurs numériques de chaque pixel dans chaque bande, les couleurs primaires se combineront en proportions différentes pour produire des couleurs différentes et distinctes (Figure 1.4).

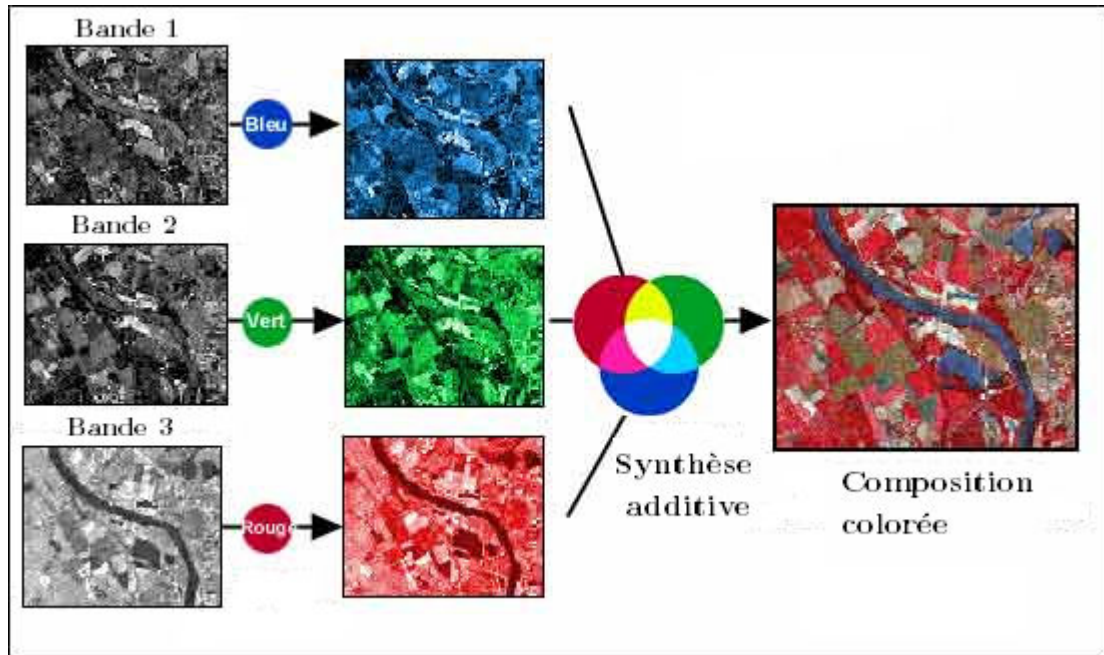


FIGURE 1.4 – Affichage d'une image en composition colorée

2.2.3 Les images hyperspectrales

L'Imagerie HyperSpectrale (IHS), aussi appelée le *spectromètre hyperspectral* fournit des informations spectrales et spatiales riches. Les spectromètres hyperspectrales sont les classes des capteurs optiques qui mesurent la radiance spectrales de chaque pixel dans une scène. Le résultat est dite IHS. Usuellement, les capteurs hyperspectraux peuvent collecter un ensemble d'images pour la représentation d'une même scène à partir des centaines de bandes spectrales contiguës et étroites avec une résolution spectrale très fine. Dans une représentation spectrale, chaque pixel p de l'IHS peut être représenté comme un vecteur de M dimensions $p = [p_1, \dots, p_z] \in \mathbb{R}^z$ avec z égale au nombre de bandes spectrales de l'IHS [Manolakis et al., 2003]. Chaque bande spectrale est assignée à une seule dimension. Dans une représentation spatiale, chaque bande correspond à une matrice $X \in \mathbb{R}^{x \times y}$ avec x et y sont deux dimensions spatiales de l'IHS. Ceci donne lieu à l'acquisition de l'IHS comme étant un cube de données de trois dimensions (matrice 3D) : deux spatiales x et y , et une spectrale z (Figure 1.5). Les deux dimensions spatiales x et y représentent la taille de chaque bande spectrale et la dimension spectrale z représente le nombre de bandes spectrales utilisées. Ainsi, chaque pixel de l'IHS est associé à un vecteur de données (radiance ou luminosité). Ces représentations spatiales et spectrales sont exploitées dans plusieurs applications d'IHS : extraction des caractéristiques, démixage spectral [Keshava and Mustard, 2002, Guerra et al., 2015,

Chen et al., 2013] et la classification [Wang et al., 2012, Teke and Sakarya, 2013, Ma et al., 2010, Chen et al., 2016].

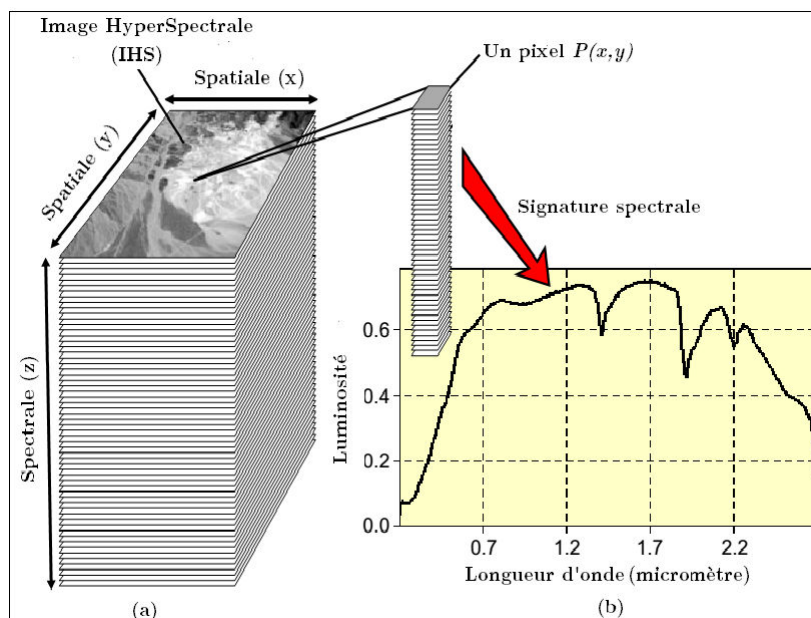


FIGURE 1.5 – (a) : Exemple d'un cube hyperspectral, (b) : Concept de la signature spectrale

2.2.4 Les images ultraspectrales

L'Imagerie UltraSpectrale (IUS), est une technologie récente permettant d'acquérir des milliers de bandes spectrales [Ehlers et al., 2003, Larar et al., 2014]. En effet, elle est très riche de plus en plus au niveau spectral que les IHSs. Dans un cadre de télédétection, ces données massives ne sont pas encore acquises par les récents capteurs satellitaires. Par conséquent, avec l'émergence et l'essor de technologie ultraspectrale, des grands défis se présentant dans leur exploitation et analyse. Ceci donne lieu à proposer et développer des nouvelles techniques supportant les grandes dimensions des IUSs.

2.3 Résolution

La résolution d'images satellitaires représente l'aptitude des capteurs à distinguer deux valeurs de deux matériaux proches l'un de l'autre. Nous distinguons quatre types de résolutions tels que : spatiale, spectrale, temporelle et radiométrique.

- **Résolution spatiale** : elle détermine la capacité du satellite pour détecter des petits matériaux. Autrement dit, la surface terrestre est couverte par un pixel d'une image satellitale. La taille d'un pixel est dépend du type de capteur satellitale et détermine la résolution de l'image. Plus les détails sont observés et clairs, plus la résolution spatiale est très importante. Les hautes

résolutions spatiales varient de 0.41 à 4 m et elles permettent la distinction entre les matériaux voisins. Tandis que les faibles résolutions spatiales varient de 30 à 1000 m , ceci donne lieu à distinguer entre deux différentes régions voisines.

- **Résolution spectrale** : elle détermine l'aptitude du capteur satellitaire à distinguer des rayonnements électromagnétiques de différentes longueurs d'ondes. Elle est définie selon le nombre de bandes spectrales et leurs largeurs. Plus la bande spectrale du capteur satellitaire est fine, plus sa résolution spectrale est très élevée. De ce fait, la résolution spectrale spécifie le nombre de bandes spectrales utilisées par le capteur satellitaire pour collecter le rayonnement réfléchi. Dans un contexte d'IHS, le nombre de bandes spectrales généralement entre 10 et 14 nm . Ainsi, dans le cas où la gamme spectrale est continue, la résolution spectrale est définie comme suit :

$$D = \frac{\text{gamme spectrale}}{\delta\lambda} \quad (1.1)$$

Avec D représente le nombre de bandes spectrales et $\delta\lambda$ est la largeur minimale d'un canal spectral. Bien que le nombre de bandes spectrales est un aspect important pour la résolution spectrale, la position des bandes spectrales dans le spectre électromagnétique est également importante.

- **Résolution temporelle** : elle détermine la fréquence de revisite d'un capteur satellitaire à une scène spécifique. Cette résolution peut être associée à la période de temps que prend un capteur pour effectuer un cycle orbital complet afin de détecter à nouveau la même scène.
- **Résolution radiométrique** : elle détermine l'aptitude du capteur satellitaire à discriminer des petites différences dans l'énergie électromagnétique. Elle représente la sensibilité du capteur satellitaire à l'intensité de l'énergie électromagnétique. Plus la résolution radiométrique d'un capteur est élevée, plus le capteur pourra déceler les petites variations dans l'énergie électromagnétique.

3 Propriétés et concepts de base de l'IHS

L'IHS est une technique qui combine l'imagerie numérique avec la spectroscopie. Elle recueille et traite des informations provenant du spectre électromagnétique en fonction de la longueur d'onde et produit des IHSs avec des instruments appelés *spectro-imageur*. Les théories de base de télédétection hyperspectrale comprend principalement les base théoriques de l'interaction du rayonnement solaire et des matériaux, les spectromètres d'imagerie, les modes d'imagerie spectrale, et les caractéristiques d'IHS, etc. [Borengasser et al., 2007]. Usuellement, les imageurs spectraux produisent et analysent les rayonnements spectraux de chaque pixel dans la scène, où les objets sont caractérisés par leur forme spatiale et leur radiance spectrale. Le rayonnement spectral d'un objet est son intensité de lumière réfléchie en fonction de la longueur d'onde, ce qui précise la matière formant l'objet. Bien que

l'œil humain puisse bien distinguer la forme spatiale, il ne peut pas distinguer les caractéristiques de rayonnement spectral.

La réflectance spectrale d'un objet est obtenue par le rapport entre la radiance et l'illumination de la scène. Tandis que, la plupart des objets peuvent être identifiés uniquement par leur réflectance spectrale. Alors que le choix des bandes spectrales pertinentes est crucial pour l'analyse d'IHS. En effet, les déflexions des courbes spectrales représentent les intervalles des longueurs d'onde pour lesquels les matériaux absorbent sélectivement l'énergie incidente. Ce phénomène est appelé communément *bande d'absorption*. Dans différents cas d'étude, la forme globale d'une courbe spectrale, la position et la force des bandes d'absorption permettent d'identifier et discriminer plusieurs matériaux. Par exemple, la végétation a une réflectance très élevée dans la plage infrarouge et une réflectance très faible de la lumière rouge que le sol. Cependant, parce que l'œil humain est un observateur relativement peu sophistiqué, il n'est pas toujours possible de distinguer les objets en fonction de leurs couleurs perçues.

3.1 Interaction du rayonnement solaire et des matériaux

La source principale du rayonnement électromagnétique sur la terre est le soleil. Lorsque le rayonnement solaire atteint le sol à travers la couche atmosphérique, certains sont réfléchis sur l'espace extérieur par la couche atmosphérique et l'autre est absorbé par l'atmosphère. En outre, certains sont dispersés pour être un rayonnement diffus et atteignent la surface de la terre. Seulement environ 31% du rayonnement solaire comme rayonnement solaire direct atteint la surface terrestre. Cette partie du rayonnement électromagnétique est en partie réfléchi par la surface de la Terre et la partie restante est absorbée par les objets terrestres. Après que la surface de la terre absorbe le rayonnement solaire, elle générera un auto-rayonnement, connu sous le nom de *rayonnement terrestre*. En effet, le rayonnement terrestre peut être divisé en rayonnement à ondes courtes ($0.3 - 2.5\mu m$) et rayonnement à ondes longues (plus de $6\mu m$). Le rayonnement à ondes courtes se concentre sur le reflet de la surface terrestre pour le soleil, et le rayonnement à ondes longues ne peut pas être considéré que le rayonnement thermique de l'objet sur la surface de la terre.

3.2 La spectroscopie

La spectroscopie, ou spectrométrie est l'étude de la lumière qui est émise ou réfléchi par des matériaux et sa variation d'énergie avec la longueur d'onde. Dans le contexte de télédétection optique, la spectroscopie traite le spectre de la lumière du soleil qui est diffusément réfléchi par des matériaux sur la surface de la Terre.

Des instruments sont appelés spectromètres (ou spectroradiomètres) sont utilisés pour effectuer des mesures au sol. En utilisant des centaines voire des milliers de détecteurs, les spectromètres peuvent effectuer des mesures spectrales de bandes en tant que étroites $0.01\mu m$ sur une large plage de longueur d'onde, typiquement au moins de 0.4 à $2.4\mu m$ dans les applications de télédétection. Pour d'autres

applications, telle que l'astronomie, différentes gammes de longueurs d'onde peuvent être utilisées (par exemple, les gammes de rayons UV, visibles, infrarouges et radio).

Capteur	Organisation	Pays	Nombre de bandes	Plage de longueur d'onde (μm)
AVIRIS	NASA	États-Unis	224	0.4 -2.5
HYMAP	Integrated Spectronics Pty Ltd.	Australie	128	0.4-2.45
CASI	Itres Research	Canada	288	0.43-0.87
AISA	Spectral Imaging Ltd.	Finlande	286	0.45-0.9
DAIS 2115	CER Corp.	États-Unis	211	0.4-12.0
PROBE-1	Earth Search Science Inc.	États-Unis	128	0.4-2.45

Tableau 1.1 – Les spectro-imageurs commerciales et de recherche [Pons, 2014]

3.3 Signature spectrale

La propriété fondamentale que nous voulons obtenir est la réflectance spectrale λ : le rapport de l'énergie réfléchie à l'énergie incidente en fonction de la longueur d'onde.

$$\text{signature spectrale } (\lambda) = \frac{\text{radiance réfléchie } (\lambda)}{\text{radiance incidente } (\lambda)} \quad (1.2)$$

Tandis que la signature spectrale est l'idée qu'un matériau possède une réflectance/émissivité spectrale qui lui discrimine par rapport aux autres matériaux. La réflectance varie avec la longueur d'onde pour la plupart des matériaux car l'énergie à certaines longueurs d'ondes est dispersée ou absorbée à différents degrés. Ces variations de réflectance sont claires lorsque l'on compare les courbes de réflectance spectrale pour différents matériaux. La Figure 1.6 montre les signatures spectrales de différents matériaux tels que : la végétation, le sol et l'eau.

3.4 Bibliothèques spectrales

Une bibliothèque spectrale constitue d'une collection des spectres de réflectance de plusieurs matériaux naturels et artificiels. Dans le cadre de différents projets de recherche, plusieurs bibliothèques sont disponibles pour un usage public. Ces bibliothèques fournissent une source de spectres de référence qui peuvent aider à l'interprétation des IHSs et IMSs. Certaines d'entre elles, sont publiquement disponibles ; parmi celles-ci :

- **Bibliothèque spectrale USGS**¹ : le laboratoire de spectroscopie d'étude géologique des États-Unis à Denver, au Colorado, a compilé une bibliothèque

1. <http://speclab.cr.usgs.gov/spectral.lib04/spectral-lib04.html>

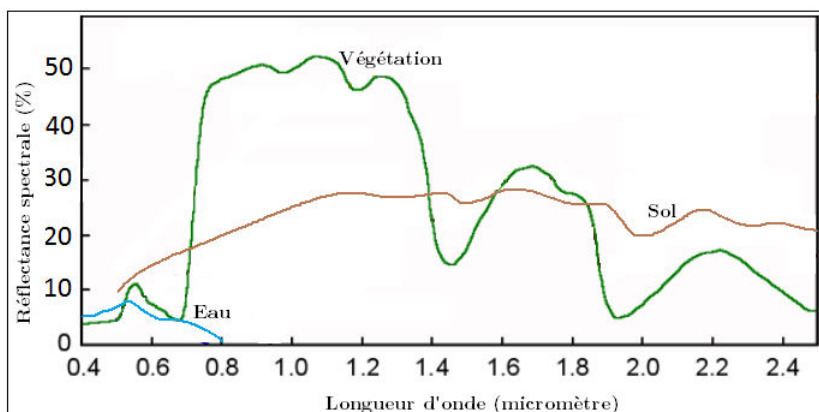


FIGURE 1.6 – Comparaison de signatures spectrales (en réflectance) correspondant à différents matériaux : en vert la végétation, en bleu l’eau et en rouge le sol.

contenant d’environ 500 spectres de réflectance des matériaux naturels (minéraux et quelques plantes) sur une plage de longueur d’onde de 0.2 à 3.0 μm .

- **Bibliothèque spectrale ASTER**² : elle a été mise à disposition par la NASA. Elle contient d’environ 2000 spectres de différents matériaux tels que : les minéraux, les roches, les sols, les matériaux artificiels, l’eau et la neige. La plupart de spectres couvrent une longueur d’onde de 0.4 à 14 μm .

3.5 Modes de représentation

Dans cette section, nous présentons différentes représentations liées à l’IHS. Nous distinguons trois types de représentations tels que : la représentation image, la représentation dans un espace spectral et la représentation dans un espace vecteur [Landgrebe, 1999].

3.5.1 Représentation image

Cette représentation permet d’afficher l’IHS sous la forme d’une image couleur RVB. De plus, elle est très riche d’un point de vue spatial puisqu’elle préserve l’information contextuelle de l’image originale, notamment les relations spatiales entre les pixels, telles que : les relations topologiques, de distance et de direction. Cependant, dans un point de vue spectral, l’information spectrale contenue dans une IHS est complètement absente dans ce mode de représentation (Figure 4.1).

3.5.2 Représentation spectrale

Dans cette représentation, l’information spectrale est perceptible et analysable. Cependant, l’information contextuelle est complètement absente. Habituellement, la représentation spectrale permet de représenter un spectre d’un ou plusieurs pixels

2. <http://speclib.jpl.nasa.gov>

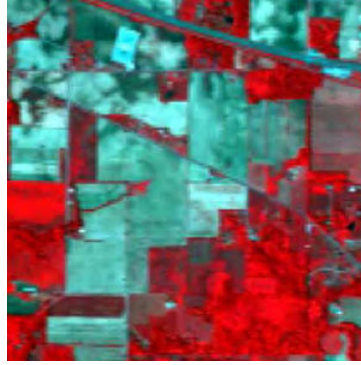


FIGURE 1.7 – La représentation image, la plus intuitive. Un exemple d’une composition colorée faite à partir de trois bandes d’une IHS (bande 50, bande 27 et bande 17).

en fonction de la longueur d’onde de rayonnement magnétique (Figure 1.6). Alors que, si la signature spectrale (radiance/longueur d’onde) transmet efficacement des informations nécessaires pour identifier le contenu d’un pixel, cela fournit une simplicité fondamentale d’un point de vue de traitement. Dans ce cas, les pixels peuvent être traités une seule fois. Il est intrinsèquement plus adapté la représentation quantitative des données.

La signature spectrale a une caractéristique très utile puisqu’elle fournit des informations spectrales souvent interprétables directement. Étant donné un détail spectral de haut degré, les caractéristiques de la réponse de pixel peuvent être liées aux propriétés physiques du contenu du pixel. Par exemple, nous pouvons déterminer facilement si un pixel contient du sol, de la végétation ou de l’eau. Ainsi, pour l’analyste, un affichage de la signature spectrale peut fournir des informations pertinentes aux propriétés physiques. Pour cette raison, les scientifiques s’orientent leur réflexion sur les données hyperspectrales du point de vue de l’espace spectral. Mais, l’analyse d’une signature spectrale pour un pixel individuel ne fournit pas toute la connaissance concernant la relation entre la réponse spectrale et l’information disponible sur la scène. La réponse spectrale pour chaque type de couverture de la surface de la Terre varie de manière caractéristique. Par exemple, la réponse spectrale des pixels du champ de maïs à un moment donné, n’est pas uniforme pour tous les pixels du champ, mais varie de manière caractéristique par rapport à une certaine valeur moyenne. Ceci est dû aux relations entre la taille et le mélange des feuilles, des tiges, du fond du sol, etc.

3.5.3 Espace vecteur

L’espace vecteur permet de réaliser une analyse spectrale fine des pixels de l’IHS, mais la notion de l’information contextuelle, notamment le voisinage spatial est absent dans ce type de représentation. Formellement, étant donné une IHS $X = [x_1, \dots, x_N] \in \mathbb{R}^{M \times N}$, avec x_i représente un vecteur d’une réponse spectrale

d'un pixel de l'IHS, M est le nombre de bandes spectrales et N le nombre de pixels. Chaque réponse spectrale a M valeurs de longueur d'onde λ . Chaque vecteur du pixel x_i est représenté dans un espace de M dimensions. Ceci est une manière utile et particulière pour représenter les réponses spectrales d'une IHS. Chaque vecteur de M dimensions représente toutes les informations spectrales liées à chaque pixel de l'image. Comme le montre la Figure 1.8, les pixels sont représentés dans un espace de 2 dimensions (Bande 1 et Bande 2). Cette représentation est considérée comme

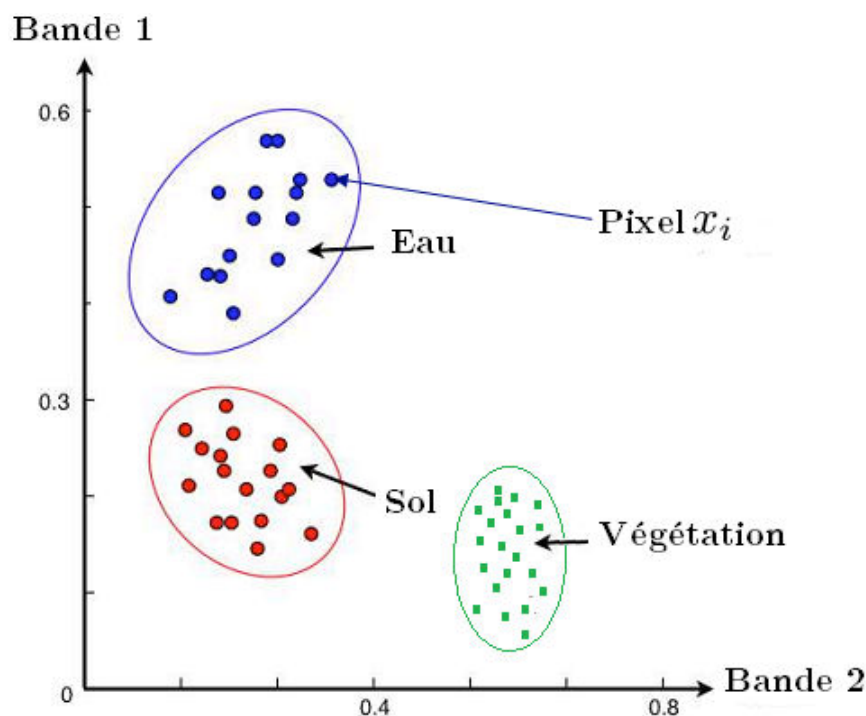


FIGURE 1.8 – Regroupement des pixels similaires dans un espace vecteur

étant une manière quantitative permettant non pas seulement de représenter les valeurs numériques de chaque pixel, mais aussi de représenter la variations des matériaux par rapport à leur valeur moyenne ou centrale. Cependant, la représentation image montre les relations spatiales entre les réponses spectrales dans un point de vue géographique et elle fournit une manière d'associer chaque pixel avec une localisation spatiale bien précise. Elle permet aussi de fournir quelques informations supplémentaires et utiles pour l'analyse d'IHS. L'espace spectral permet de relier souvent les types de matériaux avec les réponses spectrales.

4 Processus général d'analyse et d'interprétation d'IHS

Usuellement, les approches d'analyse et d'interprétation d'IHS, nous permet, à travers un processus général, à la génération des carte thématiques d'occupation des sols de la surface terrestre (Figure 1.9).

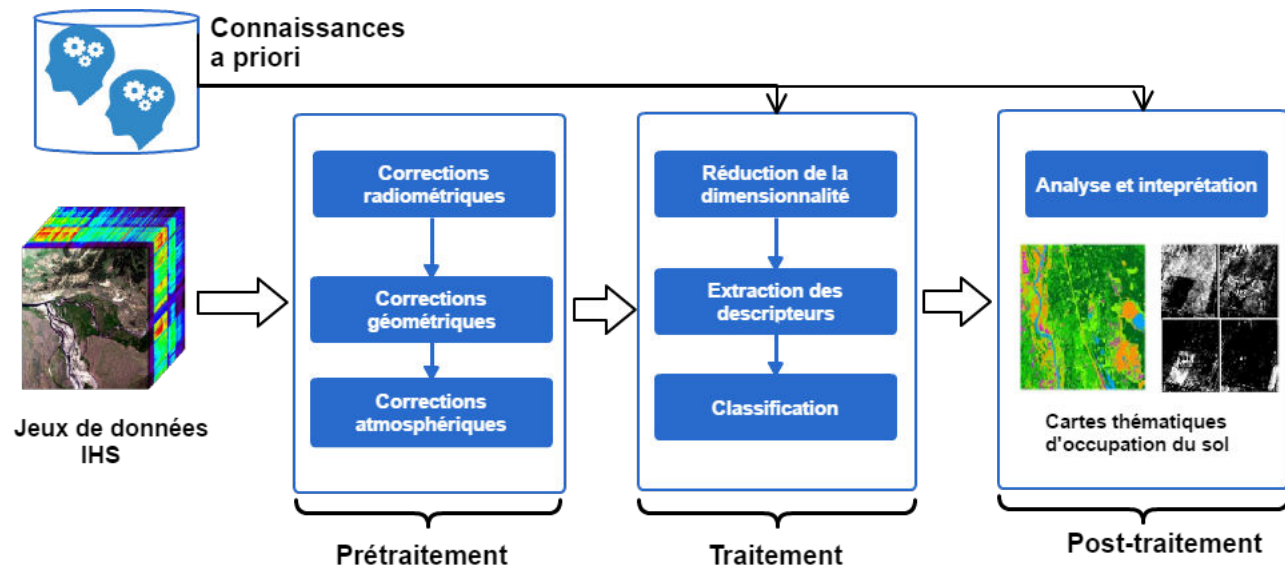


FIGURE 1.9 – Illustration d'un processus typique d'analyse et d'interprétation d'IHS

Le processus d'analyse et d'interprétation d'IHS présenté ci-dessus, comprend principalement les différentes phases qui peuvent être appliquées à une IHS avant de pouvoir l'interpréter :

- **Phase de pré-traitement** : elle vise à traiter et corriger les différents erreurs de l'IHS dus à l'acquisition à savoir les facteurs de l'atmosphère, les technologies des capteurs, etc. En raison de ces facteurs, plusieurs opérations de correction de l'IHS peuvent être appliquées. Cette phase n'est pas toujours nécessaire si les conditions d'acquisition de l'IHS sont parfaites. Nous présentons dans la section suivante plus de détails sur ces techniques de correction de l'IHS.
- **Phase de traitement** : à partir de l'IHS, un processus de réduction de dimension est nécessaire. En effet, la réduction de dimension est une étape préliminaire et primordiale pour l'analyse de l'IHS. Elle vise à réduire la haute dimensionnalité ainsi d'éviter toute sorte de redondance dans les bandes spectrales afin de mettre en ouvre facilement l'analyse et l'interprétation de l'IHS. Dans un deuxième temps, différentes caractéristiques seront retenues par les techniques de réduction de dimension pour être exploitées comme des descripteurs dans l'étape de classification. Ceci permet d'avoir une carte thématique de l'occupation des sols.
- **Phase de post-traitement** (analyse et interprétation) : Elle a pour objectif d'interpréter les données extraites et classifiées dans la phase précédente. Elle exploite généralement toutes les informations extraites dans la phase de traitement avec des connaissances *a priori*, qui vont être utiles pour aboutir à une bonne interprétation. Nous présenterons dans ce qui suit les différentes phases de l'analyse et l'interprétation d'IHS à savoir les opérations de correc-

tions et un état de l'art sur les différentes techniques de classification d'IHS.

5 Prétraitements

Tandis que l'IHS transmis par le satellite hyperspectral est brute. Pour cela, il faut effectuer un certain nombre de corrections pour la rendre exploitable et interprétable. De nombreuses opérations de prétraitements ont été développées, nous citons : les corrections radiométriques, géométriques et atmosphériques.

5.1 Corrections radiométriques

Les données hyperspectrales contiennent généralement des erreurs radiométriques causées par les capteurs et l'atmosphère, notamment l'apparition d'un bruit dû à l'atténuation du signal et le dysfonctionnement des capteurs, ce qui entraîne des perturbations d'interprétation visuelle et dégradations des résultats des traitements (classification, identification d'objets, etc). Ceci donne lieu à effectuer des corrections radiométriques pour atténuer ces altérations et réduire le bruit. Dans un cadre d'analyse d'IHS, il est habituellement désirable de corriger les bandes bruitées. Généralement, ces corrections tentent à changer la valeur radiométrique de quelques pixels par celle la plus proche de la vérité terrain.

5.2 Corrections géométriques

Ces opérations de traitements permettent de corriger l'image acquise des déformations systématiques et distorsions géométriques engendrés par les conditions de prise de l'image (distorsions panoramiques, caractéristiques et position du capteur, forme géométrique de la terre) en faisant correspondance à chaque pixel de l'image différentes couches d'information tout en ayant des informations sur les coordonnées géographiques. Ceci permet de corriger l'exactitude entre les coordonnées de localisation des objets de l'image satellitale acquise et les coordonnées spatiales des objets de la vérité terrain.

5.3 Corrections atmosphériques

La procédure de correction atmosphérique permet d'éliminer les effets de l'atmosphère. La valeur enregistrée à n'importe quel emplacement de l'image acquise ne représente pas le vrai rayonnement. En effet, une partie de la luminosité qui arrive sur la cible est due à la diffusion du soleil et la réflexion de la cible et le reste est dérivé de la luminosité de l'atmosphère elle-même. Bien que la séparation des contributions n'est pas connue a priori, l'objectif de la correction atmosphérique est de quantifier ces deux composantes. À cet égard, l'analyse peut être basée sur la réflectance de la cible ou bien sur les valeurs de la radiance.

6 La classification

La classification est un domaine de recherche fondamental et très important dans le traitement d'IHS [Richards and Richards, 1999]. Elle est considérée comme un moyen important pour l'acquisition des informations, dont le but est de classer automatiquement les types de couverture terrestres de tous les pixels de l'image. Elle permet alors d'extraire des informations utiles pour générer des cartes thématiques d'une IHS. En effet, la carte thématique peut clairement refléter la répartition spatiale des objets terrestres. Dans un contexte d'analyse d'IHS, les algorithmes de classifications peuvent être répertoriés comme classification supervisée et non supervisée, classification paramétrique et non paramétrique (voir Figure 1.10). Nous nous intéressons dans cette section à présenter les approches de classification supervisée et non supervisée.

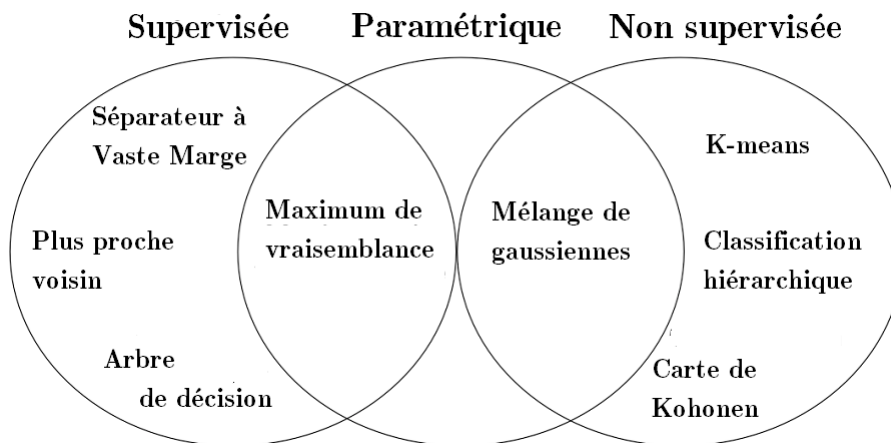


FIGURE 1.10 – Catégorisation des techniques de classification

6.1 Principe

Étant donné un ensemble des pixels $x_i, i \in \{1, \dots, N\}$ d'une IHS X , l'objectif est de produire des cartes thématiques en divisant l'ensemble des pixels en un nombre K de partitions C représentant de manière schématique différentes régions élémentaires d'une scène. Chaque partition représente un groupe de pixels similaires c'est à dire ayant des propriétés communes et qui identifie une classe d'une scène (végétation, sol, eau, etc.). Dans un contexte hyperspectral, la similarité entre pixel est double : spectrale et/ou spatiale. Usuellement, la qualité et le taux de classification doivent satisfaire certaines conditions relatives à l'utilisateur et les données : la pertinence liée à l'utilisateur tandis que l'exhaustivité et la séparabilité liées aux données.

- Pertinence : les classes doivent fournir des informations utiles pour l'utilisateur.
- Exhaustivité : tous les pixels doivent être classifiés.
- Séparabilité : chaque pixel doit être associé à une seule classe.

6.2 Classification supervisée et non supervisée

La classification la plus commune est de sélectionner des échantillons d'apprentissage en fonction des connaissances a priori afin que l'image connue au préalable soit classifiée. Dans ce cadre, les méthodes de classification peuvent être supervisées ou non supervisées.

6.2.1 Classification supervisée

Avant de réaliser la classification supervisée, il est nécessaire de sélectionner les échantillons d'apprentissage représentatifs pour chaque type d'objet terrestre à travers des mesures effectuées sur le champs ou bien sur la carte de vérité terrain. Ces mesures peuvent être statistiques, notamment l'écart type, la matrice de covariance, etc. Les échantillons d'apprentissage sont classifiés selon une fonction de discrimination de classification avec ces critères statistiques. L'approche de classification supervisée utilise ces échantillons d'apprentissage pour réaliser un classifieur.

Dans la littérature, il existe plusieurs approches de classification supervisée. Des approches qui utilisent des fonctions de discriminations qui minimisent le taux d'erreur de classification. Parmi ces approches : arbres de décision [Wang et al., 2012], fonctions linéaires [Teke and Sakarya, 2013], k plus-proches-voisins (KPP) [Ma et al., 2010], etc... D'autres approches comme le maximum de vraisemblance utilisent des critères statistiques, notamment la variance et la covariance des échantillons d'apprentissage pour évaluer la probabilité d'association d'un pixel x_i à une classe C_j . Le classifieur du maximum de vraisemblance est considéré à la fois comme étant une approche de classification supervisée et paramétrique. Tandis que parmi les approches non paramétriques et supervisées le procédé le plus utilisé est celui des k plus proches voisins [Ma et al., 2010]. Dans cette approche, chaque pixel x_i est associé à une classe majoritaire C_j qui est la plus proche dans le nuage de points, avec $C_j \in k$ échantillons d'apprentissage. Cependant, ces approches classiques ne supportent pas la grande dimensionnalité de l'IHS. Pour cela, il faut avoir un nombre très élevé d'échantillons d'apprentissage. Pour surmonter ce problème, d'autres approches de classification basée sur les noyaux ont été développées [Kuo et al., 2014]-[Gu et al., 2014]-[Chen et al., 2016]. Ces approches ont été commencé avec les Séparateurs à Vaste Marge (SVM) [Camps-Valls and Bruzzone, 2005]-[Suykens et al., 2002]. En effet, en se basant sur les noyaux, SVM vise à trouver un hyperplan à marge maximale pour séparer les échantillons d'apprentissages dans l'espace de Hilbert de grande dimension. Cet espace de grande dimension, voire infinie, permet aux échantillons d'apprentissage d'être linéairement séparable. SVM donne des résultats satisfaisants même le nombre d'échantillons est limité. Cependant, le paramétrage de l'algorithme et le choix du noyau adapté restent parmi les difficultés d'utilisation de SVM avec les noyaux. D'autres approches de classification supervisée non paramétrique populaires se focalisent sur les réseaux de neurones (*Multi-Layer Perceptron*) [Benediktsson et al., 1990]-[Kröse et al., 1993]. Ce perceptron multi-couche est constitué de plusieurs couches, au minimum de 3 couches de neurones. Un al-

gorithme d'apprentissage par rétropropagation du gradient a été largement utilisé avec ce réseau afin de paramétrer les poids associés à ces neurones. Ce dernier vise à trouver les poids optimaux en propageant progressivement l'erreur quadratique sur les échantillons d'apprentissage.

6.2.2 Classification non supervisée

Dans la classification non supervisée, la connaissance a priori est absente (échantillons d'apprentissage, nombre de classe), uniquement selon les caractéristiques des données d'image, elle utilise la similarité et la dissimilarité des caractéristiques spectrales des objets terrestres représentées par ces données hyperspectrales, de sorte que la classification non supervisée est également appelée *clustering*. Dans ce cadre, il existe principalement deux grandes catégories : la classification ascendante ou descendante hiérarchique et la classification par partitionnement. La première catégorie a pour but de générer une suite des partitions des observations dans un sens ascendant ou descendant par fusion ou division. Plusieurs critères sont utilisés pour la division ou la fusion des classes, notamment la distance euclidienne, angle spectral, etc. Ces partitions allant d'une partition monoclasse à une partition multi-classes où chaque observation représente une classe (Figure 1.11).

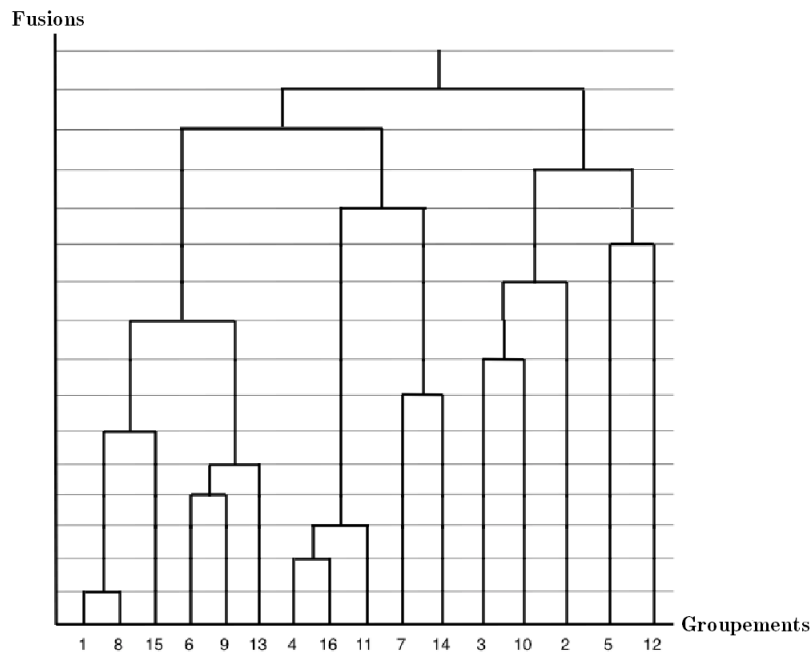


FIGURE 1.11 – Classification ascendante hiérarchique basée sur la succession de fusion

La matrice par factorisation non négative (NMF) a été utilisée pour diviser un ensemble de clusters de l'IHS [Gillis et al., 2015]. Ceci permet de stabiliser et équilibrer les clusters dans un processus de partitionnement et d'améliorer le taux

de précision de la classification. Dans un contexte de segmentation et classification d'IHS, une approche de segmentation hiérarchique récursive a été proposée [Tarabalka et al., 2012]. En effet, elle permet de produire une segmentation multi-résolution afin d'améliorer le taux de classification d'IHS en utilisant un modèle hybride entre la technique de croissance de région et le clustering spectral. Cet algorithme de segmentation tente à associer chaque région étudiée à un niveau de segmentation optimal en associant la région de niveau plus fin à une région de niveau plus grossier.

Parallèlement à ces approches dites de partitionnement, les approches de *clustering* [Elavarasi et al., 2011] tentent à diviser l'espace de représentation en un nombre fini de K groupes (clusters) disjoints en calculant le centre de chaque classe. L'estimation du nombre de groupes reste la question centrale. La méthode de clustering K-means [Kanungo et al., 2002] vise à diviser un ensemble de pixels $\{x_1, \dots, x_N\}$ de l'IHS X en un ensemble de K classes $\{C_1, \dots, C_K\}$ en minimisant la somme quadratique de la distance entre les pixels et le centre de classe de chaque cluster. L'algorithme c-moyens flous non supervisé FCM (*Fuzzy C-Means*) [Hung et al., 2011] du l'algorithme classique FCM [Groenen and Jajuga, 2001] est proposé pour améliorer la stabilité et la précision du FCM standard dans un contexte de classification des données de haute dimension, notamment l'IHS. Il est basé sur la pondération des centres des classes et l'extraction des pixels pertinents par une analyse discriminante. Une approche de partitionnement appelée propagation d'affinité AP (*Affinity Propagation*) a été proposée [Frey and Dueck, 2007]. Cette approche permet de déterminer les pixels représentants en se basant sur la notion d'ensemble pour produire des classes de manière déterministe. Elle peut être réalisée dans le cas de classification supervisée ou non supervisée, c'est à dire soit le nombre de classes K à trouver est connu ou n'est pas connu au préalable. L'algorithme itératif ISODATA (*Iterative self-organizing data analysis*) [Ghamisi et al., 2014, Tarabalka et al., 2009] a été présenté dans un contexte de classification non supervisée d'IHS. Dans une première étape, il tente à fournir un *clustering* initial. Dans une deuxième étape, il adopte une méthode itérative pour modifier et ajuster le regroupement, de manière à aboutir un *clustering* correct.

6.3 Classification paramétrique et non paramétrique

Dans une autre perspective, les approches de classification peuvent être divisées en paramétriques et non paramétriques. Par exemple, le classificateur de maximum par vraisemblance supervisé MLC (*Maximum Likelihood Classifier*) est largement utilisé dans le contexte paramétrique. De ce cadre, le MLC repose sur l'hypothèse que la fonction de densité de probabilité pour chaque classe est régie par la distribution gaussienne [Fukunaga, 2013]. En revanche, les méthodes non paramétriques ne sont pas contraintes par des hypothèses sur la distribution des données d'entrée. Par conséquent, des techniques telles que les SVM, les réseaux de neurones, les arbres de décision et les approches d'ensemble peuvent être appliquées, même si les densités conditionnelles classiques ne sont pas connues ou ne peuvent être estimées de

manière fiable [Benediktsson and Ghamisi, 2015]. De ce fait, pour les données hyperspectrales avec un nombre limité d'échantillons d'apprentissage, telles techniques peuvent conduire à des résultats de classification plus précis.

6.4 Classification spectro-spatiale

Les approches décrites dans la section précédente se focalisent sur la classification au niveau pixel en associant chaque pixel à une classe en analysant uniquement ses caractéristiques spectrales, sans tenir compte l'information spatiale, notamment l'information sur le voisinage entre les pixels. Usuellement, les matériaux de certaines classes se rassemblent ensemble. Étant donné un pixel donné x_i , s'il appartient à une classe C_j , les pixels voisins ont une forte probabilité d'appartenir à la même classe C_j . En outre, certaines structures comme les contours peuvent être extraites à partir de l'information spatiale. De ce fait, la combinaison de l'information spectrale avec l'information spatiale peuvent améliorer les résultats de classification d'IHS. Ces approches tentent d'associer chaque pixel de l'IHS à une classe en analysant ses caractéristiques spectrales et spatiales. Dans ce cadre, une classification spectro-spatiale basée sur les profils morphologiques a été présentée [Benediktsson et al., 2003]. Afin d'intégrer l'information spatiale, des opérations d'ouverture et de fermeture sont exploitées. D'autres approches basées sur les profils morphologiques ont été proposée pour le même objectif de classification spectro-spatiale d'IHS : les transformations morphologiques étendues [Plaza et al., 2005], les profils morphologiques étendus [Gu et al., 2016] basés sur le framework d'apprentissage par noyau multiples MKL, les profils morphologiques combinés avec plusieurs classifieurs (méthode d'ensemble) [Bao et al., 2016] et les profils morphologiques combinés avec les techniques d'ondelettes [Quesada-Barriuso et al., 2014]. Dans le même contexte, une approche basée sur la combinaison des champs de Markov avec SVM a été présentée [Tarabalka et al., 2010b]. Les champs de Markov constituent une technique probabiliste permettant de modéliser les interactions entre les pixels d'une IHS dans un contexte spatial. Ils construisent un champ (appelée *Markovien*) de variables aléatoires où chaque variable est associée à un pixel de l'image muni d'un espace de voisinage. Par transitivité, les interactions spatiales entre les pixels sont considérées à la fois globalement et localement. D'autres approches ont exploité les champs de Markov pour la classification spectro-spatiale : Un sous-espace projeté avec SVM est combiné avec les champs des Markov multi-niveaux [Yu et al., 2016], des champs de Markov sont utilisés avec le modèle d'apprentissage actif AL (*Active Learning*) [Sun et al., 2015a]. Des techniques de segmentation sont intégrées aussi dans le processus de classification spectro-spatiale. En effet, plusieurs algorithmes de segmentation sont combinés avec la méthode de vote majoritaire pour classifier chaque pixel de l'IHS [Tarabalka et al., 2010a]. En se basant sur un modèle bayésien, une carte de segmentation hiérarchique a été combinée avec les champs de Markov pour classifier les pixels de l'IHS en exploitant les deux informations spectrales et spatiales [Golipour et al., 2016].

6.5 Progrès récents

Dans cette section, nous énumérons quelques tendances nouvelles en classification spectro-spatiales d'IHS. Actuellement et dans ce contexte, des travaux sont basés sur une récente notion l'apprentissage profond (*Deep Learning*) ont été développés. Un nouvel algorithme d'apprentissage profond basé sur des connaissances spectrales et le réseau de neurone a été développé [Zhou et al., 2017]. Les réseaux de neurones sont utilisés dans cet algorithme pour classifier les pixels de l'IHS en exploitant des informations spectro-spatiales (Fig. 1.12). Le réseau spectro-spatial

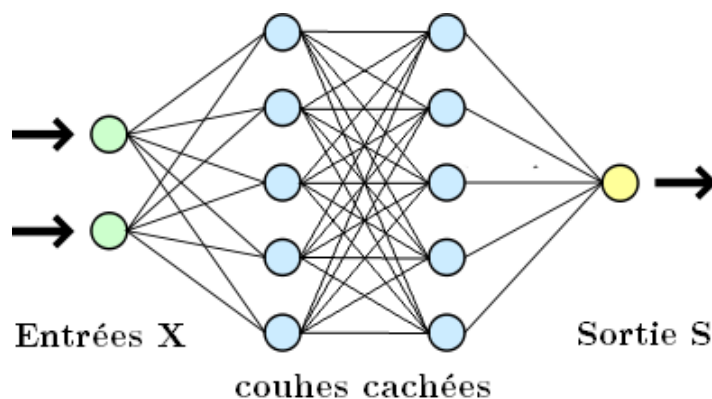


FIGURE 1.12 – Architecture classique d'un réseau multicouche

non linéaire est développé en se basant sur la structure de réseau de l'analyse en composantes principales PCA (*Principal Component Analysis*) [Pan et al., 2016]. L'information non linéaire a été utilisée dans cette approche pour générer des caractéristiques discriminantes. Un nouveau framework qui combine les techniques de réduction de dimension avec les techniques d'apprentissage profond pour l'extraction des caractéristiques spectro-spatiales a été proposé [Zhong et al., 2017]. Dans ce framework, un algorithme de discrimination local est développé pour extraire des caractéristiques spectrales à partir de l'IHS. Tandis que le réseau de neurone convolutif est utilisé pour extraire des caractéristiques spatiales à des niveaux élevés. Le modèle d'apprentissage actif semi-supervisé a été développé en exploitant à la fois les techniques d'apprentissage actif et d'apprentissage semi-supervisé [Wang et al., 2017]. Ceci permet de découvrir la représentativité et le discriminabilité des caractéristiques de l'IHS. En effet, les algorithmes de *clustering* semi-supervisé permettent de découvrir les informations structurelles améliorer la qualité des échantillons d'apprentissage.

7 De l'analyse aux défis de la grande dimensionnalité d'IHS

Parmi les majeurs défis de l'exploitation et l'analyse d'IHS est la haute dimensionnalité de cube de données fournie par les capteurs hyperspectraux. Nous citons

principalement :

- **La forte corrélation inter-bandes** : les pixels constituent de centaines de bandes et les cubes de données sont typiquement centaines de mégabytes, le nombre de données à analyser peut être écrasant. Autrement, tandis que les bandes spectrales sont étroites et contiguës, il est fortement possible d'avoir une forte corrélation entre les bandes contiguës. Ceci entraîne une classification imprécise d'IHS.
- **La malédiction de dimensionnalité** : des travaux ont signalé des propriétés géométriques, statistiques et asymptotiques distinctives de données de haute dimension à travers quelques des résultats expérimentaux [Jimenez and Landgrebe, 1998]. D'après ces expérimentations, nous distinguons que :
 - L'espace à grande dimension est presque vide, ce qui implique que les données multivariées dans R sont habituellement dans une structure dimensionnelle inférieure. En d'autres termes, les données de haute dimension peuvent être projetées dans un sous-espace inférieur sans perdre l'information pertinente en termes de séparabilité de classes [Benediktsson and Ghamisi, 2015].
 - Des travaux [Keinosuke, 1990] ont montré qu'il existe une relation entre le nombre d'échantillons d'apprentissage et le nombre de dimensions requis pour les différents classifieurs. Tandis que le nombre requis d'échantillons d'apprentissage est linéairement relatif à la dimensionnalité pour les classifieurs linéaires et au carré de la dimensionnalité pour les classifieurs quadratiques.
 - Landgrebe [Landgrebe, 2005] a montré que plusieurs bandes spectrales sont inutiles et indésirables en terme de précision de classification. Quand le nombre de bandes augmente, avec un nombre constant d'échantillons d'apprentissage, un ensemble statistique de grande dimension doit être estimé. En d'autres termes, bien que des dimensions spectrales plus élevées augmentent la séparabilité des classes, la précision de l'estimation statistique diminue. Cela entraîne une diminution de la précision des classification au-delà d'un certain nombre de bandes. Dans le contexte de classification, ces problèmes sont liés au problème de malédiction de la dimensionnalité (*curse of dimensionality*).

Bien que la dimensionnalité d'IHS augmente, plus d'informations peuvent être détectées pour aboutir à une précision de classification plus élevée. Dans le même temps, les travaux existants démontrent que les techniques classiques développées pour les données multispectrales ne peuvent pas être robuste et flexible pour les données hyperspectrales.

Les problèmes susmentionnés liés à la haute dimensionnalité des données ont une influence sur les techniques de classification supervisée [Qian et al., 2009]. Ces techniques exigent un grand nombre d'échantillons d'apprentissage (ce qui est presque impossible à obtenir dans la pratique) pour faire une estimation précise. Ce problème est encore plus sévère lorsque

la dimensionnalité augmente. De ce fait, les approches de classification développées sur les données hyperspectrales doivent être capables de traiter des données de grande dimension lorsque le nombre d'échantillons d'apprentissage disponible est très limité.

- **L'incertitude** : usuellement, elle est générée dans les phases d'acquisition et de classification des données. Cela peut influencer considérablement sur la précision de la classification et sur la qualité de la carte thématique de classification obtenue [Canter, 1997, Wang, 2015]. Il existe de nombreuses raisons pour de telles incertitudes, y compris les conditions atmosphériques au moment de l'acquisition des données, la limitation des données en termes de résolutions radiométriques et spatiales, la mosaïque de plusieurs images et bien d'autres. L'enregistrement d'IHS et la correction géométrique provoquent une incertitude de position. En outre, les erreurs algorithmiques au moment de l'étalonnage des effets atmosphériques ou topographiques peuvent entraîner des incertitudes radiométriques [Lu and Weng, 2007].
- **Influence de la résolution spatiale** : La précision de la classification peut être fortement influencée par la résolution spatiale des IHSs. La haute résolution spatiale peut réduire le problème des pixels mixtes et détecter plus de détails de la scène. Dans [Woodcock and Strahler, 1987], la précision de classification a été considérée comme le résultat du compromis entre deux aspects. Le premier se réfère à l'influence des pixels frontières sur les résultats de la classification. Dans ce cas, tandis que la résolution spatiale devient plus fine, le nombre de pixels frontières de différents objets diminuera. Le deuxième aspect se réfère à l'augmentation de la variance spectrale de différentes couvertures terrestres associées à une résolution spatiale plus fine. Lorsque les données optiques se traitent avec résolution spatiale basse ou moyenne, l'existence de nombreux pixels mixtes entre différentes classes de couverture terrestre est la principale source d'incertitude et peut influencer sur la précision de la classification. La résolution spatiale fine peut fournir des informations détaillées sur la forme et la structure des différentes couvertures terrestres. De telles informations peuvent également être transmises au système de classification afin d'augmenter encore le taux de précision de la classification et d'améliorer la qualité des cartes thématiques de classification. La prise en compte de l'information spatiale dans le système de classification est un sujet de recherche dynamique dans la communauté hyperspectrale, et il a été étudié dans de nombreux travaux tels que [Benediktsson and Ghamisi, 2015, Ghamisi et al., 2015, Zhu et al., 2015]. L'utilisation d'IHSs à haute résolution spatiale présente de nouveaux problèmes, en particulier ceux provoqués par des ombres, ce qui entraîne des variations spectrales élevées dans la même classe de couverture terrestre. Ces inconvénients peuvent réduire la précision de la classification si les classificateurs et les techniques de réduction de dimension ne peuvent pas gérer efficacement ces contraintes.

8 Conclusion

Dans ce chapitre, nous avons présenté, dans un premier lieu, les différents aspects liés à la télédétection. Dans un second lieu, nous avons focalisé sur la technologie hyperspectrale et son processus général d'analyse et d'interprétation. Nous nous sommes intéressées aussi sur les différentes techniques de classification d'IHS. Cependant, nous trouvons, d'après cet état de l'art, que de beaucoup d'enjeux liés à l'exploitation et l'analyse d'IHS, notamment la grande dimensionnalité. Par conséquent, nous avons dégagé les différents problèmes relatifs à la haute dimensionnalité, tels que : la forte corrélation inter-bandes, la malédiction de dimensionnalité, l'incertitude, etc. Partant de ces remarques, le prochain chapitre présente un état de l'art sur les différentes techniques de la réduction de dimension en IHS et présente une étude comparative entre ces techniques afin de mettre en relief la contribution de notre travail de recherche.

La réduction de dimensionnalité en imagerie hyperspectrale : un état de l'art

Sommaire

1	Introduction	36
2	Représentation des données hyperspectrales pour la RD . .	36
3	La RD en IHS	36
4	La RD basée sur les approches de projection	37
4.1	Approches par projection linéaire	38
4.2	Approches par projection non linéaire	40
5	La sélection de bandes spectrales	41
5.1	Approches par sélection supervisée	43
5.2	Approches par sélection non supervisée	45
5.3	Approches par sélection semi-supervisée	46
6	Estimation du nombre de bandes/composantes à conserver	48
6.1	Le critère de Kaiser	49
6.2	Pourcentage de la variance	49
6.3	Scree-test de Cattell	49
6.4	Rapport de croissance	51
6.5	La dimensionnalité virtuelle	51
7	Travaux existants : étude comparative	51
7.1	Tableau comparatif	51
7.2	Synthèse et discussions	54
8	Conclusion	54

1 Introduction

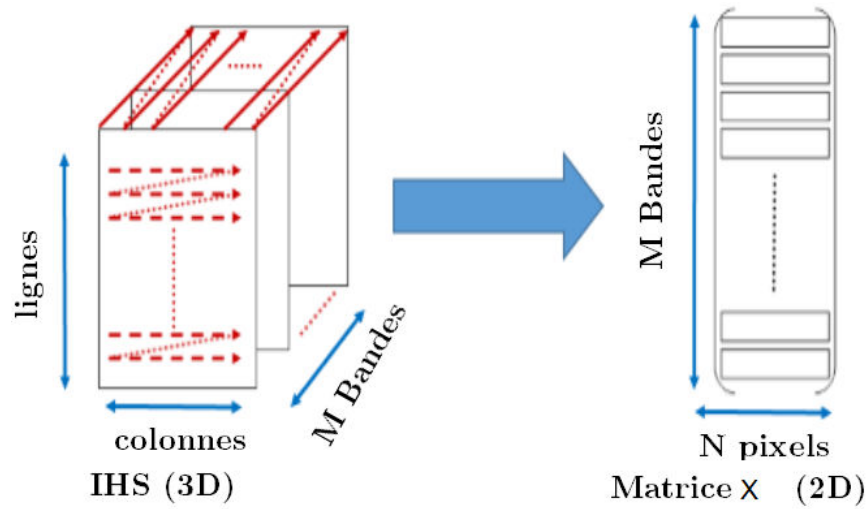
La télédétection hyperspectrale a été largement utilisée dans plusieurs domaines comme la discrimination des types d'occupations des sols et la détection des cibles à partir d'IHSs [Harsanyi and Chang, 1994, Plaza et al., 2009]. La classification d'images est considéré comme un grand défis dans plusieurs applications, par exemple, la reconnaissance militaire, le suivi de l'occupation des sols, l'évaluation des dommages, etc. Cependant, dû aux problèmes cités dans le chapitre précédent, notamment la corrélation inter-bandes, la taille massive des bandes spectrales et la complexité du calcul, la Réduction de Dimensions (RD) devient une tâche indispensable pour contrebalancer ces différents problèmes. En effet, elle permet de réduire les données massives (le nombre élevé des bandes spectrales) tout en préservant les informations utiles afin d'améliorer et faciliter l'interprétation et l'analyse d'IHSs. Dans ce cadre, nous commençons dans ce chapitre de présenter un aperçu général sur la RD. Dans un deuxième temps, nous allons détailler les différentes approches de RD en mettant l'accent sur des travaux existants. Dans une dernière section, nous finirons par une étude comparative.

2 Représentation des données hyperspectrales pour la RD

Dans le contexte de RD d'IHS, le cube d'IHS 3D (2 spatiales et 1 spectrale) est transformé en image 2D (1 spectrale et 1 spatiale), c'est à dire sous forme d'une matrice. En d'autres termes, pour appliquer les techniques de la RD, les données hyperspectrales sont considérées comme des échantillons de spectres. Formellement, étant donné une IHS_(3D) notée $I \in \mathbb{R}^{I_1 \times I_2 \times M}$, avec M est le nombre de bandes spectrales et I_1 times I_2 est la taille de chaque image de bande spectrale, c'est à dire $N = I_1 I_2$ est le nombre de pixels de l'IHS. Chaque vecteur de pixel de l'IHS est une variable aléatoire de M -dimensions. Ces vecteurs de pixels représentent les signatures spectrales de l'IHS, où ils peuvent être concaténés dans une matrice $X = [x_1, \dots, x_N] \in \mathbb{R}^{M \times N}$ (Figure 2.1). En d'autre terme, le $l^{\text{ème}}$ ligne dans X correspond au $l^{\text{ème}}$ bande spectrale de l'IHS.

3 La RD en IHS

Selon la littérature, les approches de RD peuvent être divisées en deux grandes catégories (Figure 2.2), à savoir les approches de projection et les approches de sélection des bandes [Bajcsy and Groves, 2004, Serpico and Bruzzone, 2001]. Habituellement, les techniques de projection visent à transformer (projeter) les données hyperspectrales de dimensions M dans un nouvel sous-espace réduit de taille p dimensions, avec M représente le nombre de bandes spectrales et p la nouvelle dimension de l'image réduit ($p \ll M$). Les techniques de projection visent à réduire la taille originale d'IHSs en trouvant un sous-espace de projection en maximisant la

FIGURE 2.1 – Représentation d'une IHS : $I_{(3D)}$ en matrice $X_{(2D)}$

variance, minimisant la corrélation, etc. Elles reposent essentiellement sur des critères statistiques permettant de trouver la meilleure projection. La transformation par ces techniques de projection est réalisée par des combinaisons linéaires ou non linéaires des bandes spectrales originales. En outre, les techniques de la sélection de bandes visent à sélectionner un sous-ensemble des bandes spectrales p à partir des données originales. Selon la disponibilité de la connaissance *a priori*, ces techniques de la sélection de bandes peuvent être regroupées en deux familles, à savoir la sélection supervisée et la sélection non supervisée des bandes spectrales. Nous détaillons dans la section suivante ces différentes techniques de la RD.

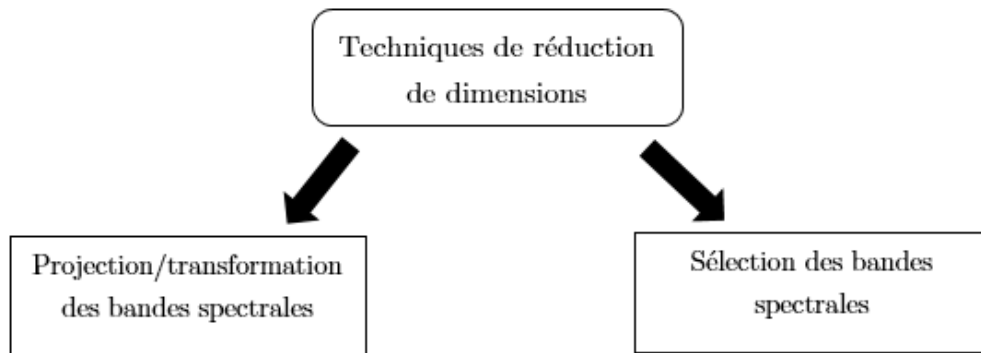


FIGURE 2.2 – Les différentes catégories des techniques de RD d'IHS

4 La RD basée sur les approches de projection

Comme mentionné précédemment, la projection des bandes spectrales vise à trouver un sous-espace optimal en faisant une combinaison linéaire ou non linéaire

des bandes spectrales. Formellement, étant donné une IHS $X = [x_1, \dots, x_N] \in \mathbb{R}^{M \times N}$, avec x_i est un vecteur pixel. M est le nombre de bandes spectrales et N représente le nombre des pixels $i \in \{1, 2, \dots, N\}$. La nouvelle image réduite Y de taille p est donnée comme suit :

$$Y = f(X) \quad (2.1)$$

avec f est une fonction de transformation linéaire/non linéaire et R représente le vecteur d'observation spectral de l'IHS. Ces techniques de projection ont l'avantage de tenir compte de toutes les bandes spectrales originales de l'IHS tout en essayant de préserver le maximum d'informations dans le nouvel sous-espace de dimension inférieure. Usuellement, la projection tente de chercher une meilleure représentation et projection des données originales en maximisant la décorrélation de ces données (Figure 2.3).

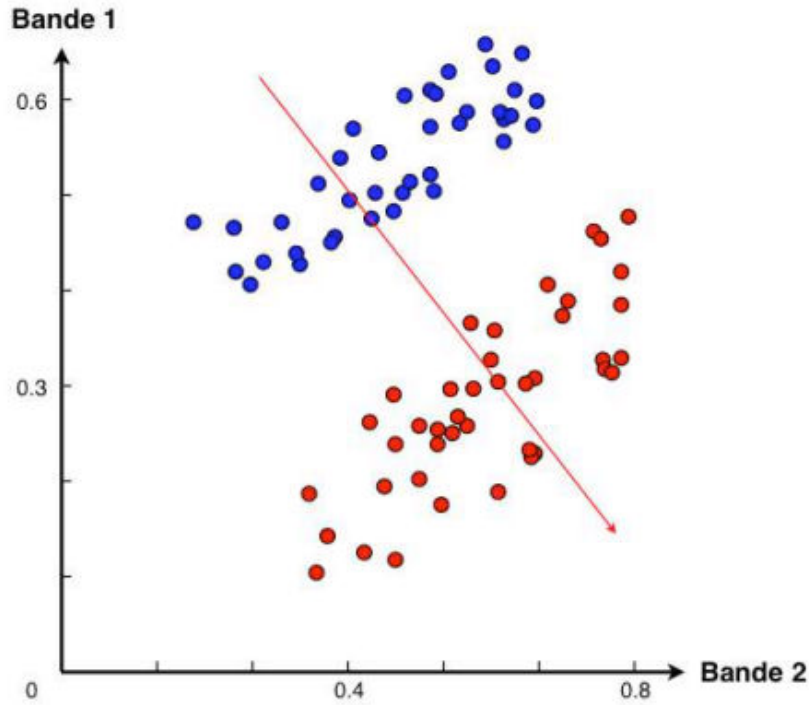


FIGURE 2.3 – Exemple de projection de deux bandes spectrales : l'axe en rouge permet de séparer les nuages inter-pixels pour discriminer deux classes différentes

4.1 Approches par projection linéaire

Usuellement, Les techniques de projection linéaire visent à trouver une meilleure projection permettant de discriminer les pixels de l'IHS en cherchant des axes qui séparent le maximum possible le nuage de pixels afin de diminuer la corrélation inter-pixels (Figure 2.3). La projection linéaire tente à extraire p caractéristiques de l'IHS X en faisant une combinaison linéaire des bandes spectrales originales, avec

$p \ll D$:

$$Y_{ki} = T_{k1}.x_{1i} + T_{k2}.x_{2i} + \dots + T_{kD}.x_{Di}, \forall k \in \{1, \dots, p\}, \forall i \in \{1, \dots, N\} \quad (2.2)$$

En notation matricielle :

$$Y = TX \quad (2.3)$$

$$\begin{bmatrix} y_{11} & y_{12} & y_{13} & \dots & y_{1N} \\ y_{21} & y_{22} & y_{23} & \dots & y_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{p1} & y_{p2} & y_{p3} & \dots & y_{pN} \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & \dots & t_{1M} \\ t_{21} & t_{22} & t_{23} & \dots & t_{2M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{p1} & t_{p2} & t_{p3} & \dots & t_{pM} \end{bmatrix} \times \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1N} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{M1} & x_{M2} & x_{M3} & \dots & x_{MN} \end{bmatrix}$$

avec T est une matrice de transformation linéaire de taille $p \times M$ et Y est l'image réduite de taille $p \times N$. Dans ce contexte, plusieurs techniques de projection linéaire ont été développées pour réduire la taille massive des données hyperspectrales. Dans ce cadre, une technique de projection linéaire non supervisée qui est l'Analyse en Composantes Principales (ACP) a été largement utilisée pour la RD d'IHS [Jolliffe, 2002, Farrell and Mersereau, 2005]. Elle vise à trouver la meilleure projection d'IHS en maximisant la variance de données originales dans un sous-espace linéaire projeté tout en essayant de réduire la corrélation entre les pixels de l'IHS (Figure 2.4). Cette technique est basée essentiellement sur le calcul de la ma-

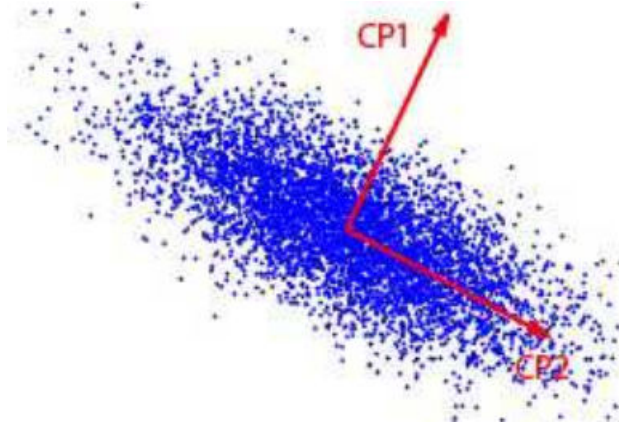


FIGURE 2.4 – Exemple de projection linéaire avec ACP

trice de covariance C entre les bandes spectrales de l'IHS X de taille $M \times N$:

$$C = \frac{1}{M} \sum_{i=1}^M x_i x_i^T \quad (2.4)$$

L'IHS originale est ensuite transformée sur les p premiers vecteurs propres de la matrice de covariance C . L'algorithme de rapport signal bruit *MNF* (Maximum Noise Fraction) [Bjorgan and Randeberg, 2015] a le même principe de projection que la technique ACP. La seule différence est que MNF son objectif est de chercher la meilleure projection qui maximise le rapport signal sur le bruit. Une autre technique est la projection préservée localement LPP (*Locality Preserving Projections*)

[Li et al., 2012] est introduite dans ce contexte. Elle vise à transformer les données hyperspectrales originales en préservant la structure géométrique originale en se basant sur la construction du graphe de voisinage laplacien entre les pixels de l'IHS originale. L'Analyse en Composantes Indépendantes ICA (*Independent Component Analysis*) [Wang and Chang, 2006, Li et al., 2014] consiste à transformer les données hyperspectrales dans un sous-espace optimal en essayant de maximiser la décorrélation entre les pixels de l'IHS. En utilisant des échantillons d'apprentissage, une technique de projection linéaire supervisée est l'analyse discriminante linéaire de Fisher LDA (*Linear Discriminant Analysis*) [Du and Younan, 2008, Yuan et al., 2013] consiste à projeter les données hyperspectrales de grande dimensions sur un sous-espace de taille inférieure, où les différentes classes de l'IHS sont bien séparées. Usuellement, l'objectif de LDA est de trouver une projection la plus discriminante des données hyperspectrales originales en maximisant la distance inter-classes et minimisant la distance intra-classes. L'avantage de cette technique qu'elle permet de mieux séparer les nuage des pixels de l'IHS, mais le problème de la malédiction de dimensionnalité reste toujours un grand défis dans le contexte hyperspectral. Dans le contexte de classification d'IHS,

4.2 Approches par projection non linéaire

Les techniques de projection linéaire ont pour objectif de chercher une meilleure projection au sens de différents objectifs : variance maximale, rapport signal sur le bruit maximal, distance intra-classes minimale, distance inter-classes maximale, etc ... Par ailleurs, la projection non linéaire d'IHS, est une autre variante de RD par projection. Le principe général de ces technique consiste à tenir en compte la non linéarité des données originales lors de la projection dans un nouvel sous-espace optimal (Figure 2.5) [Roweis and Saul, 2000] :

$$Y = f(X) \quad (2.5)$$

avec f est une fonction de transformation non linéaire. Usuellement, ces techniques

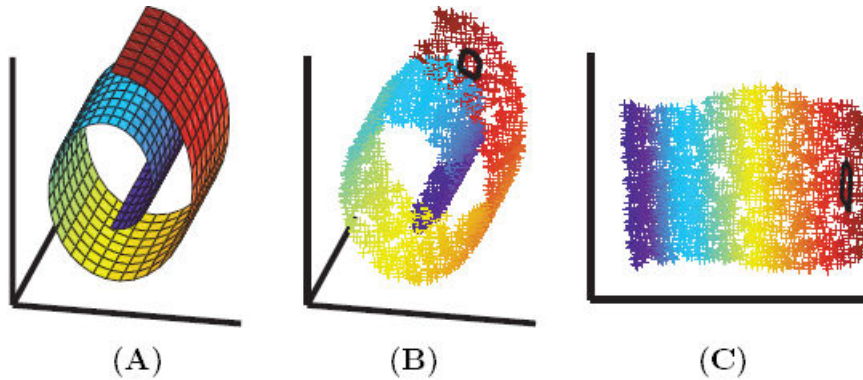


FIGURE 2.5 – Exemple de projection non linéaire : (A) : Données Swiss-roll, (B) : projection linéaire, (C) : projection non linéaire

visent à préserver les structures locales, les propriétés géométriques de l'IHS, tels que le voisinage, les distance et les relations topologiques entre les pixels. Dans ce cadre, plusieurs techniques sont basées sur l'utilisation des noyaux (*kernel*) et d'autres sur des graphes. Parmi les techniques de projection non linéaires basées sur les noyaux, L'ACP kernelisée *KPCA* (*Kernel PCA*) [Fauvel et al., 2006] vise à introduire des noyaux dans la matrice de covariance comme suit :

$$C = \frac{1}{M} \sum_{i=1}^M \Phi(x_i) \Phi(x_i)^T \quad (2.6)$$

Avec $\Phi(\cdot)$ représente le noyau de KPCA. En effet, ces noyaux permettent d'alléger la complexité du calcul. La technique ISOMAP (*Isometric Features Mapping*) [De Silva and Tenenbaum, 2003] exploite l'algorithme MDS (*Multidimensional Scaling*) [Cox and Cox, 2000] pour effectuer la projection des données hyperspectrales dans le nouvel sous-espace optimal. Cette technique tente à préserver les distances entre les pixels dans la nouvelle projection à partir de l'image originale. L'algorithme MDS a été utilisé dans cette technique pour déterminer une matrice de dissimilarité entre les pixels de l'IHS en exploitant la distance géodésique. L'objectif général de cette technique est de chercher une non linéarité (variété) contenant les données d'IHS. La technique Plongement Localement Linéaire LLE (*Locally Linear Embedding*) [Kim and Finkel, 2003, Fang et al., 2014] tente à préserver les pixels voisins les plus proches en attribuant une certaine pondération afin de trouver la meilleure combinaison non linéaire des données originales. La carte propre laplacienne LE (*Laplacian Eigenmaps*) [Belkin and Niyogi, 2003] consiste à trouver une nouvelle représentation Y de l'IHS originale X en essayant de préserver les relations de voisinage des pixels de l'IHS à l'aide du graphe de voisinage Laplacien.

5 La sélection de bandes spectrales

Dans le contexte d'analyse d'IHS, la corrélation intra-bandes spectrales rend l'extraction de l'information pertinente est très difficile. De plus, la quantité énorme des bandes spectrale génère souvent des propriétés statistiques et géométriques indésirables [Kuybeda et al., 2007, Jimenez and Landgrebe, 1998]. Figure 2.6 montre une matrice de corrélation comme une image en niveau de gris pour une IHS AVIRIS de l'Indian Pines contenant 220 bandes spectrales, où la haute intensité représente les bandes spectrales fortement corrélées. Par conséquent, il est désirable d'exploiter les données hyperspectrales de manière plus concise sans perte le sens physique de données originales et les d'informations pertinentes. Ceci donne lieu à sélectionner un sous-ensemble des bandes spectrales significatives qui contiennent l'information utile pour améliorer le taux de classification, la détection des cibles, etc. Par conséquent, les techniques de sélection des bandes visent à sélectionner automatiquement p bandes spectrales discriminantes et significatives à partir de M bandes spectrales représentant l'IHS originale. En d'autres termes, soit une IHS $X = [x_1, \dots, x_N] \in \mathbb{R}^{M \times N}$, avec x_i un vecteur pixel. M est le nombre de bandes



FIGURE 2.6 – Exemple de matrice de corrélation spectrale inter-bandes spectrales

spectrales. La nouvelle image réduite Y est de taille p avec ($p \ll M$) (Figure 2.7). L'atout des techniques de sélection des bandes est qu'elles préservent le sens physique des données d'origine. En effet, il s'agit de chercher une image réduite $Y_l = \{Y_{li}, \dots, Y_{pi}\}$ qui contient les mêmes informations spectrales que l'image originale X .

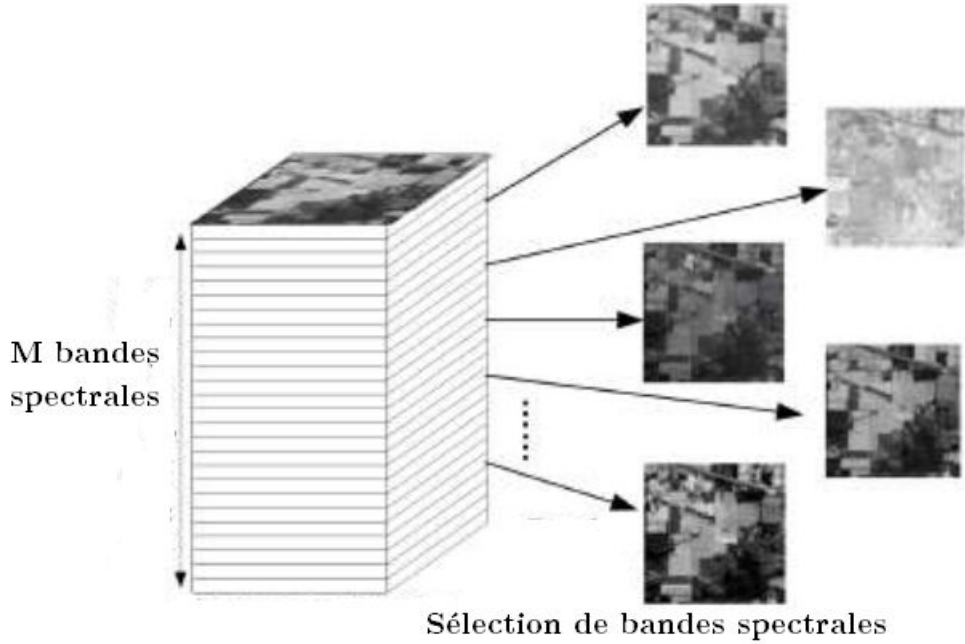


FIGURE 2.7 – Processus général de la sélection de bandes spectrales

$$\begin{bmatrix} P_{11} & P_{12} & P_{13} & \dots & P_{1M} \\ P_{21} & P_{22} & P_{23} & \dots & P_{2M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{N1} & P_{N2} & P_{N3} & \dots & P_{NM} \end{bmatrix} \text{ Sélection de bandes } \Rightarrow \begin{bmatrix} PS_{11} & PS_{12} & PS_{13} & \dots & PS_{1p} \\ PS_{21} & PS_{22} & PS_{23} & \dots & PS_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ PS_{N1} & PS_{N2} & PS_{N3} & \dots & PS_{Np} \end{bmatrix}$$

Dans le contexte hyperspectral, la sélection de bandes spectrales peut être supervisée, non supervisée et semi-supervisée. Dans la section suivante, nous détaillons les trois familles de la sélection de bandes spectrales.

5.1 Approches par sélection supervisée

La sélection supervisée de bandes spectrales tente à sélectionner les bandes discriminatives en mesurant leur pertinence avec les classes étiquetées, notamment les mesures sur la séparabilité des classes [Zhang et al., 2007], de l'information [Peng et al., 2005, Shen et al., 2013] et de l'erreur de classification [Pal and Foody, 2010].

- **La sélection de bandes basée sur la séparabilité des classes** : elle vise à sélectionner les bandes spectrales pertinentes en mesurant certaines métriques sur la séparabilité des classes. Des travaux de sélection supervisée des bandes spectrales ont été proposés [Ifarraguerri and Prairie, 2004, De Backer et al., 2005, Yan et al., 2016] en se basant sur les métriques de distance, notamment la divergence, la distance de Bhattacharyya et la distance de Jeffries–Matusita (JM). Par conséquent, les bandes spectrales qui ont fourni une grande séparabilité des classes vont être sélectionnées. Bien qu'un grand nombre des échantillons d'apprentissage étiquetés sont nécessaires pour évaluer les statistiques des classes. D'autres critères de sélection supervisée de bandes ont été proposés, à savoir mappeur d'angle spectral SAM (*spectral angle mapper*) et la divergence de projection orthogonale OPD (*orthogonal projection divergence*) [Chang, 2003]. Ces derniers mesurent la distance de classes par paire et prennent la moyenne de toutes ces distances afin d'effectuer la sélection de bandes spectrales informatives. Cela entraîne un calcul combinatoire très coûteux. Afin de remédier ce problème des approches ont été développées, notamment la stratégie de sélection séquentielle en avant SFS (*sequential forward selection*) et la stratégie de séquentielle flottante en avant SFFS (*sequential forward floating selection*) [Pudil et al., 1994].
- **La sélection de bandes basée sur les théories de l'information** : Ces mesures sont utilisés pour quantifier le gain informationnel de bandes spectrales de l'IHS [Fu et al., 2014, Hossain et al., 2014, Feng et al., 2015]. Dans la théorie de l'information, les deux critères qui ont été utilisés sont l'entropie de Shannon et l'information mutuelle MI (*mutual information*). Formellement, l'entropie de Shannon d'une variable aléatoire X_1 est l'unité fondamentale de l'information. L'entropie est notée $H(X_1)$ vise à quantifier la quantité de l'information contenue par X_1 . Elle est définie comme suit :

$$H(X_1) = - \int_{X_1} p(x_1) \log p(x_1) dx_1 \quad (2.7)$$

où $p(x_1)$ représente une fonction de densité de probabilité de la variable aléatoire X_1 . L'information partagée MI entre deux variables aléatoires X_1

et X_2 est calculée comme suit :

$$I(X_1; X_2) = \int_{X_1} \int_{X_2} p(x_1, x_2) \log \frac{p(x_1, x_2)}{p(x_1)p(x_2)} dx_1 dx_2 \quad (2.8)$$

$I(X_1; X_2)$ peut être interprétée comme étant une mesure de dépendance linéaire ou non linéaire entre ces deux variables aléatoires X_1 et X_2 . $I(X_1; X_2)$ est égale à zéro les deux variables sont complètement indépendante. Cependant, si $I(X_1; X_2)$ est faible alors une faible dépendance entre ces deux variables aléatoires. La relation entre l'entropie et MI est formulée comme suit :

$$I(X_1; X_2) = H(X_1) - H(X_1|X_2) = H(X_2) - H(X_2|X_1) \quad (2.9)$$

$$I(X_1; X_2) = H(X_1) + H(X_2) - H(X_2X_1) \quad (2.10)$$

Dans un contexte hyperspectral, les deux variables aléatoires peuvent être une bande spectrale X_1 de l'IHS et une classe étiquetée X_2 (Figure 2.8).

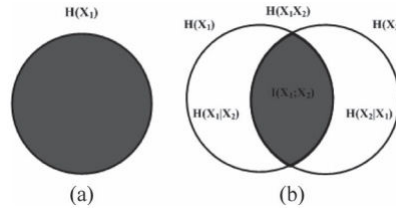


FIGURE 2.8 – Principe général de MI : (a) : entropie $H(X_1)$ (b) MI $I(X_1, X_2)$

- **La sélection de bandes basée sur la mesure de l'erreur de classification** : elle vise à sélectionner les bandes spectrales informatives en utilisant des modèles de classification tel que SVM et des mesures des erreurs de classification [Pal and Foody, 2010]. Un algorithme supervisé a été développé [Riedmann and Milton, 2003] permettant sélectionner les bandes spectrales pertinentes en se basant sur le taux de classification de vraisemblance maximale MLC (*Maximum Likelihood classification*) et le classifieur SVM. Les réseaux de neurones ont été utilisés pour la classification supervisée de bandes spectrales [Bajwa et al., 2004]. Ces modèles ont été appliqués dans le contexte de classification des zones d'agriculture. Une autre approche de sélection supervisée

Bien que ces techniques de sélection supervisée de bandes spectrales aient souvent une meilleure performance en raison de la présence d'informations *a priori*, mais il n'existe que des études limitées dans ce domaine. En effet, leur succès dépend grandement de la quantité d'échantillons étiquetés, mais la collecte des échantillons est coûteuse et prend beaucoup de temps pour les images hyperspectrales, et lorsqu'il n'y a que des pixels étiquetés limités, nous avons besoin d'informations supplémentaires pour améliorer leur performance.

5.2 Approches par sélection non supervisée

La sélection non supervisée de bandes spectrales vise à sélectionner automatiquement les bandes discriminatives et distinctives à partir des bandes spectrales originales [Jia et al., 2012, Qian et al., 2009, Sun et al., 2014, Tan et al., 2014]. Cette idée dépend de plusieurs problèmes, y compris une bonne représentation des données, un critère discriminatoire, une recherche appréciable et la taille optimale du sous-ensemble de bande à sélectionner. En général, les études connexes se concentrent principalement sur les critères de discrimination et les stratégies de recherche. Les techniques de sélection non supervisée traditionnelles décrivent souvent les bandes hyperspectrales à travers l'information spectrale naïve d'une image, et elles mesurent l'information et l'indépendance des bandes spectrales sur la base de différents ordres de statistiques [Martínez-Usó Martínez-Usó et al., 2007, Le Moan et al., 2011, Chang and Liu, 2014]. Malgré sa popularité dans la littérature, ce type de mesures ne peut être satisfaisant dans le cas où il existe des informations structurelles. Usuellement, les techniques de sélection non supervisées de bandes spectrales sont basées sur deux types : techniques basées sur le tri [Chang and Wang, 2006, Du and Yang, 2008, He et al., 2005, Kim and Rattakorn, 2011] et les techniques basées sur le regroupement (*clustering*) [Martínez-Usó Martínez-Usó et al., 2007, Qian et al., 2009, Mitra et al., 2002].

- **Les techniques basées sur le tri** : L'idée générale consiste à sélectionner les bandes distinctives et informatives en se basant sur des critères statistiques (exemple : divergence de leibler, variance, asymétrie, etc. Quelques travaux sont développés dans la littérature comme les techniques basées sur la divergence de l'information (ID) [Chang and Wang, 2006], ACP basée sur la variance maximale [Chang et al., 1999] et la sélection de bandes spectrales basée sur la similarité [Du and Yang, 2008]. Le désavantage de ce type de techniques est que la corrélation entre les bandes n'est pas prise en compte lors du calcul du contenu de l'information ou de la capacité de discrimination d'une bande spectrale. Ainsi, nous constatons parfois que les bandes spectrales redondantes peuvent être sélectionnées par ces techniques de tri.
- **Les techniques basées sur le clustering** : il s'agit d'effectuer un regroupement des bandes spectrales selon leur redondance ainsi de sélectionner une bande représentative de chaque groupe de bandes. Dans le contexte hyperspectral, il existe quelques techniques de sélection de bande basées sur le clustering comme la stratégie de liaison de Ward en utilisant la divergence (WaLuDi) ou en utilisant l'information mutuelle (WaLuMI) [Martínez-Usó Martínez-Usó et al., 2007] et la sélection de bandes spectrales en utilisant la propagation d'affinité (AP) [Qian et al., 2009]. Les techniques de sélection de bandes spectrales basées sur le clustering reposent principalement sur le critère de redondance entre les différentes bandes spectrales de l'IHS. En conséquence, il pourrait arriver que certaines bandes plus informatives et discriminantes que d'autres ne soient pas sélectionnées. Une autre technique d'élimination des bandes [Datta et al., 2014] tente à éliminer ité-

rativement une bande redondante parmi deux bandes voisines en fonction de la capacité de discrimination des bandes. L'inconvénient de cette technique qu'elle ne tient pas en compte la redondance des bandes éloignées, mais que de la corrélation entre les bandes voisines.

5.3 Approches par sélection semi-supervisée

La sélection semi-supervisée de bandes spectrales vise à chercher un sous-ensemble de bandes spectrales discriminantes et aussi informatives en utilisant à la fois des échantillons étiquetés et des échantillons non étiquetés. Dans ce contexte, ces techniques peuvent être divisées en deux familles, celles basées sur le clustering et d'autres basées sur le multi-apprentissage (*manifold learning*).

- **Les techniques basées sur le clustering** : Il s'agit de sélectionner les bandes spectrales informatives en se basant les techniques de regroupement et l'utilisation des échantillons d'apprentissage étiquetés et non étiquetés. Dans ce contexte, une approche semi-supervisée basée sur le tri est proposée sSelect (ranking-based semi-supervised feature selection) [Zhao and Liu, 2007]. Elle permet de transformer chaque bande dans un indicateur de cluster et évalue cet indicateur en mesurant sa séparabilité et consistance avec les classes étiquetées. Par conséquent, les bandes sont regroupées selon leurs valeurs d'évaluation. Dans cette approche, les critères de séparabilité et de consistance correspond à la haute information et discrimination. Cependant, la combinaison de bandes informatives et discriminatives n'est pas toujours satisfaisante vu que les bandes spectrales sont fortement corrélées et elles ne permet pas de fournir des informations supplémentaires à la classification d'IHS. Dans le même contexte, une approche semi-supervisée SSFC (*semi-supervised feature clustering*) [Quinzán et al., 2009] permet de sélectionner les bandes spectrales pertinentes en exploitant à la fois des échantillons étiquetés et non étiquetés. Les bandes sont sélectionnées en mesurant la similarité avec l'entropie conditionnelle et l'information mutuelle conditionnelle (MI). Par conséquent, dans chaque cluster, les échantillons étiquetés sont utilisés pour sélectionner les bandes spectrales qui ont une capacité discriminative très élevée et d'éliminer les bandes spectrales informatives. De façon générale, les bandes spectrales discriminatives ne peuvent être facilement détectées vu que les échantillons étiquetés sont très limités.
- **Les techniques basées sur le *manifold learning*** : elles supposent que les échantillons étiquetés et non étiquetés se situent sur un collecteur avec une dimension inférieure à celle de l'image originale. Habituellement, ce type de techniques modéliser la structure de collecteur par les graphes adjacents. Une approche de sélection discriminative sensible à la localité LSDF (*locality sensitive discriminant feature selection*) a été proposée [Zhao et al., 2008] permettant de sélectionner les bandes spectrales en exploitant un graphe inter-classe et un graphe intra-classes. Ces deux graphes visent à découvrir la structure discriminative et géométrique de données hyperspectrales. Les

bandes spectrales sont triées selon leurs degrés de préservation de structure des graphes et celles qui ont des degrés élevés sont sélectionnées. Dans le même travail de recherche, une approche de sélection de bandes basée sur la contrainte de score NCS (*new constraint score*) [Kalakech et al., 2011] adopte la même démarche que LSDF en ajoutant un graphe supplémentaire des échantillons non étiquetés. Cela permet de réduire la dépendance de LSDF sur les échantillons étiquetés. Usuellement, il existe deux variantes de LSDF, à savoir celles basées sur la pondération WLSDF (*weighted LSDF*) [Chen et al., 2010] et celles basées sur les graphes avec une réduction d'échantillons non étiquetés GUSR (*unlabeled sample reduction in semisupervised graph*) [Huang et al., 2013]. Récemment, une technique de approche de sélection semi-supervisée basée sur l'IM trivariée (TMI) et les graphes a été proposée [Feng et al., 2014]. Les graphes sont utilisés pour sélectionner les bandes spectrales informatives en exploitant des échantillons non étiquetés en minimisant la redondance. Bien que les échantillons étiquetés ne mesurent pas la redondance avec précision. Un critère semi-supervisé basé sur la discrimination et information maximale MDI (*Maximum Discrimination and Information*) a été utilisé dans ce travail permettant de sélectionner non seulement les bandes spectrales discriminatives mais aussi informatives. Pour sélectionner non seulement un sous-ensemble de caractéristiques discriminatives mais aussi informatives [Feng et al., 2015]. Cependant, le critère MDI est difficile à calculer avec précision en raison des de la taille limite des échantillons et de la complexité du calcul [Peng et al., 2005]. En outre, un nouveau critère basé sur le concept de discrimination élevée, une information élevée et une faible redondance DIR (*high discrimination, high information, and low redundancy*) visant à sélectionner les bandes spectrales informatives, discriminatives et peu redondantes en exploitant simultanément des échantillons étiquetés et non étiquetés [Feng et al., 2015]. Néanmoins, les critères de sélection semi-supervisée de bandes spectrales [Zhao and Liu, 2007, Feng et al., 2014] sont des approximations de faible ordre du critère MDI. De plus, le critère DIR nécessite des conditions d'approximation plus détendues. Dans ce cadre, un algorithme de sélection clonale CSA (*clonal selection algorithm*) a inspiré la capacité des systèmes immunitaires artificiels AIS (*Artificial Immune Systems*) pour l'apprentissage et l'adaptabilité de résolution des problèmes complexes [Feng et al., 2011]. Par rapport aux techniques populaires de sélection de bandes spectrales basées sur la recherche incrémentale [Peng et al., 2005, Shen et al., 2013], l'algorithme CSA peut détecter et chercher les bandes spectrales sélectionnées entièrement en évoluant vers une solution optimale en exploitant le mécanisme de recherche basé sur la population. Habituellement, les opérateurs de clones et de mutations adaptatifs sont conçus pour accélérer la convergence dans l'algorithme CSA.

6 Estimation du nombre de bandes/composantes à conserver

Les techniques de projection et de sélection de bandes spectrales cités précédemment produisent pour la plupart un nombre p de bandes/composantes égal au nombre de bandes spectrales M de l'IHS. La question qui est toujours ouverte et posée, quel est le nombre optimal de bandes/composantes à conserver pour effectuer la RD d'IHS sans perte d'informations pertinentes. Par conséquent, il est nécessaire de ne retenir qu'un petit nombre p de ces bandes spectrales/composantes. Le tableau 2.1 présente les différents métriques d'estimation du nombre de bandes/composantes à conserver pour différentes techniques de RD. Pour les techniques de l'ACP et

Technique	Famille de RD	Nombre de bandes/composantes p
LDA	projection linéaire supervisée	$p = k - 1$ (k : nombre de classes de l'IHS)
ACP	projection linéaire non supervisée	p estimé en fonction de la variance
MNF	projection linéaire non supervisée	p estimé en fonction de la variance
La sélection basée sur les théories de l'information et les distances	Sélection supervisée de bandes spectrales	$p = k$ (k : nombre d'échantillons d'apprentissage)
La sélection de bandes basée sur le tri	sélection non supervisée	p estimé avec VD (<i>Virtual Dimensionality</i>)
La sélection de bandes basée sur le clustering	sélection non supervisée	$p = k$ (k : nombre de clusters)

Tableau 2.1 – Estimation du nombre de bandes/composantes à retenir

MNF, le nombre de composantes à retenir est déterminé en fonction de la variance des composantes. Dans ce cadre, plusieurs métriques ont été développées pour déterminer le nombre minimal p des composantes, à savoir critère de Kaiser, pourcentage de la variance, Scree-test de Catell et le rapport de croissance. Pour la technique LDA, le nombre minimal de composantes p est égal au nombre de classes de l'IHS moins un. Concernant les techniques de sélection de bandes spectrales, le nombre de bandes à conserver est dépend essentiellement au nombre d'échantillons d'apprentissage (classes) dans le cas supervisé. Par ailleurs, dans le cas de sélection non supervisée de bandes spectrales, des travaux ont été basés sur la technique de la dimensionnalité virtuelle VD (*Virtual Dimensionality*)

[Chang and Liu, 2014, Chang and Du, 2004] et d'autres sont basés sur des techniques d'estimation de sous-espace [Asl et al., 2014, Mojaradi et al., 2009]. Ces techniques visent à estimer le nombre de signatures caractérisant les données hyperspectrales d'origine. Certains travaux [Chang and Liu, 2014, Fisher and Chang, 2010] ont montré que le nombre de bandes spectrales p à sélectionner doit être supérieur ou égal au nombre estimé VD de l'IHS ($p \geq VD$). Dans d'autres travaux [Asl et al., 2014] ont précisé que le nombre maximal optimal p de bandes spectrales à sélectionner est probablement égale à deux fois de la dimension du sous-espace du signal. Un algorithme de traçage de compacité de distribution a été proposé [Sun et al., 2015b] pour sélectionner une taille appropriée du sous-ensemble de bandes spectrales en utilisant une estimation non paramétrique de la fonction de densité de probabilité pour l'ensemble de bandes spectrales. Il repose essentiellement sur une décomposition de valeur propre de la matrice de similarité pour mesurer la compacité de distribution. En outre, il existe de nombreux travaux qui ont utilisé la technique d'estimation de la dimension intrinsèque dans le domaine du traitement du signal [Tolle et al., 2003, Camastra, 2003]. La dimension intrinsèque décrit le nombre de variables indispensables pour représenter le signal. Par conséquent, cette technique est très proche de la technique VD. Nous décrivons brièvement ces différentes méthodes d'estimation du nombre de bandes/composantes pour la RD en IHS.

6.1 Le critère de Kaiser

Chaque bande B_l observée de l'IHS X contribue à la variance totale de l'ensemble de données réduites Y . En effet, la composante Y_l représentant une quantité significative de variance doit être conservée. L'idée générale de cette approche vise à conserver et d'interpréter uniquement les composantes Y_l dont la valeur propre λ_l est supérieure ou égale à un ($\lambda_l \geq 1$).

6.2 Pourcentage de la variance

Il s'agit de conserver les composantes significatives de l'image réduite en se basant sur la variance expliquée par chaque composante Y_l , les valeurs propres λ et la somme cumulée des valeur propres. En d'autres termes, étant donné un seuil s , les composantes qui seront conservées celles ayant une proportion ou pourcentage de variance inférieure au seuil s (Figure 2.9). Dans d'autres cas, les composantes Y_l ayant des valeurs propres inférieures au seuil s sont conservées.

6.3 Scree-test de Cattell

En utilisant les valeurs propres, un histogramme est construit avec l'axe horizontal représente le nombre de composantes et l'axe vertical représente les valeurs propres λ . En se basant sur la décroissance des valeurs propres et la présence des coudes. Dans le cas où il y a des coudes entre les dimensions ayant des grandes valeurs propres et celles ayant des faibles valeurs propres, il est désirable de retenir

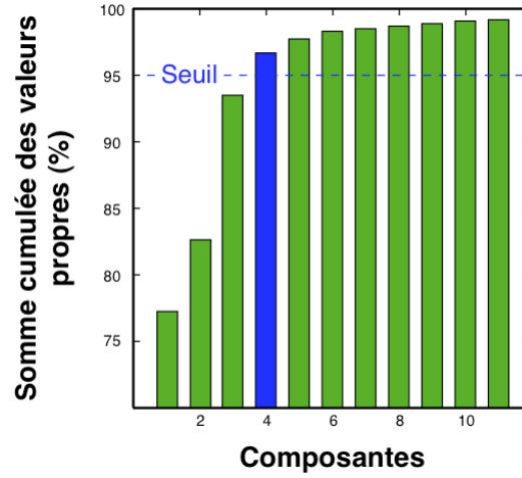


FIGURE 2.9 – Exemple d'un seuil de 95% : cas de pourcentage des valeurs propres cumulées

uniquement que les composantes se trouvant en amont, c'est à dire les composantes qui apparaissent avant la rupture sont supposées significatives et seront conservées et ceux qui apparaissent après la rupture sont supposés non significatives et ne sont pas retenues. Figure 2.10 montre un exemple où les cinq premières composantes Y_l sont conservées et sont considérées comme significatives. Usuellement, le seuil α

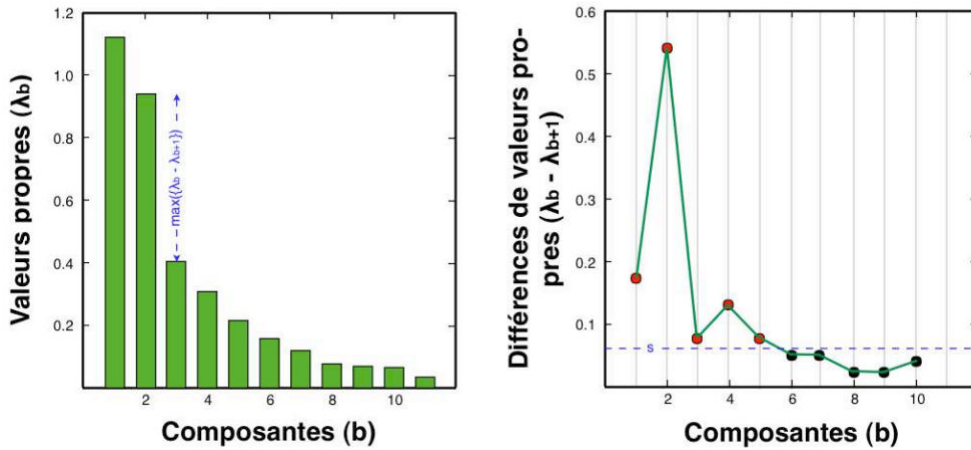


FIGURE 2.10 – Exemple de la technique du coude avec $\alpha = 10\%$.

est fixé entre 10% et 20% afin de détecter automatiquement la présence des coudes comme suit :

$$s = \alpha \max(\lambda_1 - \lambda_2, \dots, \lambda_{k-1} - \lambda_k, \dots, \lambda_{M-1} - \lambda_M) \quad (2.11)$$

D'après l'équation (2.11), le coude se forme lorsque $\lambda_{l-1} - \lambda_l < s$ de la plus petite composante l constituera. Dans ce cas, nous ne conservons plus les composantes Y_l .

6.4 Rapport de croissance

Dans cette approche, la valeur de seuil s est fixé et estimé de façon automatique automatiquement ce qui influe sur le choix du nombre de bandes spectrales à retenir et aussi sur les résultats d'analyse et du traitement d'IHS (détection de cibles, détection d'anomalies, classification, etc...). Dans ce cadre, l'estimateur de rapport de croissance a été développé [Ahn and Horenstein, 2013]. Il est basé essentiellement sur l'évaluation de la perte d'information en fonction d'élimination des composantes les unes après les autres. Usuellement, le rapport de croissance tente à détecter et déterminer le nombre de composantes faiblement significatives comme suit :

$$\lambda_k^* = \frac{\sum_{i=k}^M \lambda_i}{\sum_{i=k+1}^M \lambda_i} \quad (2.12)$$

avec λ_k correspond à la valeur propre de la k -ième composante.

6.5 La dimensionnalité virtuelle

Pour estimer le nombre initial des bandes à utiliser, la technique de dimensionnalité virtuelle VD (*Virtual Dimensionality*) [Chang and Du, 2004] largement utilisée dans la littérature afin d'estimer le nombre de bandes spectrales à sélectionner noté p . La technique de VD est une mesure de nombre des signaux spectraux distincts. En outre, il s'agit de considérer une source de signal par un pigeon et une bande spectrale par le trou de pigeon. Par conséquent une bande spectrale donnée peut être utilisée pour accueillir une source particulière. Ainsi, si une source de signal est donnée, il es possible de détecter la bande spectrale pertinente. Étant donnés les valeurs propres de la matrice de corrélation notée λ_l^* et la matrice de covariance notée λ_l avec l représente $l^{\text{ième}}$ bande spectrale. Une source signal sest présentée si $\lambda_l^* - l^{\text{ième}}$ est positive. Les valeurs propres λ sont calculées comme suit :

$$\lambda_l^* > \lambda_l > \sigma_{nl}^2 \text{ pour } l = 1, \dots, VD \quad (2.13)$$

$$\lambda_l^* > \lambda_l = \sigma_{nl}^2 \text{ pour } l = VD + 1, \dots, L \quad (2.14)$$

Avec l représente $l^{\text{ième}}$ bande spectrale et σ_{nl}^2 représente la variance de bruit pour chaque bande l . Dans cette section, nous présentons quelques travaux de RD.

7 Travaux existants : étude comparative

7.1 Tableau comparatif

Nous présontons dans cette section quelques travaux existants concernant la RD en IHS. L'objet de cette étude est d'évaluer et de comparer les différentes approches de RD afin de mieux cerner les problèmes traités dans ce travail de thèse. De ce fait, différents critères sont utilisés pour montrer les atouts et les limites de chaque travail de recherche présenté. Parmi ces critères, nous citons la préservation de l'information spectro-spatiale, la réduction de corrélation, la stabilité et l'incertitude. Les deux tableaux 3.1 et 2.3 montrent quelques travaux existants de RD en IHS.

Auteur	Technique de RD	Famille	Information spectro- spatiale	Corrélation	Stabilité	Incertitude
[Li et al., 2012]	LPP	proj. lin. non sup.	non	oui	30	oui
[Du and Younan, 2008]	LDA	proj. lin. sup.	non	non	15	oui
[Bjorgan and Randeberg, 2015]	MNF	proj. lin. non sup.	non	oui	21	oui
[Li et al., 2014]	ACI	proj. lin. non sup.	non	oui	21	oui
[Fauvel et al., 2006]	KPCA	proj. nonlin. non sup.	non	non	15	oui
[Kim and Finkel, 2003]	LLE	proj. nonlin. non sup.	oui	oui	30	oui
[Fang et al., 2014]	LLE + information spatiale	proj. nonlin. non sup.	oui	oui	37	oui
[De Silva and Tenenbaum, 2003]	ISOMAP	proj. nonlin. non sup.	oui	oui	30	oui
[Cox and Cox, 2000]	MDS	proj. lin. non sup.	non	oui	25	oui
[De Silva and Tenenbaum, 2003]	LE	proj. nonlin. non sup.	non	oui	30	oui
[Bao et al., 2016]	EL + Attributs morph. math.	proj. nonlin. non sup.	non	oui	35	oui

Tableau 2.2 – Quelques travaux de projection linéaire/non linéaire en IHS

Auteur	Technique de RD	Famille	Information spectro-spatiale	Corrélation	Stabilité	Incertitude
[De Backer et al., 2005]	distances	SB sup	non	oui	10	oui
[Yan et al., 2016]	erreur class. SVM	SB sup	non	oui	15	oui
[Chang, 2003]	PBS	SB nonsup	non	non	32	oui
[Fu et al., 2014]	MI	SB sup	non	oui	15	oui
[Hossain et al., 2014]	MI	SB nonsup	non	non	25	oui
[Feng et al., 2015]	MI	SB nonsup	non	30	oui	
[Chang and Wang, 2006]	CBS	SB nonsup	non	15	oui	
[Du and Yang, 2008]	MI + distance	SB nonsup	non	15	oui	
[He et al., 2005]	score	SB sup tri	oui	oui	25	oui
[Kim and Rattakorn, 2011]	score +ACP	SB nonsup	oui	non	20	oui
[Tan et al., 2014]	SVM + morphologie math.	SB sup	oui	oui	30	oui
[Qian et al., 2009]	PA	SB nonsup	non	non	15	oui
[Mitra et al., 2002]	PA	SB nonsup	non	non	30	oui
[Chen et al., 2010]	SVM + Graphes	SB semi-sup	oui	oui	35	oui
[Huang et al., 2013]	Graphes	SB semi-sup	oui	oui	35	oui
[Feng et al., 2014]	MI + sélection clonale	SB semi-sup	oui	oui	35	oui

Tableau 2.3 – Quelques travaux de sélection de bandes spectrales

7.2 Synthèse et discussions

D'après une large littérature sur les techniques de RD en IHS, les techniques de sélection de bandes ont l'avantage de préserver le sens physique de données originales. Dans un contexte supervisé, ces techniques exigent un grand nombre d'échantillons étiquetés afin de sélectionner les bandes spectrales informatives. De ce fait, un problème est très connu dans l'analyse d'IHS est que le nombre d'échantillons d'apprentissage étiquetés sont insuffisants (la malédiction de dimensionnalité). Vu ce problème, les techniques supervisées de RD ne sont pas satisfaisante en terme de redondance et de précision de l'information fournie. Par conséquent, des bandes informatives uniquement sont préservées mais la redondance d'informations n'est pas pris en compte. En outre, des bandes distinctives et discriminatives peuvent être sélectionnées par les techniques de sélection non supervisée de bandes spectrales. Néanmoins, les techniques non supervisées ne souffrent pas du problème de la malédiction de dimensionnalité, mais les bandes sélectionnées ne sont pas informatives. En d'autres termes, les informations contextuelles ne sont pris en compte dans cette démarche de sélection de bandes spectrales. Les techniques de projection tentent à chercher un sous-espace optimal en fournissant des informations significatives et utiles en combinant toutes les bandes spectrales de l'IHS. Cependant, les images réduites se sont des images obtenues après une transformation linéaire ou non linéaire de bandes spectrales, c'est à dire de passer de l'espace original vers un nouvel espace réduit. Ceci provoque une déformation du sens physique de l'image originale ainsi une perte d'information fournie par les composantes obtenues.

8 Conclusion

La RD est une tâche primordiale pour l'analyse d'IHS, où chaque famille de technique de RD a ses propres atouts au terme de préservation du sens physique de données hyperspectrales, l'élimination de la redondance et la forte corrélation entre les bandes spectrales, ainsi l'intégration de l'information spatiale et contextuelle. En se basant sur les atouts de chacune, la sélection de bandes permet d'avoir des informations de l'image originale ainsi les techniques de projection fournissent des nouvelles caractéristiques qui portent une quantité importante des informations significatives. Pour exploiter la complémentarité d'informations entre les techniques de projection et les techniques de sélection de bandes spectrales, des approches hybrides peuvent exploiter cette richesse d'information dont le but de pouvoir des images réduites contenant les informations utiles qui peuvent être exploitées par la suite dans une tâche d'analyse d'IHS. Dans le chapitre suivant, nous présenterons notre contribution dans ce manuscrit. En effet, dans une première partie nous allons montrer une approche de RD spectro-spatiale qui a pour objectif de réduire les données hyperspectrales en tenant compte l'information spectro-spatiale afin d'améliorer le taux de classification d'IHS. Dans une deuxième partie, nous tentons à développer une approche hybride qui appréhendera deux techniques de RD intégrant des connaissances sémantique pour pourvoir des bandes spectrales informatives, discriminatives

et distinctives qui devront être utiles pour l'interprétation sémantique en IHS.

Réduction adaptative de dimensions pour l'interprétation sémantique d'IHS

Sommaire

1	Introduction	58
2	Motivations et ébauche	58
3	De la RD vers l'extraction de la sémantique bas-niveau / haut-niveau	59
3.1	Sémantique bas-niveau	59
3.2	Sémantique haut-niveau	60
3.3	Le passage de la sémantique bas-niveau vers la sémantique haut-niveau	60
4	Représentation et approximation tensorielle d'IHS	61
4.1	Représentation tensorielle d'IHS	61
4.2	Aplatissement tensoriel des vecteurs et matrices en mode-n	61
5	Édification de l'approche proposée de RD spectro-spatiale	63
5.1	Architecture générale	63
5.2	La projection par TLPP	64
5.3	La sélection de bandes spectrales avec CBS	65
5.4	La classification spectro-spatiale avec SVM	66
5.5	La fusion de classification avec la théorie de Dempster-Shafer	67
5.6	Synthèse et discussion	68
6	Approche de RD adaptative pour l'interprétation sémantique d'IHS	69
6.1	Motivations	69
6.2	Architecture générale	70
6.3	Extraction des connaissances à partir des caractéristiques	70
6.4	Sélection adaptative des bandes spectrales pertinentes sous contraintes sémantiques	73
6.5	Discussion	75
7	Conclusion	76

1 Introduction

Dans le contexte de hyperspectrale, les techniques existantes de RD tentent à réduire la masse de données fournie par la technologie hyperspectrale tout en essayant de préserver des informations significatives. Comme nous avons présenté dans le chapitre précédent, les approches de projection exploitent la totalité des données hyperspectrales, c'est à dire en combinant toutes les bandes spectrales de l'IHS ces techniques vise à trouver un sous-espace optimal pour représenter les données originales en fournissant des informations significatives. Les approches par sélection, c'est une autre variante de RD permettant de sélectionner les bandes informatives dans le cas supervisé et les bandes spectrales discriminatives et distinctives dans le cas non supervisé. Cependant, les techniques de projection provoquent une perte d'informations de données originales ainsi le sens physique de données est n'est pas préservé. De plus, certaines techniques de sélection de bandes se focalisent essentiellement sur la présence des échantillons d'apprentissage, où ces données d'apprentissage ne sont pas toujours disponible. D'autres techniques de sélection non supervisée se reposent essentiellement sur des critères statistiques permettant de regrouper les bandes spectrales similaires ou de trier les bandes pertinentes discriminatives selon un critère donné. Dans les sections suivantes, nous essayons de présenter deux approches de RD ; la première est une approche de RD sepectro-spatiale intégrant l'information spectrale et spatiale conjointement afin d'améliorer la classification d'IHS. Par ailleurs, la deuxième est une approche de RD basée sur les techniques d'extraction de connaissances sémantiques permettant de sélectionner automatiquement et de façon adaptative les bandes spectrales discriminatives, distinctives et informatives pour une interprétation sémantique.

2 Motivations et ébauche

Usuellement, le problème de RD [Khoder et al., 2012] peut être défini comme suit : Considérons un ensemble de données X représenté par $N \times M$ constitué par N vecteurs de données $i \in \{1, 2, \dots, N\}$ avec M dimensions, et p est la dimensionnalité intrinsèque (où $p < M$, et souvent $p \ll M$). La dimensionnalité intrinsèque signifie que les points dans l'ensemble de données (pixels) de X sont situés sur ou près d'un collecteur de p -dimensions qui est incorporé dans l'espace de M dimensions. Usuellement, les techniques de RD transforment l'ensemble de données X de dimension M en un nouvel sous-espace optimal Y de p -dimensions, tout en préservant autant que possible la géométrie des données. En général, ni la géométrie du collecteur de données, ni la dimension inférieure p de l'ensemble de données X sont connues. Par conséquent, la RD est un problème majeur qui ne peut être résolu qu'en supposant certaines propriétés des données telle que la RD. Dans le contexte de l'IHS, les méthodes de projection extraient les p composantes, avec $p \ll M$, tel que :

$$Y = TR \tag{3.1}$$

où T est la matrice de transformation linéaire, Y est la matrice réduite et R est les différentes signatures spectrales.

Cependant, ces approches par projection linéaire sont des techniques basées sur l'algèbre matricielle qui s'appuient sur l'aspect spectral uniquement négligeant l'information spatiale. En effet, elle permettent de vectoriser les données hyperspectrales en deux dimensions (2D), ce qui provoque une déformation de la structure spatiale dans un pixel spectral ou dans les objets de l'image. Pour résoudre ces problème, une nouvelle représentation d'IHS a été développée basée sur les tenseurs [Renard and Bourennane, 2009]. Cette approche tente à incorporer et analyser les deux informations spatiales et spectrales conjointement en utilisant le modèle de tenseur et la décomposition de Tucker. Le traitement tensoriel [De Lathauwer, 1997] a montré son efficacité dans plusieurs domaines : télécommunication [Sidiropoulos et al., 2000], traitement d'images [Muti and Bourennane, 2007, Vasilescu and Terzopoulos, 2005] et l'analyse d'IHS [Renard and Bourennane, 2008, Letexier et al., 2008]. Nous présentons détaillons dans la section suivante le modèle tenseur pour la représentation d'IHS.

3 De la RD vers l'extraction de la sémantique bas-niveau / haut-niveau

La RD visent à réduire la masse de données hyperspectrales tout en essayant de préserver l'information significative, en particulier l'information spectrale. Dans ce contexte, les approches de projection exploitent toutes les bandes spectrales pour chercher la meilleure projection qui représente le plus fidèle possible l'image originale. Par conséquent, la projection obtenue suite à cette transformation de toutes les bandes spectrales représentent des nouvelles informations significatives qui peuvent être utiles pour l'interprétation d'IHS. Nous visons à extraire la sémantique c'est à dire la signification physique de ces caractéristiques obtenus par les approches de projection afin d'enrichir la sélection de bandes spectrales en préservant l'information pertinente pour améliorer l'interprétation d'une scène hyperspectrale. Par conséquent, la sémantique peut être définie sur deux niveaux, à savoir la sémantique bas-niveau et la sémantique haut-niveau.

3.1 Sémantique bas-niveau

La transformation de bandes spectrales par les approches spectrales fournit de nouvelles caractéristiques représentant une image réduite Y_l de l'IHS d'entrée X . Chaque valeur du point de nuage de point obtenu Y_l est une combinaison de toutes les bandes spectrales (Fig. 3.1). Par conséquent, à partir de ces données projetées la sémantique bas-niveau peut être définie comme étant la description de nouveaux points de données obtenues qui peuvent fournir des nouvelles informations significatives qui ne peuvent pas être détectées dans l'image originale. Pour cela, des techniques d'extraction de connaissances peuvent être implémentées pour extraire de la

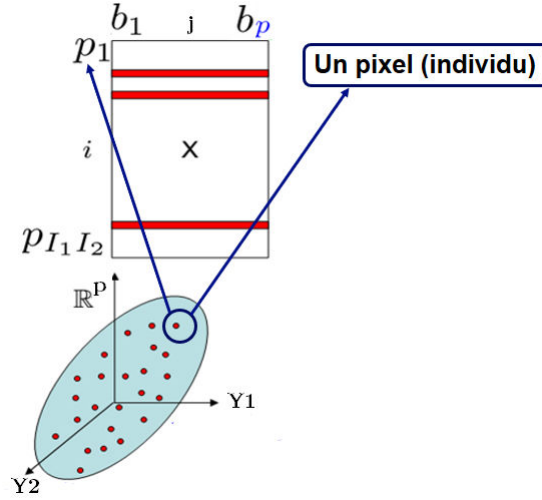


FIGURE 3.1 – Principe de projection de bandes spectrales

connaissance sémantique bas-niveau décrivant les différentes informations issues des approches de projection. En effet, la sémantique bas-niveau peut être relative à la similarité, la corrélation et la dépendance de nouveaux points de données, ce qui permet d'avoir des connaissances sur les classes de l'image originale. En outre, des informations spectrales peuvent être obtenues qui représentent l'importance des caractéristiques pour la discrimination des classes de l'image originale.

3.2 Sémantique haut-niveau

Cette sémantique correspond essentiellement à la description de la scène hyperspectrale en attribuant un sens spectro-spatial de données hyperspectrales. Usuellement, dans un contexte hyperspectral, la description d'une scène avec une sémantique haut-niveau permet d'avoir des connaissances relatives aux classes de l'image, notamment le nombre de classes, leur composition et leur relations spatiales. De ce fait, il est désirable de préserver l'information pertinente ainsi le sens physique de données hyperspectrales dans la RD afin d'améliorer l'interprétation de l'imagerie hyperspectrale. En effet, les bandes pertinentes doivent contenir des informations relatives à l'identification et la discrimination de l'image originale, ce qui permet d'avoir une sémantique haut-niveau décrivant les entités spectrales de la scène. Bien que la sémantique haut-niveau vise à fournir des informations relatives aux objets de la scène, la sémantique bas-niveau décrit la scène au niveau pixel, c'est à dire au niveau les points de données obtenus par les approches de projection.

3.3 Le passage de la sémantique bas-niveau vers la sémantique haut-niveau

Partant par la sémantique bas-niveau qui décrit l'ensemble de pixels obtenus par le biais des approches de projection d'IHS, des connaissances relatives à cette sémantique

tique peuvent être utiles pour enrichir la description de la scène hyperspectrale. De ce fait, la sélection de bandes tentent à préserver l'information pertinente de l'image originale en préservant le sens physique de données hyperspectrales. Par conséquent, des connaissances de la sémantique bas-niveau (pixels) peuvent améliorer la sémantique haut-niveau en donnant une certaine signification spectrale, notamment au niveau de la discrimination de classes. En d'autres termes, les données projetées avec les approches de projection peuvent donner des nouvelles informations utiles qui peuvent être exploitées par la sélection de bandes afin de préserver l'information pertinente et utile pour améliorer l'interprétation d'IHS en préservant le sens physique. Ce passage de la sémantique bas-niveau vers la sémantique haut-niveau a pour but d'avoir des connaissances relatives aux pixels des caractéristiques qui vont être des informations significatives et utiles pour la sémantique haut-niveau de la scène hyperspectrale, c'est à dire pour l'identification et la discrimination des classes en préservant les bandes spectrales pertinentes.

4 Représentation et approximation tensorielle d'IHS

4.1 Représentation tensorielle d'IHS

Naturellement, dans une IHS $I_{(3D)}$ chaque pixel x_i est représenté comme étant un vecteur et les données sont typiquement représentées par un cube hyperspectral 3D. Par conséquent, un tenseur de *troisième-ordre* (tableau multidimensionnel) peut être utilisé comme un modèle mathématique pour l'analyse d'IHS. En effet, cette représentation tensorielle permet de préserver la structure spatiale initiale de l'IHS et de conserver les contraintes de voisinage entre les pixels de l'image. Usuellement, les dimensions spatiales sont respectivement associées au *premier-ordre* (1-mode) et *deuxième-ordre* (2-mode) du tenseur et la dimension spectrale est associée au *troisième-ordre* du tenseur (Figure 3.2). En d'autres termes, les données d'IHS sont entièrement considérées comme un tenseur de *troisième-ordre*, c'est à dire les dimensions de l'IHS sont accessibles par 3 indices qui sont notées par $R \in \mathbb{R}^{M \times I_2 \times I_1}$, avec les éléments sont arrangés comme suit : r_{m_1, i_2, i_1} , où $m_1 = 1, \dots, M$; $i_2 = 1, \dots, I_2$, $i_1 = 1, \dots, I_1$ et \mathbb{R} est le collecteur réel. Chaque indice est appelé mode. La représentation tensorielle est mathématiquement fondée sur l'algèbre multi-linéaires qui étudie les propriétés des données de tenseur R de mode n . Parmi les opérations de tenseur d'ordre 3 avec les algèbres multi-linéaires est l'aplatissement tensoriel des vecteurs et matrices en mode- n .

4.2 Aplatissement tensoriel des vecteurs et matrices en mode- n

Un vecteur mode- n de R est un vecteur de n -dimensions en variant l'indice i_n et préservant les autres indices fixes. Considérons tous les vecteurs de mode- n comme colonnes, nous pouvons obtenir une matrice $Mat_{R_n} \in \mathbb{R}^{I_n \times (I_1 \dots I_{n-1} I_{n+1} \dots I_N)}$. La matrice Mat_{R_n} obtenue est peut être considérée comme étant un aplatissement de mode- n (n -mode flattening) de tenseur R . En d'autres termes, La matrice Mat_{R_n}

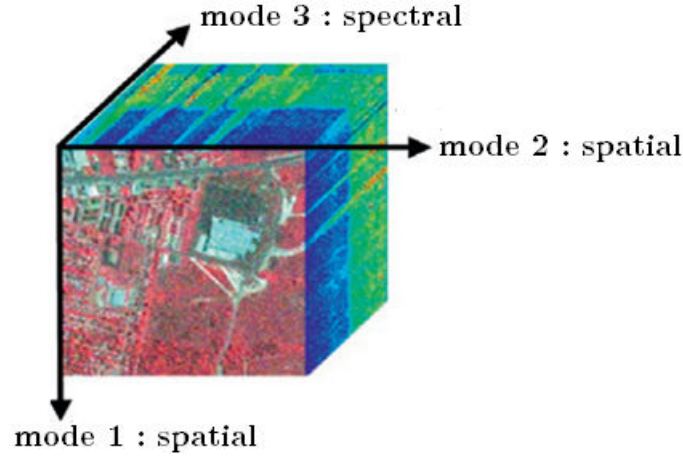


FIGURE 3.2 – Illustration de la représentation tensorielle d'une IHS : $R \in \mathbb{R}^{M \times I_2 \times I_1}$

est l'ensemble des vecteurs de le mode- n obtenu en préservant les indices fixes et en variant l'indice i_n (Figure 3.4). Dans le contexte hyperspectral, la matrice de pliage

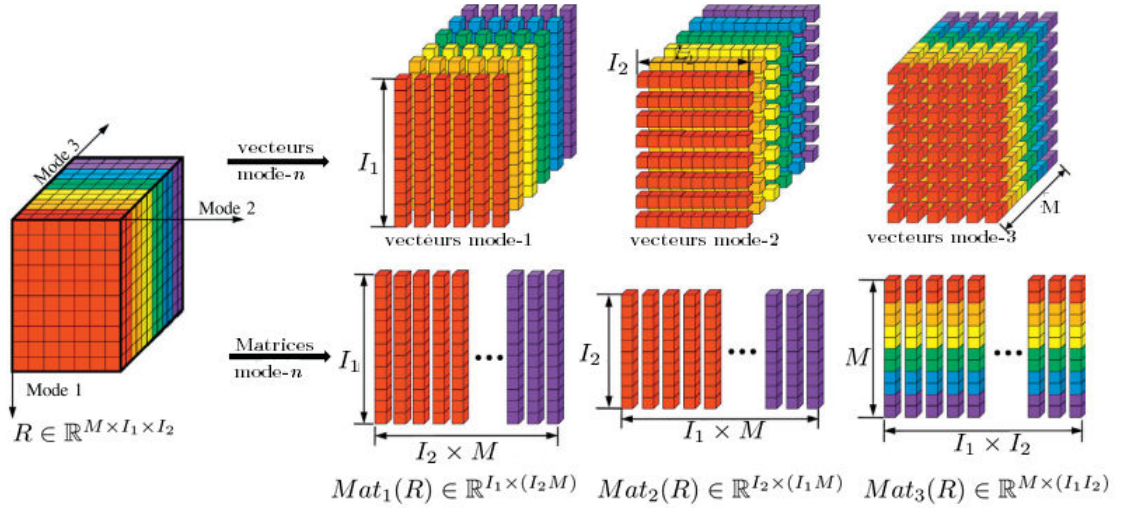


FIGURE 3.3 – Illustration de l'aplatissement avec les tenseurs en mode-3

en mode spectral (mode 3) d'un tenseur est un point très intéressant l'étude et l'analyse des données hyperspectrales par rapport aux matrices de pliage des modes spatiaux (mode-1 et mode-2). En effet, la matrice de pliage dans le mode spectral permet d'avoir une représentation physique concrète des données spectrales dans lesquelles chaque colonne de la matrice de pliage représente le spectre d'un pixel, contrairement aux deux matrices de pliage des modes spatiaux qui sont plus difficiles à interpréter (Figure ??). En effet, la matrice en 3 modes de pliage est actuellement utilisée pour l'analyse spectrale. Cependant, l'introduction d'informations spatiales permet souvent d'augmenter les performances [Schweizer and Moura, 2001]. Habituellement, cette représentation des tenseurs utilise l'algèbre multi-linéaire et en

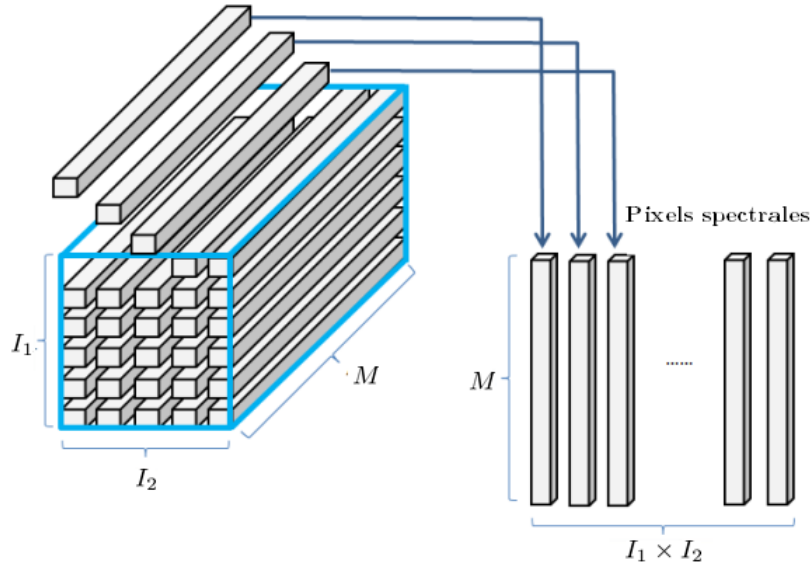


FIGURE 3.4 – Représentation tensorielle en mode-3 d'IHS

particulier les méthodes de décomposition des tenseurs et d'approximation. La décomposition du tenseur est exprimée comme suit :

$$R = C \times U^1 \times U^2 \times U^3 \quad (3.2)$$

où C est le cœur du tenseur, et U^n est la matrice de vecteurs propres associée à la matrice de covariance de mode- n : $R_n R_n^T$.

5 Édification de l'approche proposée de RD spectro-spatiale

5.1 Architecture générale

Après l'étude approfondie sur les principales techniques de la RD en s'appuyant sur leurs limites. Nous présentons dans cette section notre approche proposée. Figure 3.5 représente les trois phases de l'approche proposée ; La première phase vise à réduire l'HSI d'origine avec deux méthodes de réduction de la dimensionnalité. En fait, afin de prendre conjointement en compte l'information spatiale et spectrale, nous avons étendu la méthode LPP en intégrant le modèle tensoriel et nous avons utilisé la technique CBS [Chang and Wang, 2006] pour sélectionner les bandes pertinentes en minimisant la corrélation et la dépendance inter-bandes. La deuxième phase consiste à classifier les images réduites obtenues avec le classifieur SVM. Enfin, nous avons appliqué un processus de décision de fusion basé sur la Théorie de Dempster-Shafer (DST) pour obtenir une classification précise de l'IHS.

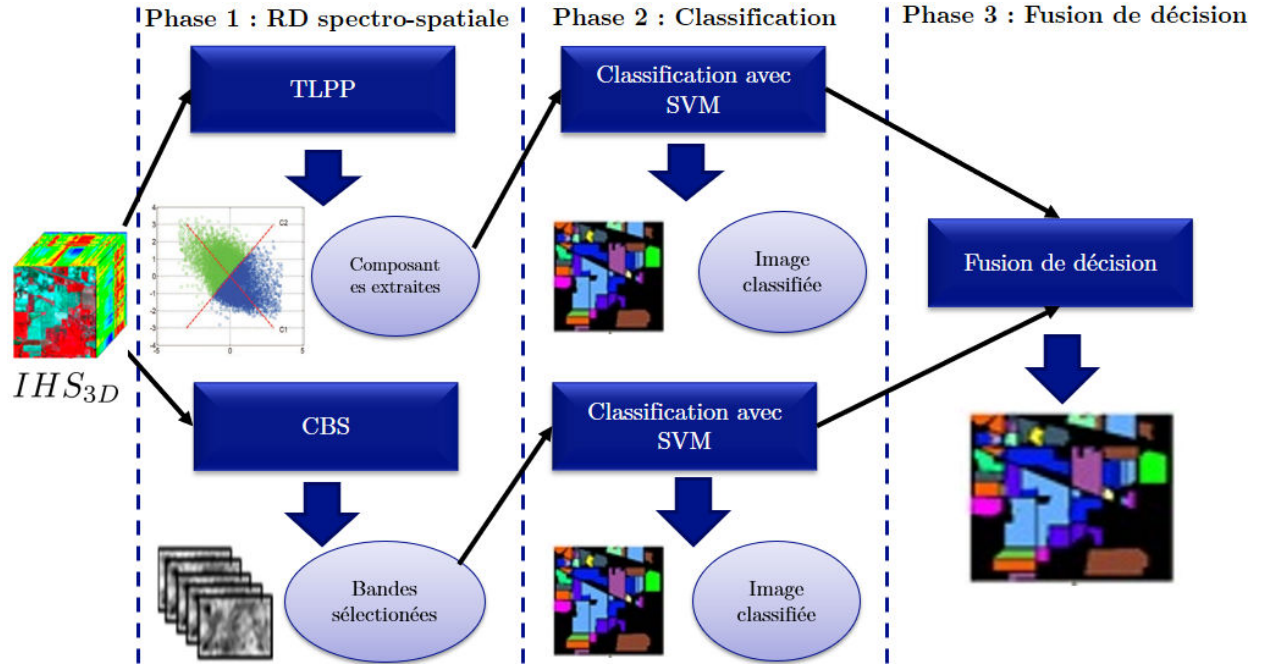


FIGURE 3.5 – Scénario générale de l'approche de RD spectro-spatiale

5.2 La projection par TLPP

Afin de conserver la structure spatiale initiale et de préserver les informations du voisinage entre les pixels de l'IHS ainsi d'exploiter les caractéristiques spectrales de, nous proposons d'utiliser l'algèbre multi-linéaire [Hemissi et al., 2013] qui est basée sur la modélisation des tenseurs dans la technique LPP. Cette approche sera considérée comme étant une étape de projection linéaire d'IHS préservant l'information spectro-spatiale. Dans ce cadre, un tenseur de *troisième ordre* (mode-3) peut représenter l'IHS entièrement. Formellement, notons que $R \in \mathbb{R}^{I_3 \times I_2 \times I_1}$ est un tenseur d'ordre-3 avec ses éléments sont arrangés comme suit : $r_{i_1 i_2 i_3}$, où $i_1 = 1, 2, \dots, I_1$; $i_2 = 1, 2, \dots, I_2$; $i_3 = 1, 2, \dots, I_3$ sont indexés sur chaque dimension. Les éléments (I_1 et I_2 représentent les deux dimensions spatiales de l'IHS et I_3 représente la troisième dimension spectrale (le nombre de bandes spectrales de l'IHS). Le modèle de tenseur peut être utilisé en exploitant la technique LPP [Wang and He, 2011] dont le but est de préserver l'information spatiale et spectrale pour améliorer le taux de classification d'IHS. Par conséquent, la RD peut être représentée en utilisant le modèle de tenseur comme suit :

$$Y = R \times T \quad (3.3)$$

où Y est le tenseur réduit d'ordre 3 ($Y \in \mathbb{R}^{I_1 \times I_2 \times p}$) contenant les p composantes. En effet, cette approche consiste à chercher la représentation de dimensionnalité faible en minimisant la somme pondérée des distances carrées entre les pixels de l'IHS voisins à l'aide d'un graphe laplacien comme suit :

1. Construire un graphe indirect $G = (X, \varepsilon)$ où les sommets représentent les

pixels et les arcs sont définis en fonction de la proximité entre les sommets (pixels).

2. Calculer les poids pour les arcs ε du graphe G en utilisant l'une des formules suivantes :
 - $W_{i,j} = \exp(-|x_i - x_j|^2 / t)$ si arc existe entre x_i et x_j avec $t \in R$ as a heat kernel.
 - $W_{i,j} = 1$ si les sommets x_i et x_j sont connecté, en utilisant le noyau chaud *heat kernel* $t = 1$, et $W_{i,j} = 0$ sinon.
3. Calculer les plus petites valeurs propres et vecteurs propres du problème du vecteur :

$$XLX^T a = \lambda XD X^T a \quad (3.4)$$

où D est la matrice diagonale de degré pondérée définie par $D_{i,i} = \sum_j W_{ij}$ et $L = D - W$ est la matrice laplacienne qui est une matrice symétrique positive. Étant donné le vecteur colonne a_0, a_1, \dots, a_{m-1} est la solution de l'équation (3.4), ordonnée selon leurs valeurs propres, $\lambda_0 < \lambda_1 < \dots < \lambda_{m-1}$. Ainsi, l'intégrateur (*embedding*) est défini comme suit :

$$x_i \rightarrow y_i = A^T x_i \quad A = (a_0, a_1, \dots, a_{m-1}) \quad (3.5)$$

avec y_i est un vecteur de m -dimensions et A est une matrice de taille $n \times m$. Par conséquent, lorsque le IHS est remodelée, la nouvelle formule basée sur la modélisation des tenseurs est définie comme suit :

$$Y_{LPP} = A^{(3)T} \times R \quad (3.6)$$

avec Y_{LPP} est un tenseur d'ordre 3 avec ($Y_{LPP} \in \mathbb{R}^{p \times I_1 \times I_2}$) contenant les p composantes et $A^{(3)T}$ représentent les vecteurs propres des composantes.

5.3 La sélection de bandes spectrales avec CBS

La technique CBS peut être considérée comme un problème de minimisation qui consiste à sélectionner des bandes tout en minimisant la corrélation ou la dépendance entre ces différentes bandes spectrales. Formellement, étant donnée $\{B_l\}_{l=1}^L$ l'ensemble de toutes les bandes spectrales originales de l'IHS, où L est le nombre total de bandes spectrale. Chaque bande B_l peut être représentée par un vecteur colonne de dimension MN désigné par b_l . Soit w_l un vecteur colonne de dimension MN qui est utilisé pour spécifier le filtre de réponses d'impulsions finies FIR (*Finite Impulse Response*) qui est conçu pour le vecteur de bande spectrale b_l et y_l soit la sortie du filtre spécifié par :

$$y_l = w_l^T b_l \quad (3.7)$$

La sortie de filtre des moindres carrés moyennes est donnée par :

$$\left(\frac{1}{L}\right) \sum_{l=1}^L y_l^2 = \left(\frac{1}{L}\right) (w_l^T b_l) (w_l^T b_l)^T = (w_l^T) \left(\left(\frac{1}{L}\right) \sum_{l=1}^L b_l b_l^T\right) w_l \quad (3.8)$$

avec $Q = (\frac{1}{L}) \sum_{l=1}^L b_l b_l^T$ désigne la matrice de corrélation des bandes spectrales. Un problème d'optimisation similaire à la minimisation d'énergie contraint CEM (*Constrained Energy Minimization*) peut être obtenue pour le problème CBS comme suit :

$$\begin{cases} \min_{w_l} \{w_l^T Q w_l\} \\ \text{subject to : } b_l^T w_l = 1 \end{cases} \quad (3.9)$$

la solution est donnée par :

$$w_l^{CEM-CBS} = (b_l^T Q^{-1} b_l)^{-1} Q^{-1} b_l \quad (3.10)$$

Dans ce travail, nous avons utilisé la technique CBS en exploitant le critère de minimisation de corrélation de bandes *BCM* (Band Correlation Minimization) pour minimiser la corrélation entre les différentes bandes spectrales vu que les bandes spectrales de l'IHS sont fortement corrélées. L'erreur de moindre carré LSE (*least-square error*) du $w_l^{CEM-CBS}$ est $\rho_l = (w_l^{CEM-CBS})^T Q w_l^{CEM-CBS}$, qui représente la corrélation minimale du vecteur du bande spectrale b_l avec l'IHS entière au sens du moindre carré en contraignant la bande spectrale b_l . Cela implique que si ρ_l est très élevé est que la bande spectrale b_l est très discriminative.

5.4 La classification spectro-spatiale avec SVM

Dans cette phase, nous utilisons le classifieur SVM linéaire [Burges, 1998] pour classifier les deux images réduites : ensemble de bandes spectrales sélectionnées Y_{CBS} et les différentes composantes Y_{TLPP} obtenues avec les techniques CBS et TLPP respectivement. Usuellement, le classificateur SVM caractérise les classes en utilisant un critère géométrique. Formellement, considérons un ensemble d'apprentissage composé de N vecteurs de l'espace des caractéristiques de d -dimensions ($x_i \in \mathbb{R}^d, i = 1, 2, \dots, N$), avec des classe étiquetées $y_i \in \{+1, -1\}$. En effet, le classifieur SVM est représenté par la fonction $f(x, \alpha) \rightarrow y$, où α désigne les paramètres du classifieur. L'idée générale du SVM vise à trouver un hyperplan optimal qui maximise la distance entre les échantillons d'apprentissage à l'hyperplan de sorte que :

- Les échantillons avec différentes étiquettes sont situés de chaque côté de l'hyperplan.
- La distance des vecteurs les plus proches de l'hyperplan de chaque côté est maximale. Ces vecteurs sont appelés *supports vectors* ainsi la distance permet de les séparer à l'hyperplan de la marge optimale.

Généralement, l'hyperplan est défini par $w.x + b = 0$, où w et b sont les paramètres de l'hyperplan. Les vecteurs qui ne sont pas sur cet hyperplan vérifient : $w.x + b \geq 0$. Ainsi, le SVM peut être défini comme $f(x, \alpha) = \text{sgn}(w.x + b)$. Pour trouver un tel hyperplan, les paramètres b et w devront être estimés pour que :

$$y_i(w.x_i + b) > 0, \quad \text{with } i = 1, 2, \dots, N \quad (3.11)$$

La distance entre l'échantillon d'apprentissage le plus proche et l'hyperplan de séparation peut être écrit comme $1/||w||$ avec une simple renormalisation des paramètres de l'hyperplan w et b tels que

$$\min_{i=1,2,\dots,N} \{y_i(w.x_i + b)\} \geq 1 \quad (3.12)$$

Par conséquent, la marge géométrique entre les deux classes est égale à $2/||w||$. En effet, la maximisation de la marge conduit au problème d'optimisation suivant :

$$\begin{cases} \text{minimize : } \frac{1}{2}||w||^2 \\ \text{subject to : } y_i(w.x_i + b) \geq 1, \quad i = 1, 2, \dots, N \end{cases} \quad (3.13)$$

Dans cette phase, nous avons utilisé le classifieur SVM linéaire pour classifier les deux images réduites obtenues avec les techniques TLPP et CBS. Cependant, ces images classifiées avec SVM sont imprécises et incertaines. Par conséquent, nous proposons d'appliquer un processus de fusion [Farah et al., 2008a] afin d'aboutir à une bonne classification avec taux de classification performant et très élevé. La section suivante décrit le processus de décision de fusion utilisé.

5.5 La fusion de classification avec la théorie de Dempster-Shafer

La première phase de notre approche proposée a été réalisée pour réduire la dimensionnalité de l'IHS tout en préservant autant que possible l'information spectro-spatiale en se basant sur les deux techniques TLPP et CBS. Le classifieur SVM linéaire a été utilisé pour classifier les deux images réduites Y_{TLPP} et Y_{CBS} indépendamment. Afin d'exploiter la complémentarité d'informations fournies par les deux images classifiées et d'obtenir une classification satisfaisante de l'IHS, nous proposons à fusionner ces deux images classifiées en utilisant la théorie de Dempster-Shafer DST (*Dempster Shafer Theory*) [Farah et al., 2008b, Fontani et al., 2013]. La DST est considérée comme une généralisation de la théorie bayésienne. Elle consiste à combiner des informations évidentielles provenant de multiples sources indépendantes. Plus précisément, elle est utilisée pour combiner des fonctions de masse. Considérons que l'espace des hypothèses est désigné par Θ , qui correspond dans les applications de classification d'images à l'ensemble des classes d'occupation du sol. La DST calcule une fonction de masse pour chaque sous-ensemble $A \subset \Theta$ ($A \in 2^\Theta$). Pour les problèmes de classification, cela signifie que non seulement les classes individuelles sont considérées (aussi appelés singletons), mais aussi toute union de classes. Dans ce qui suit, les hypothèses sur les singletons et les hypothèses sur les unions de classes sont respectivement appelées hypothèses simples et hypothèses composées. Plus précisément, elle fournit une représentation à la fois de l'imprécision et de l'incertitude par la définition de deux fonctions : la plausibilité et la croyance, qui sont dérivées d'une fonction de masse :

$$\begin{cases} m : 2^\Theta \rightarrow [0, 1] \\ X \rightarrow m(X) \end{cases} \quad (3.14)$$

avec

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{X \in 2^\theta} m(X) = 1 \end{cases} \quad (3.15)$$

Dans notre cas, nous avons deux images classifiées C_{TLPP} et C_{CBS} , et 16 classes C_1 à C_{16} . Pour le calcul des masses d'éléments focaux, nous procédons comme suit : la masse des classes simples est égale aux probabilités conditionnelles a posteriori, alors que la masse des classes composées est égale à la somme des masses des classes simples incluses dans l'union. De plus, en utilisant les règles de DST, nous avons :

$$m_{1,2}(D) = (m_1 \oplus m_2)(D) = \frac{1}{1-K} \sum_{A \cap B = D \neq \phi} m_1(A)m_2(B) \quad (3.16)$$

où m_1 et m_2 sont respectivement les fonctions de masse correspondant à C_{TLPP} et C_{CBS} et

$$K = \sum_{A \cap B \neq \phi} m_1(A)m_2(B) \quad (3.17)$$

L'algorithme 1 résume les différentes phases présentées dans l'approche proposée :

Algorithm 1 Algorithme proposé de RD specro-spatiale d'IHS

Require: HSI, GT : Vérité terrain

Ensure: Image fusionnée classifiée

- 1: $Im \leftarrow \text{ReadImage}(HSI)$
 - 2: **Réduction de dimension spectro-spatiale**
 - 3: *Méthode de TLPP*
 - 4: $Y_{TLPP} \leftarrow TLPP(Im)$
 - 5: *Méthode de CBS*
 - 6: $Y_{CBS} \leftarrow CBS(Im)$
 - 7: **Classification avec SVM**
 - 8: $C_{LPP} \leftarrow SVM(Y_{TLPP}, GT)$
 - 9: $C_{CBS} \leftarrow SVM(Y_{CBS}, GT)$
 - 10: **Fusion de décision avec DST**
 - 11: $C_{DST} \leftarrow DST(C_{LPP}, C_{CBS})$
-

5.6 Synthèse et discussion

Dans un premier travail, nous avons proposé une nouvelle approche appelée extraction des caractéristiques basée sur le modèle de tenseur et la projection préservant la localité TLPP (*features extraction based on Tensor modelling and locality preserving projections*) qui a été utilisée pour maintenir la distribution spatiale des pixels dans l'IHS originale et qui est actuellement négligée dans la plupart des techniques de projection linéaire. Par la suite, nous avons utilisé la technique CBS pour

6. Approche de RD adaptative pour l'interprétation sémantique d'IHS

sélectionner les bandes spectrales discriminatives et décorréées à partir de l'IHS originale. Nous avons implémenté le classifieur SVM en utilisant les deux techniques de RD (TLPP et CBS). Les résultats de classifications obtenues sont fusionnées en utilisant la technique de DST afin de considérer et modéliser l'incertitude entachée par ces deux images classifiées. La classification fusionnée s'est révélée efficace par rapport aux techniques classiques telles que TPCA, TICA et TLPP. Cette approche peut être appliquée dans d'autres domaines d'applications d'IHS comme la détection de cibles, etc.

6 Approche de RD adaptative pour l'interprétation sémantique d'IHS

6.1 Motivations

Dans un premier travail dans cette thèse, nous avons essayé de proposer une approche de RD spectro-spatiale qui vise d'exploiter la complémentarité d'informations des techniques de projection et de sélection de bandes en utilisant le tenseur et les modèles d'algèbre multi-linéaire [Sellami and Farah, 2016]. Nous visons dans une deuxième partie d'étendre cette approche de RD spectro-spatiale en améliorant le modèle proposé. Par conséquent, l'approche proposée permet uniquement d'avoir une image réduite composite de deux images réduites obtenues par TLPP et CBS. Cependant, les composantes réduites obtenues par TLPP sont incertaines et imprécises, mais les nouvelles caractéristiques peuvent être utiles pour l'interprétation sémantique d'IHS, comme les nouveaux points de données des nuages de points peut représenter les classes de l'image originale avec une certaine signification sémantique. De plus, les bandes spectrales sélectionnées par la technique CBS sont uniquement discriminatives et ne sont pas informatives. En d'autres termes, l'ensemble de bandes spectrales obtenu par CBS représente des bandes pertinentes qui sont discriminantes et aussi décorréées. Par contre, ces bandes spectrales ne sont pas informatives, c'est à dire elle ne fournissent pas des informations relatives au contexte, notamment les classes de l'images. D'une manière générale, nous avons proposé dans la première partie du travail de cette thèse, une approche hybride entre deux différentes techniques de RD mais qui n'est pas adaptative selon le contexte. Nous visons dans une deuxième partie d'étendre cette approche et la rendre adaptative et générique qui ne nécessite qu'un minimum de connaissances *a priori*. L'idée générale de la deuxième approche est d'exploiter des techniques d'extraction des connaissances pour avoir une RD adaptative pour améliorer l'interprétation des données hyperspectrales. Dans la section suivante, nous détaillons l'approche proposée de RD adaptative en mettant l'accent sur la phase d'extraction des connaissances à partir de la projection et la sélection de bandes discriminatives et informatives en exploitant des connaissances sémantiques.

6.2 Architecture générale

La technique de projection TLPP est une approche basée sur les tenseurs exploitant à la fois l'information spectrale et spatiale de l'IHS. Son objectif est double ; le premier consiste à préserver la distance spectrale entre les pixels voisins de l'IHS et le deuxième vise à incorporer le réarrangement spatial, notamment la structure spatiale dans la RD. Par conséquent, dans la projection obtenue par TLPP des nouvelles caractéristiques sont construites en exploitant la totalité de bandes spectrales ce qui rend leur interprétation et exploitation facile. Néanmoins, que cette transformation provoque une perte d'informations dû la déformation du sens physique et de changement de l'espace d'origine. Bien qu'il y a une perte d'informations au niveau cette projection par TLPP, mais ces caractéristiques sont considérées comme significatives et leurs informations sont riches et qui peuvent être utiles pour la sélection de bandes spectrales pertinentes. Pour cela, les caractéristiques fournies par TLPP seront exploitées comme étant des connaissances qui seront injectées par la suite dans le processus de sélection de bandes spectrales afin de rendre l'approche adaptative et générique ce qui permettra de sélectionner automatiquement les bandes spectrales à la fois discriminantes et informatives selon le contexte sans aucune connaissance *a priori*. Cela évite le problème de la malédiction de dimensionnalité qui exige un nombre important d'échantillons étiquetés dans la sélection de bandes spectrales informatives. Formellement, considérons $X = [x_1, \dots, x_N] \in \mathbb{R}^{M \times N}$ est l'ensemble de données de l'IHS, où x_i représente un vecteur de réponse spectral pour un pixel de l'IHS. M est le nombre de toutes les bandes spectrales et N est le nombre de pixels $i \in \{1, 2, \dots, N\}$ dans une seule bande $\{B_m\}_{m=1}^M$. L'objectif de cette approche est de sélectionner p bandes spectrales à partir de X ($p \ll W$ de manière adaptative et automatique. En d'autres termes, il s'agit de sélectionner automatiquement les bandes discriminatives et informatives sans connaissances *a priori*. De plus, cette approche doit s'affranchir au problème de la non disponibilité des échantillons étiquetés au préalable ainsi d'introduire le nombre de bandes à sélectionner. La figure 3.6 montre l'approche proposée de sélection adaptative de bandes spectrales.

6.3 Extraction des connaissances à partir des caractéristiques

Usuellement, l'objectif de cette phase est d'extraire des connaissances à partir des données projetées avec la technique TLPP afin d'enrichir et alimenter la sélection de bandes spectrales. Cet enrichissement peut aider la sélection de bandes spectrales par des connaissances relatives à la contribution de bandes spectrales pour la construction du nouvel sous-espace optimal. En d'autres termes, les techniques de projection permettent de fournir des nouvelles informations dans le nouvel sous-espace en combinant et exploitant toutes les bandes spectrales de l'IHS originale. Cependant, la sélection de bandes spectrale se limite à une sélection d'un sous-ensemble sans recours à exploiter la totalité de bandes spectrales. Par conséquent, l'idée a été inventée dans ce cadre afin d'enrichir la sélection de bandes par des nouvelles informations qui seront extraites à partir le nuage de points des composantes transformées par

6. Approche de RD adaptative pour l'interprétation sémantique d'IHS71

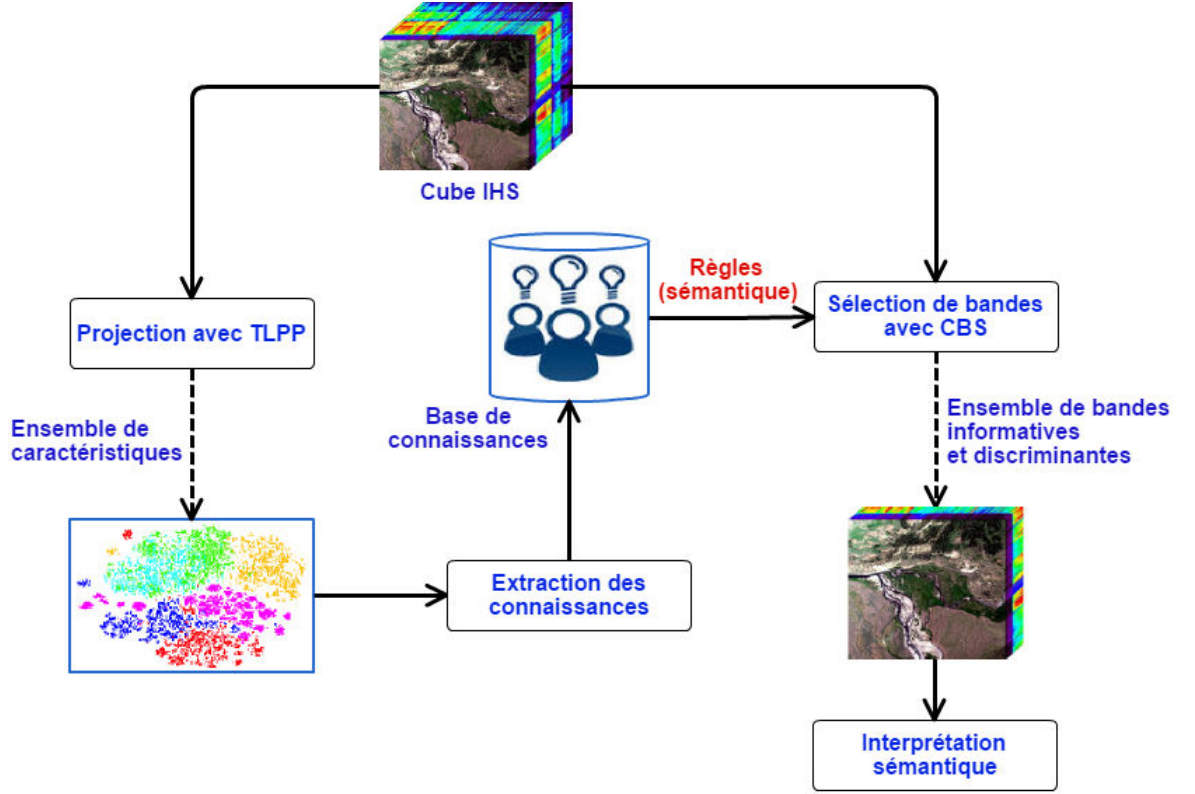


FIGURE 3.6 – Approche proposée d'interprétation sémantique basée sur une RD adaptative

la technique de TLPP. Formellement, étant donné l'ensemble de points de données N dans l'espace des caractéristiques de p -dimensions $\vec{Y}_{LPP} = [\vec{y}_1, \dots, \vec{y}_N] \in \mathbb{R}^{p \times N}$ obtenu avec la technique TLPP. Nous visons à utiliser l'algorithme de clustering K-means pour partitionner les N points de données dans K clusters afin de regrouper les points de données similaires et avoir des informations supplémentaires sur chaque cluster. En effet, k-means est considéré comme étant un algorithme de partitionnement rapide et facile à mettre en œuvre [Kanungo et al., 2002]. Son objectif est de minimiser la distance euclidienne carrée moyenne de chaque point de données à son centroid le plus proche. Le centroid $\vec{\mu}$ dans un cluster C peut être calculé comme suit :

$$\vec{\mu} = \frac{1}{|C|} \sum_{\vec{y} \in C} \vec{y} \quad (3.18)$$

De plus, nous pouvons étiqueter de façon automatique les différents points de données en utilisant les vecteurs des centroides obtenus $\vec{\mu} = [\vec{\mu}_1, \dots, \vec{\mu}_K]$. Par conséquent, pour effectuer un étiquetage des échantillons, nous calculons la distance euclidienne

entre les échantillons y et les centroides μ comme suit :

$$dist(y, \mu) = \sqrt{\sum_{n=1}^p (y_n - \mu_n)^2} \quad (3.19)$$

En outre, afin de trouver l'ensemble des composantes qui contribuent pour la discrimination de chaque classe, nous utilisons le classifieur SVM linéaire pour générer un ensemble de règles [Pinto et al., 2016, Diederich, 2008]. Plus précisément, ces règles visent à calculer les contributions des différentes caractéristiques extraites Y_{LPP} . Formellement, désignons par T l'ensemble des données d'apprentissage composé de l'espace des caractéristiques de p - dimensions ($t_i \in \mathbb{R}^p, i = 1, \dots, N$), avec les étiquettes des classes $C_j \in \{C_1, \dots, C_K\}$. Dans ce cas, SVM cherche à concevoir un hyperplan qui classe tous les vecteurs d'apprentissage. Usuellement, l'hyperplan de SVM est défini par $g(\vec{Y}) = \vec{\omega} \cdot \vec{Y} + \omega_0$, où $\vec{\omega}$ est le vecteur des poids (*support vectors*). Par conséquent, nous pouvons générer un ensemble de règles comme suit :

$$IF \ \vec{\omega}_1 \cdot \vec{Y}_1 + \dots + \vec{\omega}_p \cdot \vec{Y}_p + \omega_0 \ THEN \ CLASS \quad (3.20)$$

où p est le nombre des caractéristiques.

Figure 3.7 illustre un exemple de génération de règles. Bien que $\vec{\omega}$ est minimale, la séparabilité des classes sera maximale. La minimisation de $\vec{\omega}$ est une tâche d'optimisation non linéaire, résolue par les conditions de Karush-Kuhn-Tucker (KKT) [Diederich, 2008] en utilisant la théorie de *Lagrange multipliers* λ_i :

$$\begin{cases} \vec{\omega} = \sum_{i=0}^N \lambda_i t_i \vec{Y}_i \\ \sum_{i=0}^N \lambda_i t_i = 0 \end{cases} \quad (3.21)$$

L'extraction des connaissances à partir TLPP en utilisant K-means/SVM une étape

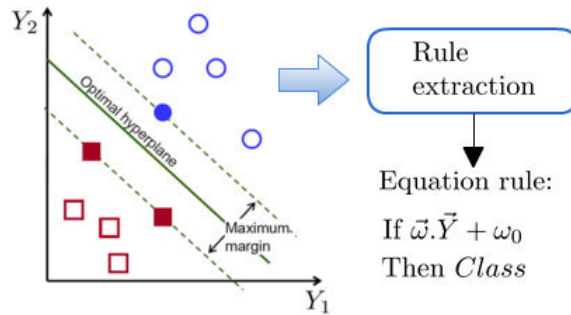


FIGURE 3.7 – Les règles générées par SVM

permettant d'extraire des informations sémantiques relatives à l'importance de composantes pour la discrimination des classes de données hyperspectrales. l'algorithme k-means permet de regrouper les nouveaux points de données sous forme des clusters qui vont être utilisés par la suite avec le classifieur SVM pour générer des règles. Ces règles sont considérées comme des connaissances sémantique permettant de fournir

6. Approche de RD adaptative pour l'interprétation sémantique d'IHS

Algorithm 2 Extraction des connaissances à partir de TLPP

Require: $Y_{TLPP} = [y_1, \dots, y_N]$, K : integer

Ensure: \vec{E}_R : Règles extraites

```

1: Étape 1 : Clustering et préparation des échantillons
2: procedure K-MEANS ( $Y_{TLPP}, K, \vec{\mu}$ )
3: Étape 2 : Étiquetage des échantillons
4: for  $i = 1 : N$  do
5:    $min \leftarrow dist(y_i, \mu_1)$  ▷ Eq.(3.19)
6:    $ind \leftarrow 1$ 
7:   for  $j = 2 : K$  do
8:     if  $min > dist(y_i, \mu_j)$  then
9:        $min \leftarrow dist(y_i, \mu_j)$ 
10:       $ind \leftarrow j$ 
11:     end if
12:   end for
13:    $t(i) \leftarrow ind$ 
14: end for
15: Étape 3 : Extraction des règles
16: procedure SVM( $Y_{TLPP}, T, \vec{E}_R$ ) ▷ Règles extraites

```

des informations significatives sur la contribution des composantes et des bandes spectrales pour les classes de l'IHS. Ceci permet de pouvoir sélectionner de manière automatique et adaptative les bandes spectrales discriminatives, distinctives et informatives pertinentes. Cette étape est décrit dans la section suivante.

6.4 Sélection adaptative des bandes spectrales pertinentes sous contraintes sémantiques

La sélection de bandes spectrales CBS est une technique non supervisée permettant de chercher les bandes discriminantes en minimisant un certain critère, notamment la corrélation (BCM) et la dépendance (BDM). L'ensemble de bandes spectrales sélectionnés par CBS se sont des bandes spectrales pertinentes discriminantes, mais ne sont pas informatives. En d'autres termes, la technique de CBS ne permet pas d'avoir des bandes spectrales portant des informations relatives au contexte, comme l'urbain, la végétation, etc. Par conséquent, afin de réduire la taille massive des données hyperspectrales en évitant les problème de leurs exploitations et analyse, nous visons à améliorer la technique CBS en injectant des informations sémantiques issues par les règles extraites par SVM. Ces connaissances vont guider la CBS à trouver les bandes spectrales pertinentes qui sont discriminantes, décorréliées, informatives et contextuelles. Formellement, notons par \vec{E}_R les règles extraites dans la deuxième phase de l'approche proposée. En se basant sur ces règles, nous tentons à calculer des scores permettant d'évaluer la pertinence des caractéristiques Y_i au terme de discrimination des différentes classes de l'IHS C_j qui sont définis

comme suit [Bai et al., 2015] :

$$score(Y_l/C_j) = \max_{i=1..R} \{\vec{E}_R(i, l) | \vec{E}_R(i, p+1) = j\} \quad (3.22)$$

où l est l'index des caractéristiques. Par conséquent, pour chaque classe nous pouvons détecter les caractéristiques pertinentes en utilisant cette formule de score. En effet, nous nous intéressons dans cette étape à sélectionner automatiquement p bandes spectrales pertinentes qui sont à la fois discriminatives et informatives à partir de M bandes spectrales de l'IHS d'origine se basant sur les règles extraites \vec{E}_R . Ceci permet de préserver le sens physique des données hyperspectrales. Par conséquent, en fonction du calcul du score de caractéristiques Y_l et de l'équation (3.6), l'élément F^{th} de caractéristique y_{lf} est obtenu par

$$y_{lf} = \sum_{m=1}^M a_{lm} r_{mf}, \quad l = 1, \dots, p \quad (3.23)$$

où a_{lm} représente le poids de $m^{ième}$ bande concernant la caractéristique Y_l . Cela permet de noter que le poids a_{lm} montre la quantité d'informations pertinentes fournies par la $m^{ième}$ bande à travers la $l^{ième}$ caractéristique et une classe donnée. Par conséquent, nous pouvons estimer l'importance de chaque bande spectrale pour toutes les caractéristiques Y_{lpp} en calculant le coefficient du poids absolu moyen $\omega_{b_m}^-$ comme suit [Du et al., 2003] :

$$\omega_{b_m}^- = \frac{1}{p} \sum_{l=1}^p |a_{lm}| \quad (3.24)$$

où $m = 1, \dots, M$.

En utilisant l'équation 3.24, le coefficient du poids absolu moyen peut être trié pour toutes les bandes spectrales, nous obtenons ainsi la séquence de tri suivante :

$$[\omega_{b_1}^-, \dots, \omega_{b_m}^-, \dots, \omega_{b_M}^-] \quad (3.25)$$

Par conséquent, la bande avec un poids plus élevé contribue plus aux caractéristiques Y_l que d'autres bandes spectrales. En d'autres termes, les bandes qui contribuent plus pour les différentes caractéristiques sont plus informatives et discriminatives pour les toutes classes de l'IHS. En outre, nous pouvons considérer les contributions des bandes spectrales comme étant des règles, comme il est indiqué dans l'équation (3.26) :

$$IF \ Y_l \ THEN \ \omega_{b_1}^- \geq \dots \geq \omega_{b_m}^- \geq \dots \geq \omega_{b_M}^- \quad (3.26)$$

Soit $\{SB\}_{(E_R)}$ est l'ensemble des bandes spectrales sélectionnées en fonction des règles extraites E_R et $\{SB\}_{(CBS)}$ sont les bandes sélectionnées avec la technique CBS. La sélection adaptative des bandes spectrales sous contraintes sémantiques vise à sélectionner p bandes spectrales informatives et discriminatives à partir de M bandes spectrales de l'IHS comme suit :

$$\{SB\}_{(ACBS)} = \{SB\}_{(CBS)} \cap \{SB\}_{(E_R)} \quad (3.27)$$

6. Approche de RD adaptative pour l'interprétation sémantique d'IHS75

en se basant sur l'équation 3.27, si le nombre des bandes sélectionnées avec $\{SB\}_{(ACBS)}$ est inférieure à p , c'est à dire le nombre minimal de bandes spectrales à sélectionner, nous utilisons les règles extraites dans l'équation 3.26 pour évaluer le reste des bandes spectrales en fonction de leur contribution pour toutes les caractéristiques Y_l . Cela permet de sélectionner les bandes avec un poids plus élevé jusqu'à obtenir p bandes pertinentes. l'algorithme proposé *ACBS* est présenté comme suit (Algorithme 3) :

Algorithm 3 Adaptive Constrained Band Selection (ACBS)

Require: Given $\vec{E}_R = [\vec{E}_{R_1}, \dots, \vec{E}_{R_K}]$, A : transformation matrix, p : integer
Ensure: $SB_{(ACBS)}$: Bandes sélectionnées par ACBS

```

// Calculer le score de chaque caractéristique  $Y_l$  pour chaque classe  $C_j$ 
2: //  $p + 1$  Indice d'une classe
   for  $j = 1 : K$  do
4:   for  $l = 1 : p$  do
        $scr(Y_l/C_j) \leftarrow \max_{i=1..R} \{\vec{E}_R(i, l) | \vec{E}_R(i, p+1) = j\}$ 
6:   end for
       // Chercher  $Y_l$  qui contribue plus à la classe  $C_j$ 
8:    $Y_l \leftarrow \arg \max_v \{scr(Y_v/C_j)\}$ 
       // Sélectionner les bandes contribuant plus à  $Y_l$ 
10:   $SB_{(E_R)}(j) \leftarrow \max_{m=1..M} \{|A(l, m)|\}$ 
   end for
12: // Sélectionner  $p$  bandes à partir de l'IHS avec CBS
    $SB_{(CBS)} \leftarrow procedure\_CBS(IHS, p)$ 
14: // Sélection de bandes avec ACBS
    $SB_{(ACBS)} \leftarrow SB_{(CBS)} \cap SB_{(E_R)}$ 
16: // Calculer les bandes rejetées
    $SB_{(RB)} \leftarrow SB_{(CBS)} \cup SB_{(E_R)} \setminus SB_{(ACBS)}$ 
18: // Calculer les contributions des bandes pour toutes les caractéristiques
   for  $m = 1 : M$  do
20:    $\vec{w}(m) \leftarrow (\frac{1}{p} \sum_{l=1}^p |a_{lm}|)$  ▷ Eq.(3.24)
   end for
22: while  $\#SB_{(ACBS)} < p$  do
    $B_m \leftarrow \arg \max_{B_i \in SB_{(RB)}} \{\vec{w}(i)\}$ 
24:    $SB_{(ACBS)} \leftarrow SB_{(ACBS)} \cup \{B_m\}$ 
    $SB_{(RB)} \leftarrow SB_{(RB)} \setminus \{B_m\}$ 
26: end while

```

6.5 Discussion

Dans cette approche, nous avons essayé de présenter une approche de RD adaptative permettant de sélectionner automatiquement les bandes spectrales informa-

tives et discriminatives tout en profitant des connaissances sémantiques issues par des méthodes de projection spectro-spatiales, à savoir la technique TLPP. Dans un premier lieu, les données hyperspectrales sont projetées en utilisant TLPP. Cette transformation nous a permis d'avoir des nouvelles connaissances extraites à partir du sous-espace optimal obtenu par TLPP. Dans ce cadre, des règles ont été extraites par les deux techniques de fouille de données et la découverte des connaissances, notamment k-means et SVM. K-means a été utilisé pour partitionner l'ensemble de nuage de points obtenu par TLPP afin de regrouper les points de données similaires et d'avoir des clusters discriminants. Dans une seconde étape, le classifieur SVM a été exploité pour générer des règles permettant d'évaluer l'importance de chaque composante Y_l par rapport à une classe donnée C_k . En d'autres termes, nous avons essayé d'extraire des règles à partir des composantes extraites par TLPP et les clusters obtenus par k-means. Les clusters de k-means ont été exploités comme étant des classes dans SVM. Ceci permettra d'avoir une approche générique et adaptative ne nécessite qu'un minimum de connaissances *a priori*. Par la suite ces règles ont permis d'aider la technique de sélection de bandes CBS d'avoir des connaissances sur les bandes spectrales par rapport aux classes de l'IHS. Par conséquent, l'approche de sélection proposée permet de sélectionner les bandes spectrales discriminatives en utilisant les deux critères de CBS, à savoir la corrélation et la dépendance et le critère de l'information sémantique qui est les règles extraites à partir TLPP pour aussi pouvoir sélectionner les bandes informatives avec un nombre minimal de bandes spectrales pertinentes.

7 Conclusion

La technologie hyperspectrale est riche au niveau spectral mais son exploitation est un grand défi jusqu'au aujourd'hui. Par conséquent, afin de rendre les données hyperspectrales facilement exploitable et interprétable, il est désirable de réaliser un meilleur compromis entre la réduction et la préservation de la quantité d'information acquise par l'IHS originale. Le principal objectif de ce travail a été inscrit dans le cadre général de développement d'une nouvelle approche adaptative de réduction et d'interprétation sémantique d'IHS. Une nouvelle voie dans la réduction et la sélection de l'information spectrale pertinente d'IHS a été touchée dans ce travail. Son objectif est d'exploiter ces données massives et palier essentiellement à de nombreux problèmes tels que la préservation de l'information spectro-spatiale, la sélection de bandes spectrales les plus pertinentes pour la discrimination des objets de la scène de l'IHS. L'approche proposée dans ce manuscrit de thèse est une approche de RD adaptative par rapport au domaine d'application qui améliore l'interprétation sémantique d'IHS qui ne nécessite qu'un minimum de connaissances *a priori*.

Résultats expérimentaux

Sommaire

1	Introduction	78
2	Jeux de données	78
2.1	Indian Pines	78
2.2	Pavia University	79
3	Approche de RD spectro-spatiale	81
3.1	La RD spectro-spatiale : modèle de tenseur	81
3.2	Classification avec SVM	82
3.3	Fusion de décision avec DST	84
3.4	Évaluation de temps de calcul	86
3.5	Synthèse et discussions	87
4	Approche de RD adaptative sémantique d'IHS	88
4.1	Techniques de comparaison	89
4.2	Ajustement des paramètres	89
4.3	Extraction des règles	89
4.4	Résultats de classification et analyse de performance de la sélection de bandes spectrales	91
4.5	Impact de la sélection aléatoire des échantillons d'apprentissage sur la sélection de bandes spectrales	95
4.6	Interprétation des bandes spectrales sélectionnées	98
4.7	Etude de l'impact de la variation du paramètre K dans l'algorithme K-means	99
4.8	Synthèse et discussions	99
5	Conclusion	100

1 Introduction

2 Jeux de données

Les résultats expérimentaux sont démontrés sur deux données hyperspectrales capturées par les capteurs AVIRIS (*Airborne Visible/ Infrared Imaging Spectrometer*) et ROSIS (*Reflective Optics System Imaging Spectrometer*), avec différents contextes (zones urbaines et de végétation), différentes résolutions spatiales (1.3 m et 20 m et différentes résolutions spectrales : le nombre de bandes spectrales de 120 à 220 bandes spectrales. Ces deux données hyperspectrales sont détaillées dans la section suivante.

2.1 Indian Pines

L'image Indian Pines est une zone de végétation située au nord-ouest d'indiana pines, qui est issue par le capteur AVIRIS en 1992. Ses dimensions spatiales de 145×145 pixels et une résolution spatiale de 20 m par pixel. Chaque pixel est représenté dans 220 bandes spectrales couvrant la plage de 375 – 2500 nm. La vérité terrain de cette IHS contient 16 classes étiquetées. L'avantage de cette image que la vérité terrain est disponible, ce qui permet de l'exploiter pour l'apprentissage et le test. La figure 4.1 montre une image composite colorée de l'IHS avec l'image de la vérité terrain. Le tableau 4.1 montre les différentes classes de vérité terrain de

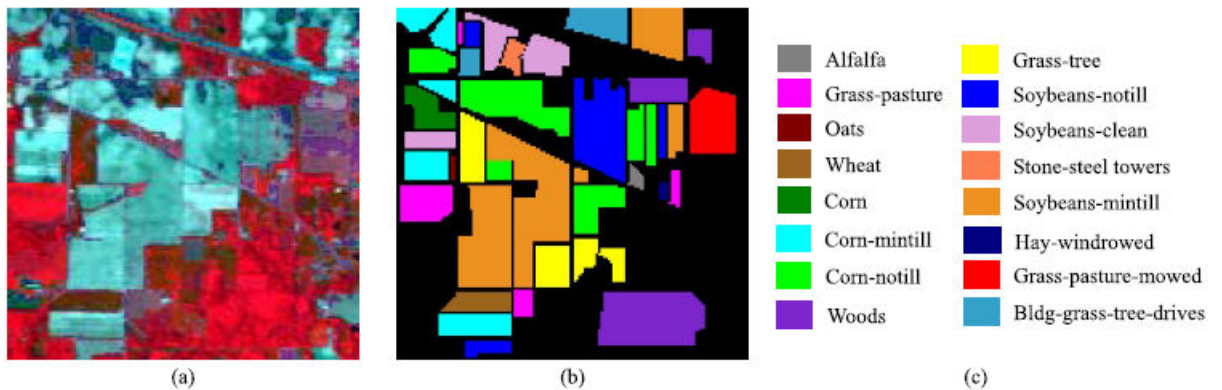


FIGURE 4.1 – Image Indian Pines. (a) : image couleur composite, (b) image de vérité terrain, (c) référence des classes

l'Indian Pines et leurs distributions pour l'apprentissage et le test. La figure 4.2 représente les différentes signatures spectrales de 16 classes de la vérité terrain de l'IHS de l'Indiana Pines. Chaque classe est représenté dans un vecteur de dimension 200 bandes spectrales .

Tableau 4.1 – Classes de vérité terrain et leurs distributions (Indian Pines)

Classes		Nombre d'échantillons
No	Nom	
1	Alfalfa	54
2	Corn-notill	1334
3	Corn-min	834
4	Corn	234
5	Grass/Pasture	497
6	Grass/Trees	704
7	Grass/Pasture-mowed	26
8	Hay-windrowed	499
9	Oats	20
10	Soybeans-notill	968
11	Soybeans-min	2468
12	Soybeans-clean	614
13	Wheat	212
14	Woods	1294
15	Building-Grass-Trees-Drives	380
16	Stone-steel Towers	95
Total		10366

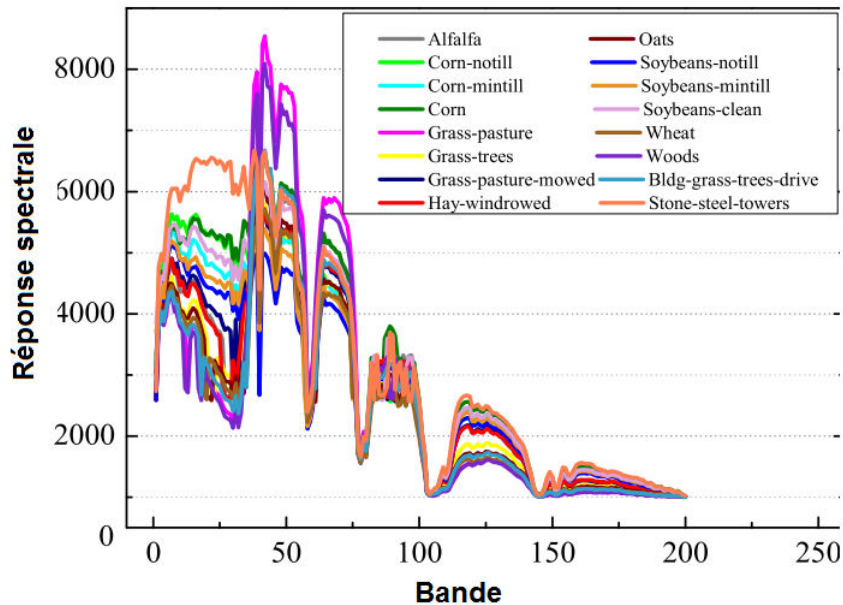


FIGURE 4.2 – Les différentes signatures spectrales de 16 classes de l'Indiana Pines

2.2 Pavia University

Cet ensemble de données hyperspectrales a été utilisé à plusieurs reprises. Il a été capturé sur la ville de Pavia, en Italie, par l'instrument aéroporté d'imagerie par spectrophotographie optique (ROSIS-03). Le vol sur la ville de Pavia a été exploité par Deutschen Zentrum für Luft- und Raumfahrt (DLR, l'Agence aérospatiale allemande) dans le cadre du projet HySens, géré et parrainé par l'Union européenne. L'IHS Pavia University représente une zone urbaine de l'université Pavia en Italie.

Elle est capturée par le capteur ROSIS-03 (*Reflective Optics Spectrographic Imaging System*). Les dimensions spatiales sont de taille 610×340 pixels et la résolution spatiale est de 1.3 m par pixel. Le nombre de bandes spectrales de l'image acquise est de 115 bandes spectrales avec une plage de 430 à 860 nm . L'ensemble de données hyperspectrales contient l'école d'ingénieur de l'université de Pavia et constitue de différentes classes : constructions, matériaux et arbres. La vérité terrain contient 9 classes d'intérêts. Figure 4.3 montre l'image couleur composite et la vérité terrain de cette IHS.

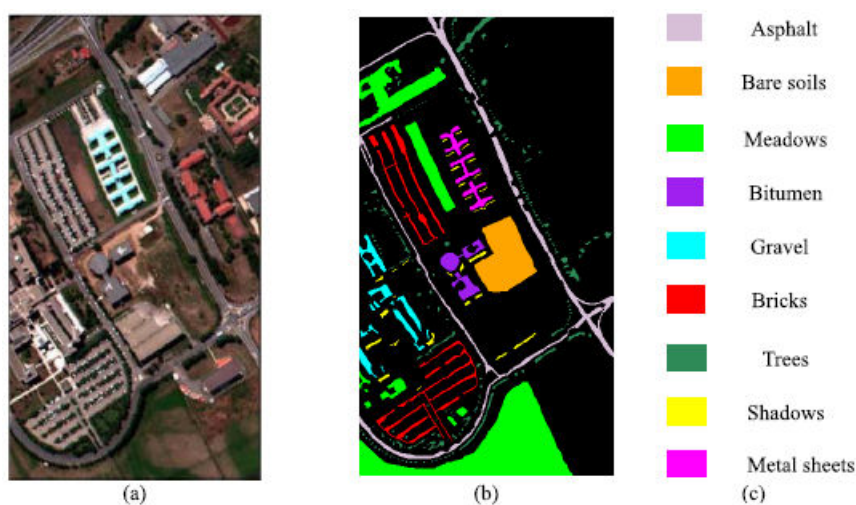


FIGURE 4.3 – Image Pavia University. (a) : image couleur composite, (b) image de vérité terrain, (c) référence des classes

Le tableau 4.2 montre les différentes classes de vérité terrain de Pavia University et leurs distributions pour l'apprentissage et le test. La figure 4.4 représente les

Tableau 4.2 – Classes de vérité terrain et leurs distributions (Pavia University)

Classes		Nombre d'échantillons
No	Nom	
1	Asphalt	6631
2	Bare-S	4529
3	Bitumen	1330
4	Gravel	2099
5	Meadows	18694
6	Painted-M	1345
7	Self-B	3682
8	Shadows	947
9	Tree	3064
Total		42321

différentes signatures spectrales de 16 classes de la vérité terrain de l'IHS de l'Indiana Pines. Chaque classe est représenté dans un vecteur de dimension 200 bandes spectrales .

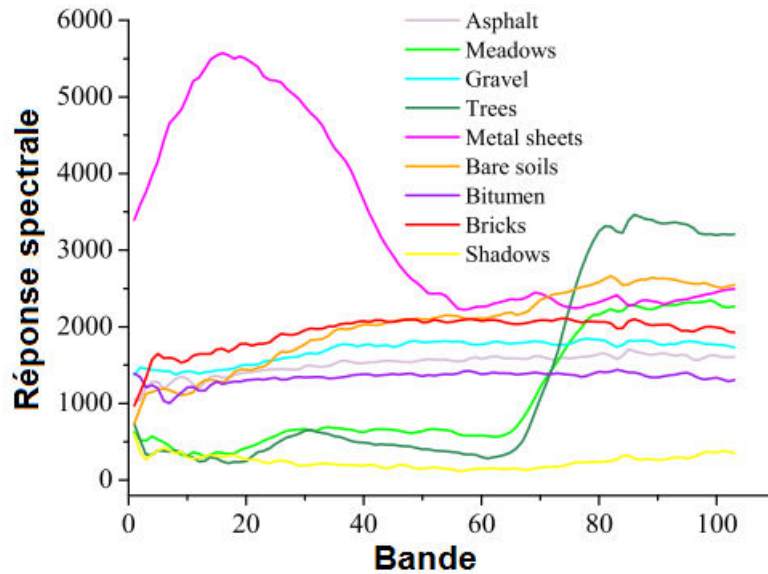


FIGURE 4.4 – Les différentes signatures spectrales de 9 classes de Pavia University

3 Approche de RD spectro-spatiale

Dans cette section, nous présentons les résultats expérimentaux pour tester et évaluer la performance et l'efficacité de l'approche proposée de la réduction spectro-spatiale pour la classification haut-niveau. Nous comparons tout d'abord notre approche proposée par d'autres méthodes de réduction de dimensionnalité à savoir la technique Analyse en Composante Principale avec Tenseur (TPCA), et l'Analyse en Composante Indépendante avec Tenseur (TICA), etc. Pour valider cette approche nous avons utilisé comme données une image hyperspectrale de type AVIRIS. Par la suite, nous montrons les résultats de classification obtenue en utilisant le classifieur SVM dans les deux résultats obtenus dans la première phase de réduction hybride. Finalement, les deux classifications obtenues ont été fusionnées et les résultats de notre approche ont montré leurs efficacité et performance par rapport aux méthodes existantes.

3.1 La RD spectro-spatiale : modèle de tenseur

Dans cette phase, deux techniques de RD basées sur le tenseur, à savoir TICA et TPCA sont appliquées pour calculer leur ensemble des caractéristiques intrinsèques, avec le nombre de dimension à conserver p est égal à trois ($p = 3$). En effet, nous avons appliqués deux techniques de RD basées sur le tenseur, à savoir TICA et TPCA. Dans la figure ?? les courbes de dispersions des différentes composantes obtenues Y_{TPCA} et Y_{TICA} sont comparés. D'après les résultats obtenus des nuages des points et les images réduites des trois premières composantes Y_{TPCA} et Y_{TICA} , les nouveaux points de données sont fortement corrélées par rapport à l'approche proposée TLPP qui clairement sépare les différents points de données de chaque

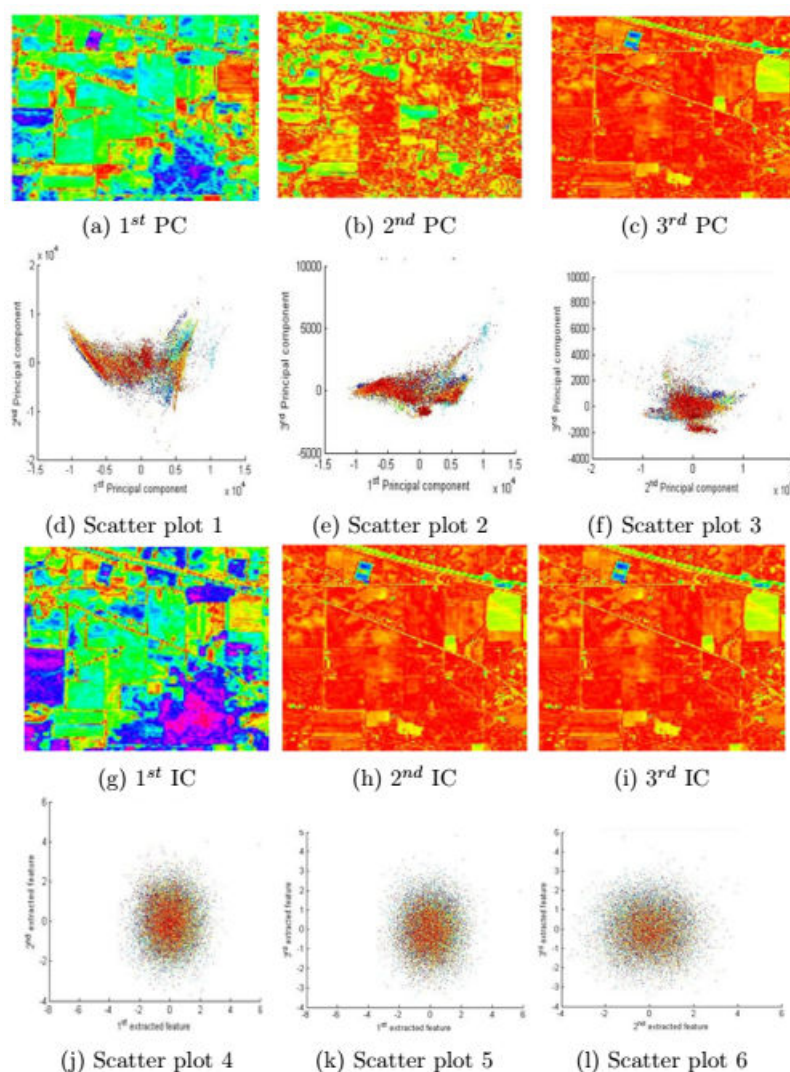


FIGURE 4.5 – Composantes extraites avec TPCA et TICA

composante obtenue. En d'autres termes, les deux techniques TICA et TPCA ne permettent pas de bien séparer les pixels de l'image réduite, ce qui provoque un taux de classification très faible. De plus, la technique de sélection des bandes CBS a été utilisée pour sélectionner les bandes spectrales pertinentes, c'est à dire les bandes discriminatives de l'IHS originale en minimisant la corrélation et la dépendance entre les différentes bandes spectrales. Le nombre de bandes à sélectionner a été fixé comme suit : $p = 3, 9$ et 15 .

3.2 Classification avec SVM

Pour la tâche de classification, nous avons utilisé une image de référence qui contient 16 classes prédéfinies. De ce fait, 20% des pixels (2078 pixels) a été aléatoi-

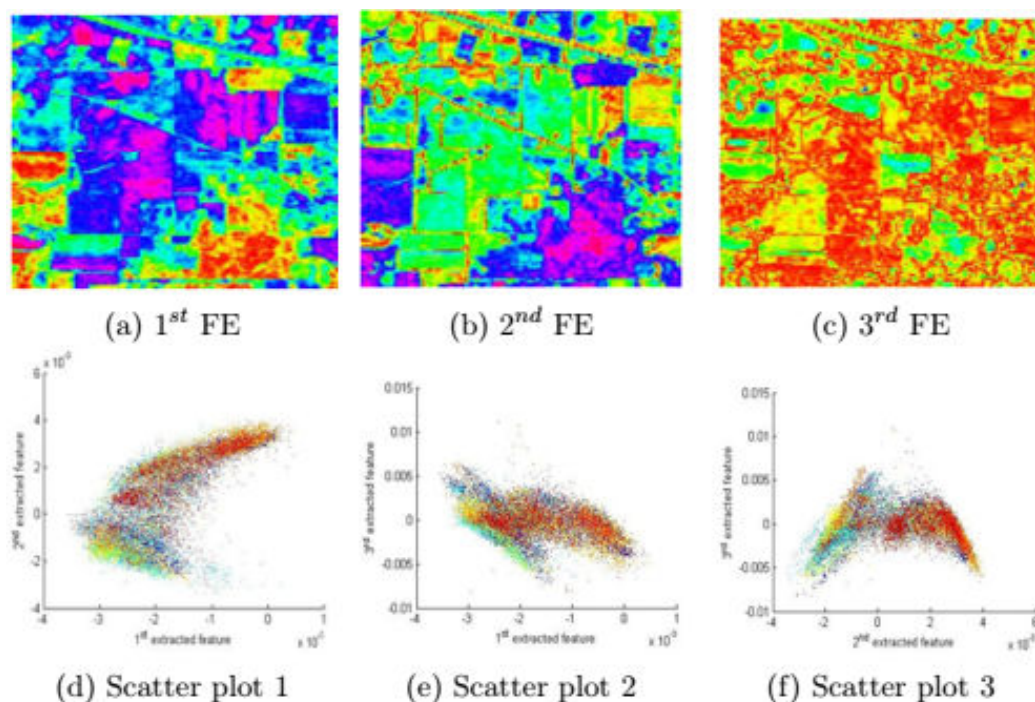


FIGURE 4.6 – Composantes extraites avec la technique de TLPP

Tableau 4.3 – bandes sélectionnées en utilisant CBS, AP, MI et BandClust

Nombre de bandes	CBS	AP	MI	BandClust
3	29, 167, 174	4, 45, 181	18, 32, 147	9, 16, 189
9	153, 167, 174, 177, 178, 179, 180, 181, 203	4, 98, 167, 177, 178, 181, 189, 200, 201	142, 165, 167, 177, 178, 179, 188, 193, 201	8, 15, 89, 92, 131, 152, 167, 169, 180
15	28, 29, 30, 32, 148, 153, 164, 167, 170, 175, 176, 179, 180, 181, 203	2, 9, 12, 96, 98, 167, 177, 178, 179, 181, 189, 195, 199, 200, 201	17, 29, 30, 36, 143, 154, 156, 162, 169, 170, 181, 189, 195, 197, 201	9, 15, 27, 28, 33, 39, 78, 82, 89, 120, 132, 175, 167, 179, 198

rement choisi pour l'apprentissage ainsi 80% (8288 pixels) pour le test. Les labels des classes ont été utilisé pour l'ensemble d'apprentissage et du test. Nous avons appliqué le classifier SVM sur les composantes extraites par la techniques TLPP ainsi les bandes sélectionnées par la technique CBS. les classes étiquetées aussi bien que les distributions des ensembles de test et d'apprentissages sont reportés dans le tableau 4.4. D'autres méthodes ont été appliquées aussi pour une évaluation et une comparaison de notre approche proposée, notamment TPCA et TICA et la technique proposée TLPP a montré son efficacité et sa performance par rapport aux différentes techniques de projection. Aussi, la technique CBS donne des meilleurs résultats par rapport aux autres méthodes de sélection des bandes (AP, BandClust et MI) (Figure 4.7). Figure 4.8 représente les différentes classification obtenues SVM/CBS et SVM/TLPP.

Tableau 4.4 – classes de vérité terrain et leurs distributions entre l'ensemble d'apprentissage et de test

Classes		Échantillons	
No	Nom	Apprentissage	Test
1	Alfalfa	11	43
2	Corn-notill	287	1147
3	Corn-min	167	667
4	Corn	47	187
5	Grass/Pasture	100	397
6	Grass/Trees	150	597
7	Grass/Pasture-mowed	6	20
8	Hay-windrowed	98	391
9	Oats	4	16
10	Soybeans-notill	194	774
11	Soybeans-min	494	1974
12	Soybeans-clean	123	491
13	Wheat	43	169
14	Woods	259	1035
15	Building-Grass-Trees-Drives	76	304
16	Stone-steel Towers	19	76
Total		2078	8288

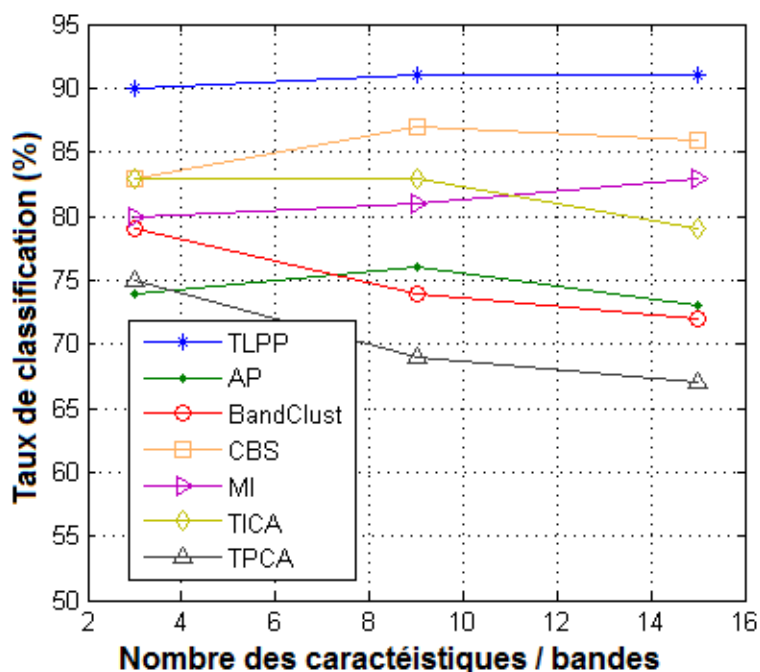


FIGURE 4.7 – Taux de classification avec TLPP, TPCA, TICA, CBS, AP, MI et BandClust

3.3 Fusion de décision avec DST

Les deux classifications obtenues en utilisant les deux techniques CBS et TLPP sont performantes et prometteuses. La fusion de ces deux résultats devient une amélioration de la tâche de classification tout en tenant compte le problème d'incer-

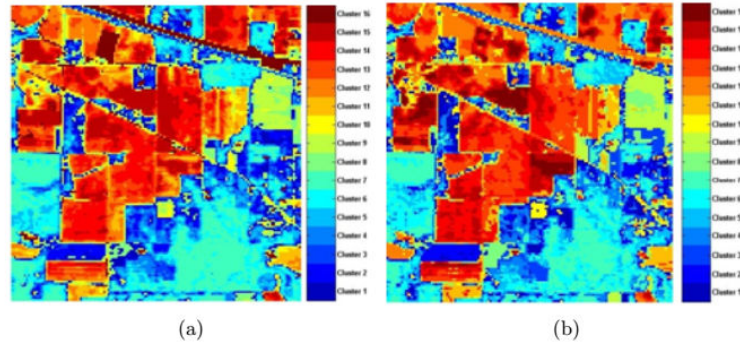


FIGURE 4.8 – Les images classifiées avec SVM : (a) : SVM/CBS, (b) : SVM/TLPP

titude. De ce fait, nous avons utilisé la technique DST (*Dempster Shafer Theory*) pour appliquer la fusion de classification dont le but d'améliorer la tâche de classification, vu que ces données peuvent être entachées par le bruit conditionné par les problèmes et les contraintes d'acquisition d'images hyperspectrales. De plus, le classifieur SVM n'est pas le seul classifieur adéquat et performant qui permet d'obtenir des bonnes classifications. Figure 4.9 montre l'image fusionnée et la figure

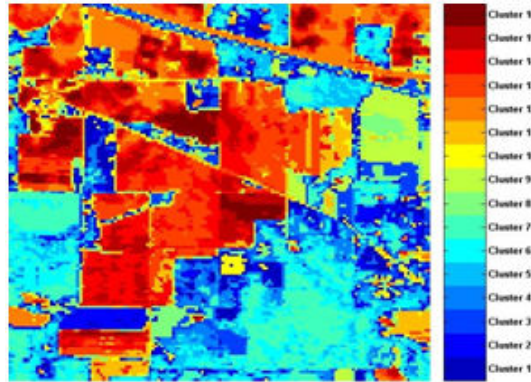


FIGURE 4.9 – Fusion d'images classifiées avec DST

4.10 illustre les performances de la classification fusionnée comparée aux deux techniques CBS et TLPP. D'après la classification obtenue, la fusion avec DST a fourni des résultats satisfaisants en terme de taux de classification par rapport aux deux techniques CBS et TLPP. Nous avons comparé la matrice de confusion permettant de comparer les résultats de classification fusionnée avec les informations fournies par la vérité terrain (Tableau 4.5). D'après l'analyse de cette matrice de confusion, nous pouvons constater que 96.68% des pixels sont correctement classifiés avec une bonne précision globale $OA = \frac{8013}{8288} = 96.68\%$ (*overall accuracy*) et le coefficient de Kappa $k = 96.22\%$ qui est qualifié comme étant parfait [Landis and Koch, 1977]. En outre, aucune des classes de test n'a été difficile à classifier car toutes les précisions de l'utilisateur UA (*user's accuracies*) sont $\geq 93\%$. De plus, le tableau 4.5 montre un bon compromis entre la précision utilisateur UA et la précision du producteur

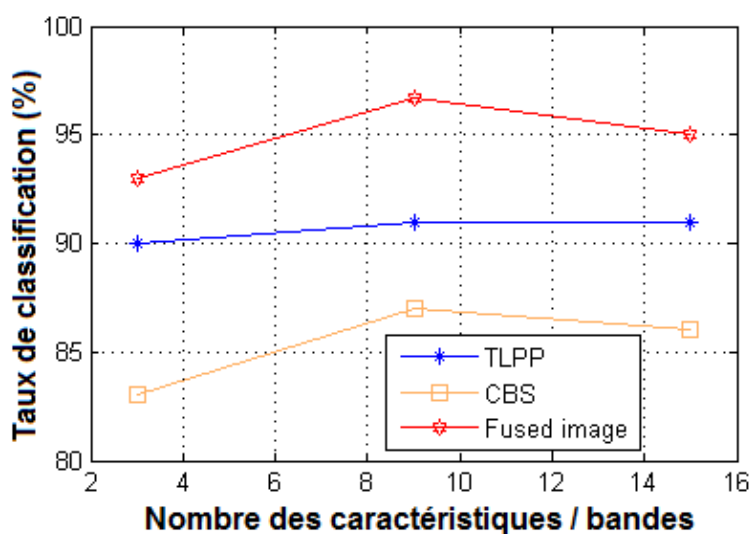


FIGURE 4.10 – Résultats de la classification (CBS, TLPP et image fusionnée)

PA (*producer's accuracies*).

Tableau 4.5 – Matrice de confusion de l'image classifiée avec SVM ($p = 9$)

No.	Vérité terrain																UA (%)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
2	0	1089	13	3	0	0	0	0	0	7	15	0	0	0	0	0	96.62
3	0	31	627	2	0	0	0	0	0	0	0	11	0	0	0	0	93.44
4	0	0	0	182	0	0	0	0	0	0	0	0	0	0	0	0	100
5	0	2	0	0	388	10	1	0	0	0	0	0	0	0	8	0	94.86
6	0	0	0	0	6	580	0	0	0	0	0	0	0	0	0	0	98.97
7	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	100
8	0	0	0	0	0	0	0	356	0	0	0	0	0	0	0	0	100
9	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	100
10	0	0	0	0	1	3	0	0	0	744	8	0	0	0	2	3	97.76
11	1	24	20	0	0	0	0	0	0	23	1950	8	0	0	0	0	96.24
12	0	1	7	0	0	4	0	0	1	0	1	472	0	0	0	0	97.11
13	0	0	0	0	0	0	0	0	0	0	0	0	163	0	0	0	100
14	0	0	0	0	0	0	0	0	0	0	0	0	0	1021	2	0	99.80
15	0	0	0	0	2	0	0	35	0	0	0	0	6	0	292	0	87.16
16	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	73	83.90
PA (%)	97.67	94.94	94	97.32	97.73	97.15	95	91.04	93.75	93.70	98.78	96.13	96.44	98.64	96.05	96.05	

3.4 Évaluation de temps de calcul

Dans ce section, nous présentons les temps de calcul des techniques utilisées dans les différentes expérimentations en utilisant un ordinateur personnel avec un CPU de

2.26GHz et une mémoire RAM de 4.0 GB. Figure 4.11 montre le temps d'exécution

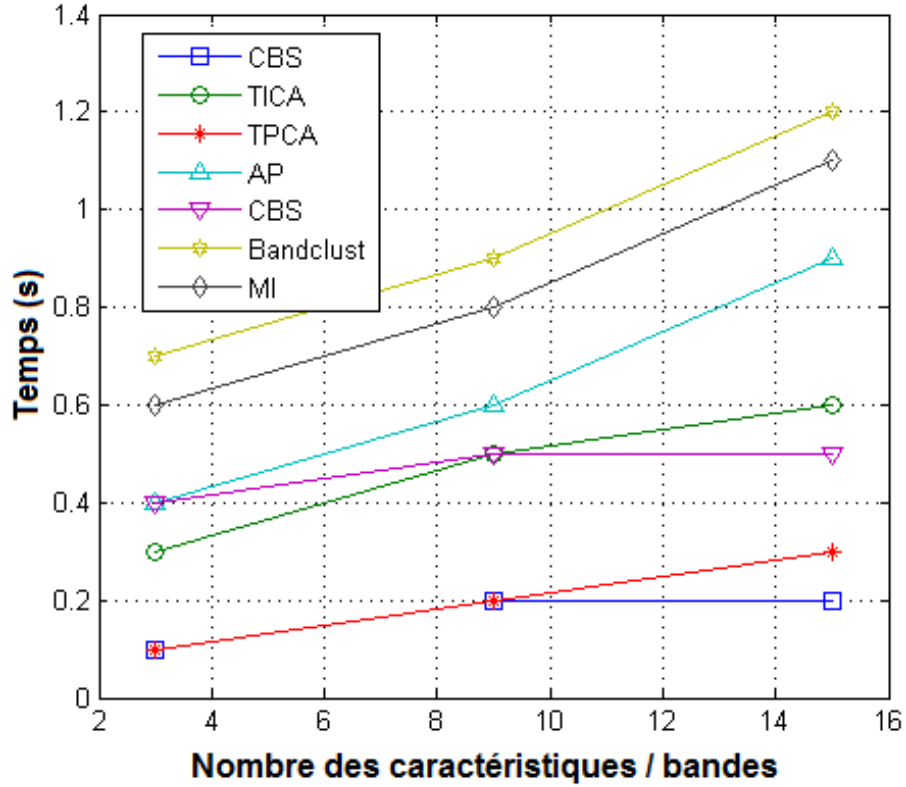


FIGURE 4.11 – Temps de calcul pour les différentes techniques de RD

de TPCA, TICA, MI, AP, BandClust, CBS et TLPP dans l'étape de RD. Nous pouvons constater que les deux techniques TLPP et TPCA sont plus rapides avec un temps d'exécution moyen (de 0.2 s). De plus, la technique de sélection CBS par rapport aux autres techniques de sélection de bandes. Concernant l'étape de classification, la figure 4.12 montre le temps d'exécution du classifieur SVM utilisé avec TPCA, TICA, MI, AP, BandClust, CBS et TLPP. Nous pouvons clairement constater que l'approche proposée TLPP/SVM aussi bien la méthode BandClust/SVM a un temps d'exécution faible (de 5 s).

3.5 Synthèse et discussions

Dans ce premier travail, nous avons essayé de proposer une approche de projection linéaire appelée extraction des caractéristiques basée sur le modèle de tenseur et la projection préservant la localité TLPP (features extraction based on Tensor modelling and Locality Preserving Projections). Le modèle de tenseur a été utilisé pour maintenir la distribution spatiale des pixels de l'IHS originale, qui est actuellement négligé dans la plupart des techniques de projection de bandes spectrales. Nous avons utilisé aussi la technique de CBS pour sélectionner un ensemble de

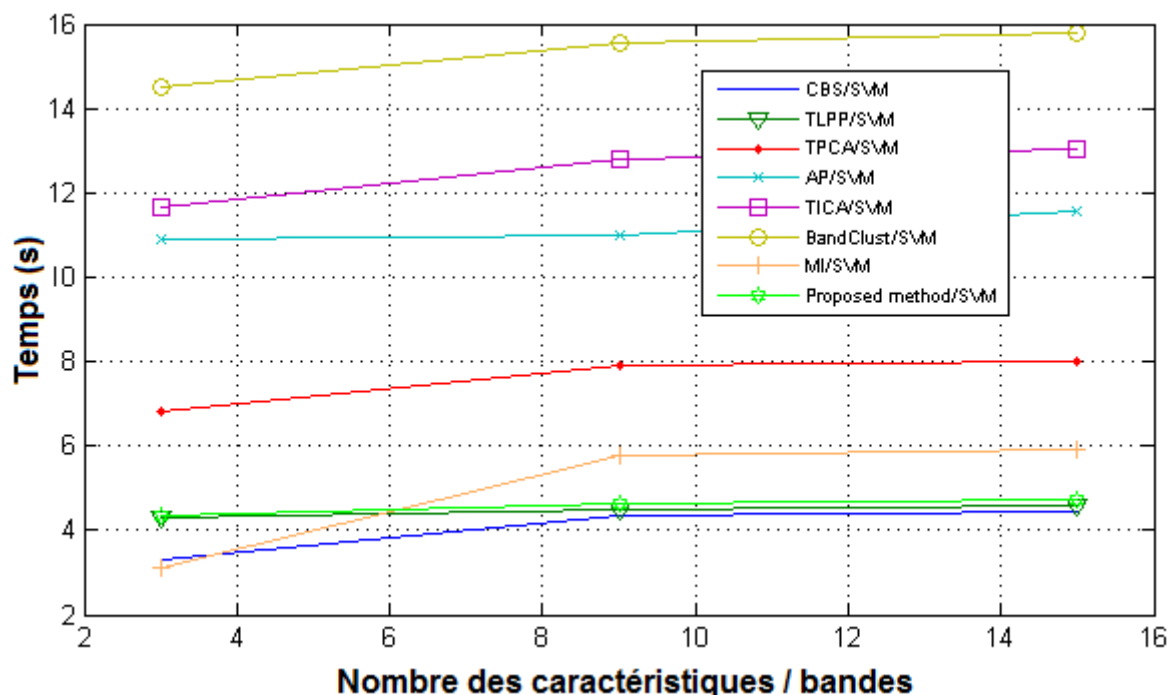


FIGURE 4.12 – Temps de calcul pour la classification des différentes images réduites avec SVM

bandes spectrales discriminantes et décorréliées à partir de données hyperspectrales originales. Pour exploiter la complémentarité d'informations fournies par les caractéristiques extraites avec TLPP et les bandes spectrales sélectionnées par CBS, nous avons implémenté le classifieur SVM en utilisant ces deux techniques de RD (TLPP et CBS). Les résultats de classifications obtenues *CBS/SVM* et *TLPP/SVM* ont été fusionnées par la théorie de DST pour considérer l'imperfection et l'incertitude des informations. Par conséquent, la classification fusionnée obtenue par DST a montré son efficacité et ses performances par rapport aux techniques de RD classiques, notamment TPCA, TICA et TLPP.

4 Approche de RD adaptative sémantique d'IHS

Dans cette section, nous démontrons l'efficacité de l'approche de sélection adaptative de bandes spectrales avec deux images, à savoir Indian Pines AIRIS et Pavia University ROSIS. Dans un premier lieu, chaque IHS a été projetée avec la technique TLPP. Dans ce cas, des nouvelles caractéristiques sont extraites contenant des informations utiles pour chaque image projetée. Le classifieur SVM a été utilisé pour évaluer la performance de différentes techniques de RD dans l'expérimentation. Trois aspects de performance, c'est-à-dire OA, AA et le coefficient de kappa sont adoptés pour évaluer les résultats de classification. En effet, la métrique OA est la proportion des pixels classifiés correctement par rapport à tous les pixels de l'image.

Le coefficient kappa mesure l'accord entre le résultat de classification avec la vérité terrain.

4.1 Techniques de comparaison

Plusieurs techniques de RD sont appliquées dans ce travail afin d'évaluer l'efficacité de l'approche proposée, c'est à dire la méthode ACBS (*Adaptive Constrained Band Selection*). En effet, des techniques de sélection de bandes spectrales basées la pondérations sont implémentées, à savoir l'ACI en utilisant la sélection de bandes BSICA (*band selection using independent component analysis*) [Du et al., 2003], la classifieur SVM avec l'entropie de similarité des caractéristiques pondérées SESVM (*SVM with similarity entropy feature weighting*) [Qi et al., 2014] et l'extraction des caractéristiques en utilisant des échantillons d'apprentissage pondéré FEWT (*feature extraction using weighted training samples*) [Imani and Ghassemian, 2015]. De plus, nous avons comparé l'approche proposée par d'autres techniques récentes de RD, notamment la sélection de bandes basée sur un modèle hypergraphe spectro-spatial SSHM (*band selection based on spectral-spatial hypergraph model*) [Bai et al., 2015], la RD spectro-spatiale SSDR (*spectro-spatial DR*) [Sellami and Farah, 2016], et nous avons utilisé aussi l'ensemble de données de l'IHS comme étant une référence qui est appelée *ORIGIN*.

4.2 Ajustement des paramètres

Toutes les expérimentations sont implémentées en utilisant l'environnement MATLAB sur 2.4-GHz CPU avec une mémoire RAM de 8-GB. Nous extrayons p caractéristiques et bandes spectrales, avec $p = 3, \dots, 21$ à partir de ces deux différents ensembles de données hyperspectrales, pour évaluer l'effet de la taille de dimensionnalité à retenir pour chaque méthode. Concernant la matrice de poids W de l'algorithme TLPP, nous considérons toute l'image pour calculer W . En fait, la variable $N = 12$, avec N représente le nombre de voisins entre les pixels x_i et y_j (sommets) dans le graphe G . Dans la phase d'extraction des règles, nous fixons le paramètre $k = 16$ clusters dans l'algorithme K-means Pour l'ensemble de données d'Indian Pines et $k = 9$ pour l'IHS de Pavia University, où k représente le nombre de classes pour chaque IHS. De plus, afin d'extraire l'ensemble des règles relatives à la contribution des caractéristiques pour la discrimination les classes, il est nécessaire d'entraîner le classifieur SVM en utilisant des échantillons étiquetés. Par conséquent, pour l'image d'Indiana Pines, nous choisissons aléatoirement 10% de pixels de l'IHS comme un ensemble d'échantillons d'apprentissage, et le reste des pixels sont considérés pour le test. Ainsi, pour l'image de Pavia University, nous choisissons 2% des pixels pour l'apprentissage et 98% de pixels pour le test.

4.3 Extraction des règles

La première phase de l'approche proposée vise à projeter les données hyperspectrales originales avec la technique TLPP. En se basant sur ces nouvelles caractéris-

tiques extraites par TLPP, nous tentons à partitionner les nouveaux points obtenus par l'algorithme de K-means afin de regrouper les points de données similaires des caractéristiques obtenues. De plus, nous utilisons ces groupe pour entraîner le classifier SVM afin de générer des règles permettant de chercher les bandes spectrales pertinentes pour la discrimination des classes. Par conséquent, nous montrons quelques règles obtenues pour Indian Pines et Pavia University avec $p = 3$ (Tableaux 4.6 et 4.7).

Tableau 4.6 – Règles obtenues pour Indiana Pines

R_1	: IF $0.4y_1 + 0.3y_2 + 0.16y_3 + 0.14$ THEN Alfafa
R_2	: IF $0.18y_1 + 0.39y_2 + 0.41y_3 + 0.02$ THEN Corn-notill
R_3	: IF $0.09y_1 + 0.7y_2 + 0.2y_3 + 0.01$ THEN Corn-mintill
R_4	: IF $0.21y_1 + 0.02y_2 + 0.57y_3 + 0.20$ THEN Corn
R_5	: IF $0.61y_1 + 0.08y_2 + 0.23y_3 + 0.08$ THEN Grass-pasture
R_6	: IF $0.03y_1 + 0.7y_2 + 0.2y_3 + 0.07$ THEN Grass-trees
R_7	: IF $0.3y_1 + 0.20y_2 + 0.2y_3 + 0.10$ THEN Grass-pasture-mowed
R_8	: IF $0.40y_1 + 0.09y_2 + 0.51y_3 + 0.0$ THEN Hay-windrowed
R_9	: IF $0.05y_1 + 0.3y_2 + 0.6y_3 + 0.04$ THEN Oats
R_{10}	: IF $0.41y_1 + 0.05y_2 + 0.37y_3 + 0.17$ THEN Soybeans-notill
R_{11}	: IF $0.51y_1 + 0.13y_2 + 0.28y_3 + 0.08$ THEN Soybeans-mintill
R_{12}	: IF $0.7y_1 + 0.09y_2 + 0.2y_3 + 0.01$ THEN Soybeans-clean
R_{13}	: IF $0.3y_1 + 0.4y_2 + 0.18y_3 + 0.12$ THEN Wheat
R_{14}	: IF $0.15y_1 + 0.37y_2 + 0.44y_3 + 0.04$ THEN Woods
R_{15}	: IF $0.08y_1 + 0.7y_2 + 0.3y_3 + 0.01$ THEN Bldg-Grass-Trees-Drive
R_{16}	: IF $0.24y_1 + 0.19y_2 + 0.54y_3 + 0.03$ THEN Stone-Steel-Towers

Tableau 4.7 – Règles obtenues pour Pavia University

R_1	: IF $0.5y_1 + 0.2y_2 + 0.10y_3 + 0.08$ THEN Asphalt
R_2	: IF $0.2y_1 + 0.37y_2 + 0.40y_3 + 0.03$ THEN Meadows
R_3	: IF $0.08y_1 + 0.8y_2 + 0.2y_3 + 0.01$ THEN Gravel
R_4	: IF $0.18y_1 + 0.05y_2 + 0.67y_3 + 0.10$ THEN Trees
R_5	: IF $0.31y_1 + 0.38y_2 + 0.20y_3 + 0.05$ THEN Metal sheets
R_6	: IF $0.7y_1 + 0.2y_2 + 0.04y_3 + 0.06$ THEN Bare soils
R_7	: IF $0.5y_1 + 0.2y_2 + 0.26y_3 + 0.04$ THEN Bitumen
R_8	: IF $0.38y_1 + 0.19y_2 + 0.43y_3 + 0.00$ THEN Bricks
R_9	: IF $0.04y_1 + 0.7y_2 + 0.2y_3 + 0.06$ THEN Shadows

D'après ces règles obtenues, différentes pondérations sont attribuées pour chaque composante afin de déterminer les composantes pertinentes pour chaque classe C_j . Par exemple, pour la classe '*Woods*' la composante qui permet de représenter et discriminer cette classe est la composante y_2 avec une pondération $w = 0.7$. Par conséquent, pour la composante y_2 , nous pouvons déterminer les différentes bandes

qui ont contribué pour la construction de la composante y_2 en utilisant le coefficient du poids absolu moyen.

4.4 Résultats de classification et analyse de performance de la sélection de bandes spectrales

Dans une première expérience, nous essayons de comparer les différentes techniques de projection en se basant sur le modèle tensoriel avec la technique TLPP. Par conséquent, nous comparons la technique TLPP avec les techniques TPCA (*Tensor-Principal Component Analysis*), TICA (*Tensor-Independent Component Analysis*) et TPP (*Tensor-Projection Pursuit*) [Renard and Bourennane, 2009] sur les deux données hyperspectrales d'Indian Pines et Pavia University. La figure 4.13 montre les taux de classification obtenus OA pour les deux images en utilisant les différentes techniques de projection par rapport au nombre des caractéristiques. En se basant sur les résultats obtenus, nous pouvons prouver que la technique TLPP est la plus meilleure que TPCA, TICA, et TPP avec OA est égal à 92.3% pour AVIRIS Indian Pines et 94.5% pour ROSIS Pavia University.

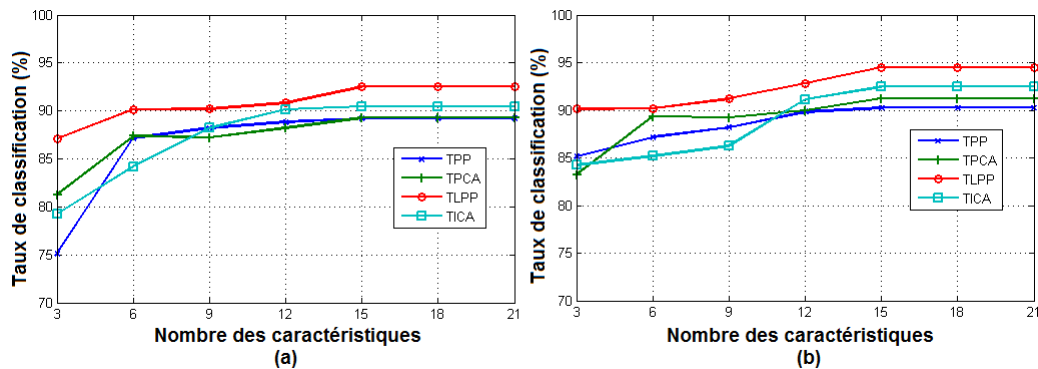


FIGURE 4.13 – Comparaison des taux de classifications OA par rapport au nombre des caractéristiques p en utilisant TPP, TPCA, TLPP et TICA : (a) Indian Pines (b) Pavia University.

De plus, nous appliquons d'autres méthodes de sélection de bandes spectrales en utilisant des règles extraites à partir de la technique TLPP afin d'évaluer la méthode proposée ACBS pour la classification d'IHS. De ce fait, nous utilisons la méthode de critère basé sur la divergence de Kullback-Leibler WaLuDi (*Kullback-Leibler divergence-based criterion*) (WaLuDi), le critère basé sur l'information mutuelle WaLuMI (*mutual information-based criterion*) [Martínez-Usó Martínez-Usó et al., 2007] et la propagation d'affinité PA (*propagation affinity*) [Qian et al., 2009]. La figure 4.14 montre les résultats obtenus pour la classification de deux images en utilisant les différentes méthodes de sélection de bandes spectrales basées sur les règles extraites à partir de la technique TLPP. Selon les résultats obtenus, nous affirmons que la méthode proposée ACBS peut fournir de meilleurs résultats dans la classification en comparant par d'autres méthodes de

sélection de bandes spectrales, où le taux de classification OA obtenu avec ACBS est égal à 95% pour Indian Pines et 97.5% pour Pavia University.

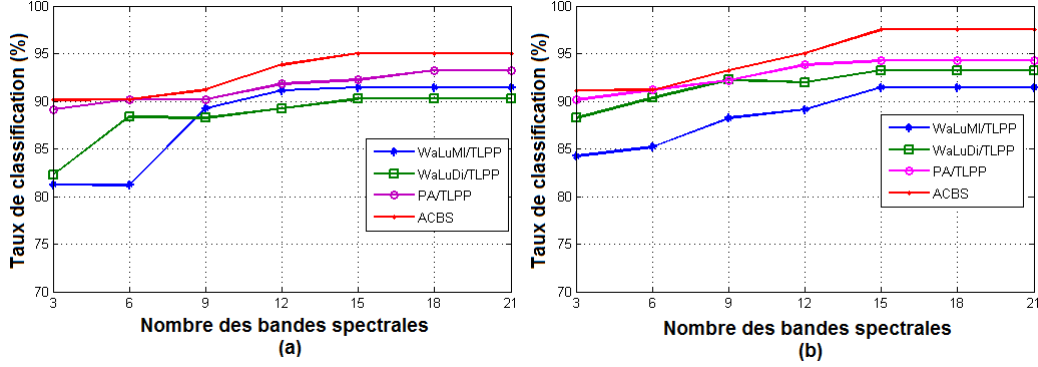


FIGURE 4.14 – Comparaison des taux de classifications OA par rapport au nombre des bandes spectrales p en utilisant WaLuDi/TLPP, WaLuMI/TLPP, PA/TLPP et ACBS : (a) AVIRIS Indian Pines (b) ROSIS Pavia University.

Pour les méthodes de RD spectro-spatiales, nous utilisons le modèle hypergraphe spectro-spatial SSHM [Bai et al., 2015], et la méthode de RD spectro-spatiale SSDR [Sellami and Farah, 2016] dans la phase d'extraction des règles pour substituer la méthode de projection TLPP. Par conséquent, la figure 4.15 montre les taux de classifications obtenus OA en utilisant les différentes méthodes de RD spectro-spatiale. En se basant sur les résultats obtenus, l'approche ACBS basée sur les règles générées à partir de TLPP reste prometteuse et donne des meilleurs résultats de classification pour l'image Indian Pines et Pavia University, par rapport aux autres méthodes de RD spectro-spatiale. Nous montrons dans la figure 4.16 la variation de OA par

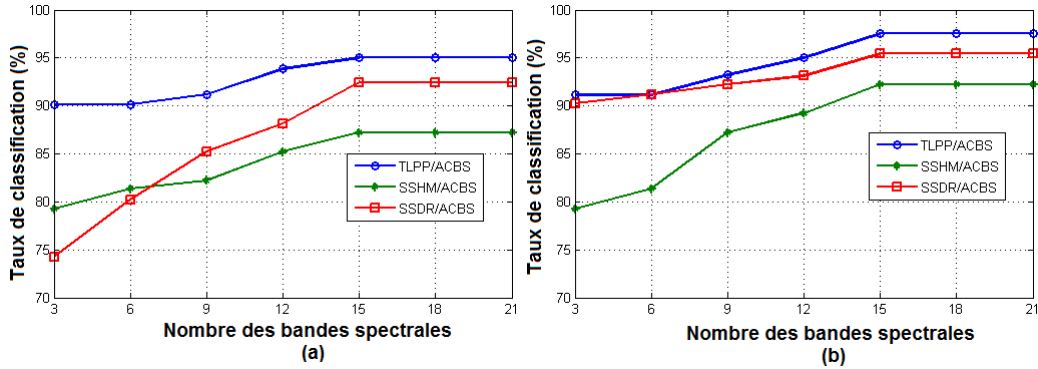


FIGURE 4.15 – Comparaison des taux de classifications OA par rapport au nombre des bandes spectrales p en utilisant TLPP/ACBS, SSHM/ACBS, et SSDR/ACBS : (a) AVIRIS Indian Pines (b) ROSIS Pavia University.

rapport au nombre de dimensions à conserver p et elle montre qu'en augmentant le nombre de composantes à extraire, OA reste stable et indique que l'approche proposée est meilleure par rapport aux autres techniques de RD. Ceci montre que les

bandes spectrales de l'IHS sont fortement corrélées et nous pouvons sélectionner un nombre minimal de bandes spectrales discriminatives et informatives par rapport aux différentes classes de l'IHS, ce qui permet de sélectionner les bandes contenant des informations utiles pour avoir un taux de classification satisfaisant.

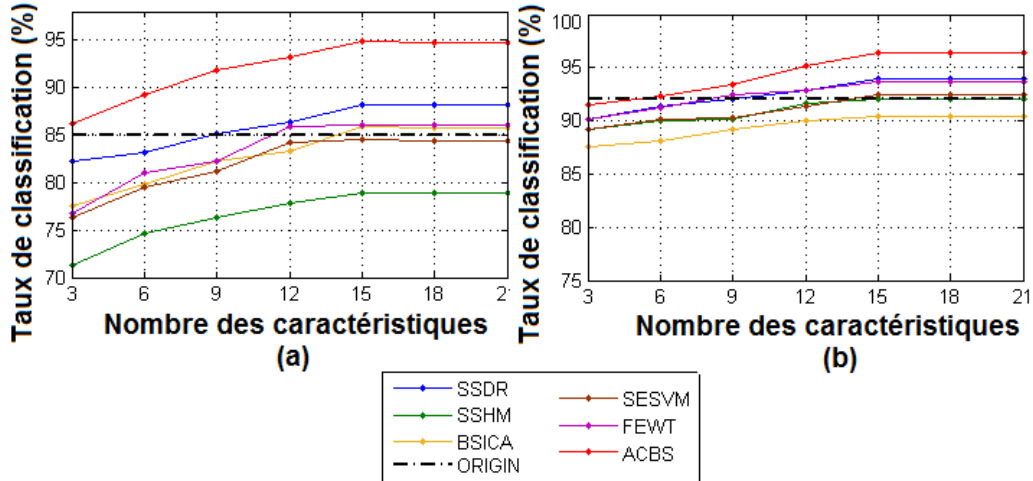


FIGURE 4.16 – Comparison des taux de classifications OA par rapport au nombre des caractéristiques : (a) the AVIRIS Indian Pines (b) ROSIS Pavia University.

Comme déjà montré dans la même figure 4.16, quand le nombre de composantes $p = 15$, OA est stable. Dans ce cadre, nous présentons les taux de classification pour chaque classe : OA, AA, le coefficient de kappa et le temps de calcul (en secondes) en fixant le nombre de bandes à conserver $p = 15$. Les deux tableaux 4.8 et 4.9 montrent les classifications obtenues pour chaque classe des images Indian Pines et Pavia University avec leurs taux de classifications en utilisant des différentes techniques de RD. De plus les deux figures 4.17 et 4.18 montrent les deux IHS classifiés obtenues avec les différentes techniques de RD. Par conséquent, les images réduites peuvent aboutir à des résultats de classification satisfaisants que les données originales de l'IHS qui sont fortement corrélées et les informations fournies par eux sont trop redondantes.

L'approche proposée ACBS sélectionne automatiquement et de façon adaptative p bandes spectrales pertinentes, autrement des bandes informatives et discriminatives entre eux avec des poids très élevés, ce qui améliore le taux de classification d'IHS.

D'après les résultats de classification obtenue, l'approche ACBS peut avoir un taux de classification très élevé de 8 parmi 16 classes de l'IHS d'Indian Pines et de 5 parmi 9 classes de l'IHS de Pavia University. Par conséquent, les taux de classification OA, AA et le coefficient k obtenus par l'approche proposée sont trop élevés et donnent des meilleurs résultats de classification par rapport aux techniques classiques de RD. Nous avons essayé d'évaluer le temps d'exécution pour chaque technique de RD. Par conséquent, les techniques SSSR et FEWT ont une performance de classification meilleure que les techniques BSICA et SSHM avec un temps du calcul plus élevé.

Tableau 4.8 – Taux de classification pour l’IHS de Indian Pines AVIRIS ($p = 15$)

Classe	Échantillons		Techniques						
	Train	Test	ORIGIN	SSDR	SSHM	BSICA	SESVM	FEWT	ACBS
Alfalfa (C1)	5	49	81.63	85.71	71.42	93.78	83.67	81.63	89.79
Build-G (C2)	38	342	87.71	89.18	83.33	87.71	96.78	80.40	93.56
Corn (C3)	23	211	85.30	93.36	86.25	90.04	80.56	90.52	91.46
Corn-M (C4)	83	751	86,55	93,34	79,89	86,55	86,68	86,55	97,20
Corn-N (C5)	143	1291	92,95	95,35	65,99	92,95	75,91	92,95	94,88
Grass-P-W (C6)	5	21	85,71	80,95	61,90	85,71	90,47	85,71	95,23
Grass-P (C7)	50	447	89,48	87,24	71,58	72,48	89,93	84,56	98,88
Grass-T (C8)	75	672	86,30	89,28	77,38	88,39	87,20	87,64	89,13
Hay-W (C9)	49	440	88,40	95,68	88,63	89,54	86,36	90,68	97,72
Oats (C10)	2	18	72,22	88,88	94,44	66,66	88,88	83,33	88,88
Soyb-C (C11)	62	552	81,52	82,24	79,71	82,06	82,60	81,52	97,82
Soyb-M (C12)	247	2221	85,45	85,36	80,36	85,09	85,45	84,87	94,32
Soyb-N (C13)	97	871	79,10	79,21	78,07	78,87	80,59	79,21	93,57
Stone-S (C14)	10	85	92,94	87,05	84,70	78,82	63,52	81,17	91,76
Wheat (C15)	21	191	69,10	90,57	90,05	88,48	91,09	93,19	89,52
Woods (C16)	130	1164	78,52	87,37	85,91	86,94	87,62	85,73	96,64
OA	–	–	85.05	88.18	78.95	85.84	84.49	86.04	94.82
AA	–	–	83.93	88.17	79.98	84.63	84.83	85.60	93.77
k	–	–	84.03	87.90	78.30	84.95	84.05	85.72	94.36
Temps (s)	–	–	215	8	25	36	17	27	19

L’approche proposée ACBS est assez rapide, mais le temps de calcul est satisfaisant et acceptable pour les applications d’analyse d’IHS. En se basant sur les résultats obtenus précédemment, nous pouvons conclure que les techniques de RD peuvent chercher la dimensionnalité intrinsèque des pixels de l’IHS et améliorer les performances de classification d’IHS. Bien que ces techniques de RD tentent à améliorer le taux de classification, la performance de classification obtenue par notre approche ACBS est meilleure par rapport aux techniques classiques de RD avec un taux de classification OA de 96.44% pour l’IHS de Pavia University et de 94.82% pour l’IHS d’Indian Pines.

Nous présentons la liste des bandes spectrales sélectionnées avec la sélection de bandes basée sur l’extraction des règles $SB_{\{ER\}}$, CBS : $SB_{\{CBS\}}$ et l’approche proposée $SB_{\{ACBS\}}$ technique. En effet, le tableau 4.10 présente une liste des bandes spectrales sélectionnées par l’approche proposée ACBS, où ces bandes sélectionnées peuvent être considérées comme étant discriminantes, distinctives, et informatives. Nous pouvons noter que les bandes informatives sont sélectionnées par le biais de l’information sémantique, c’est à dire les règles extraites à partir de la techniques TLPP notées par $SB_{\{ER\}}$, les bandes discriminantes et distinctives sont sélectionnées par la technique CBS notées par $SB_{\{CBS\}}$.

Tableau 4.9 – Taux de classification pour l'IHS de Pavia University ROSIS ($p = 15$)

Classe	Echantillons		Techniques						
	Train	Test	ORIGIN	SSDR	SSHM	BSICA	SESVM	FEWT	ACBS
Asphalt (C1)	132	6499	87,70	97,24	84,62	90,78	97,09	90,96	96,93
Bare-S (C2)	100	4929	87,23	91,49	93,32	91,70	91,09	90,99	97,38
Bitumen (C3)	26	1304	92,02	95,09	91,25	92,02	95,09	92,94	95,85
Gravel (C4)	41	2058	97,13	96,20	86,49	96,30	92,07	91,83	97,18
Meadows (C5)	373	18321	96,88	92,78	96,61	90,06	90,02	96,89	95,62
Painted-M (C6)	26	1319	83,39	94,08	90,14	90,52	98,86	90,82	97,80
Self-B (C7)	73	3609	90,05	92,24	83,12	84,37	94,48	91,99	97,25
Shadows (C8)	19	928	90,51	96,87	94,71	90,08	92,02	89,87	98,06
Tree (C9)	61	3003	83,25	97,23	93,80	87,07	93,63	90,07	96,57
OA	–	–	92,06	93,96	92,10	90,04	92,46	93,65	96,44
AA	–	–	89,80	94,80	90,45	90,32	93,82	91,82	96,96
k	–	–	91.14	93.94	91.95	90.12	93.24	92.89	96.60
Temps (s)	–	–	45	5	12	14	9	10	12

Tableau 4.10 – Les bandes sélectionnées avec ACBS. Classes d'Indian Pines : Alfalfa (C1) et Corn-N (C5). Classes de Pavia University : Asphalt (C1) et Meadows(C5)

HSI	$SB_{\{CBS\}}$	$SB_{\{ER\}}$		$SB_{\{ACBS\}}$	
Indian Pines	Toutes les classes	Alfalfa (C1)	Corn-N (C5)	Alfalfa (C1)	Corn-N (C5)
	04-12-21-43-65-76-81-82-94-101-107-117-184-201-212	02-09-20-35-43-55-62-75-81-94-97-101-103-167-184	05-12-20-21-31-45-52-76-78-81-87-92-118-175-198	43-94-101	21-76
Pavia University	Toutes les classes	Asphalt (C1)	Meadows (C5)	Asphalt(C1)	Meadows (C5)
	03-09-12-17-23-31-45-55-72-77-81-95-101-102-110	02-08-12-19-23-45-51-66-69-71-77-85-95-112-114	02-03-11-18-21-31-45-52-61-63-81-87-89-90-91	12-23-45-77	03-31-81

4.5 Impact de la sélection aléatoire des échantillons d'apprentissage sur la sélection de bandes spectrales

Dans cette section, nous analysons l'impact des variations des échantillons d'apprentissage sur la bandes spectrales sélectionnées. De ce fait, nous choisiss-

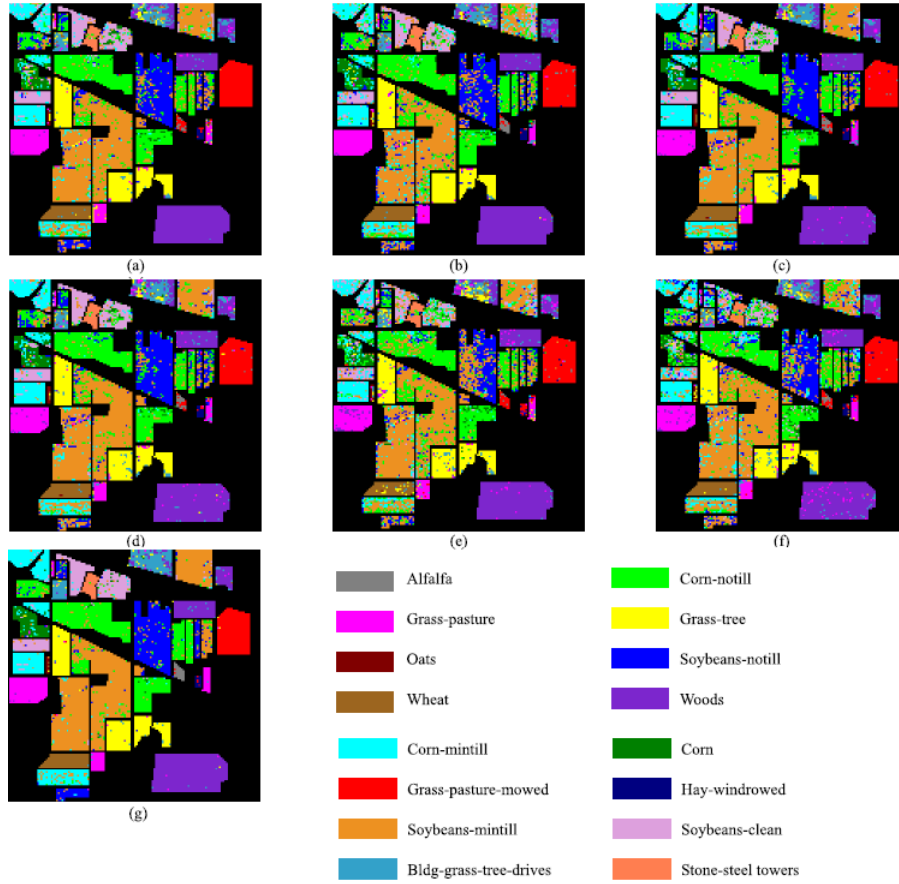


FIGURE 4.17 – Les résultats de classifications obtenues par SVM de l’Indiana Pines en utilisant les différentes approches de RD : (a) ORIGIN, (b) SSDR, (c) SSHM, (d) BSICA, (e) SESVM, (f) FEWT, (g) ACBS

sons aléatoirement 10% échantillons par classe (1040 samples) à partir de la vérité terrain (voir Tableau 4.8) pour l’IHS d’Indian Pines et 851 échantillons par classe (voir Tableau 4.9) pour l’IHS de Pavia University. En effet, pour ce faire, nous avons répété les expériences dix fois avec différents échantillons d’apprentissage pour chaque exécution. Pour tester la stabilité des bandes sélectionnées, nous avons utilisé trois mesures : 1) Indice Jaccard JI (*Jaccard Index*) [Real and Vargas, 1996], 2) Indice de stabilité de Kuncheva KSI (*Kuncheva’s Stability Index*) [Kuncheva, 2007], et 3) Indice de stabilité de l’information (ISI) (*Information Stability Index*) [Pascual et al., 2010, Gulgezen et al., 2009].

Le tableau 4.11 rapporte les mesures de stabilité des bandes pour quinze bandes sélectionnées ($p = 15$) calculées sur dix périodes d’exécution. En fait, nous comparons la stabilité des bandes sélectionnées avec l’approche proposée par rapport aux bandes sélectionnées par le critère basé sur la divergence de Kullback-Leibler (WaLuDi), le critère basé sur l’information mutuelle (Wa-

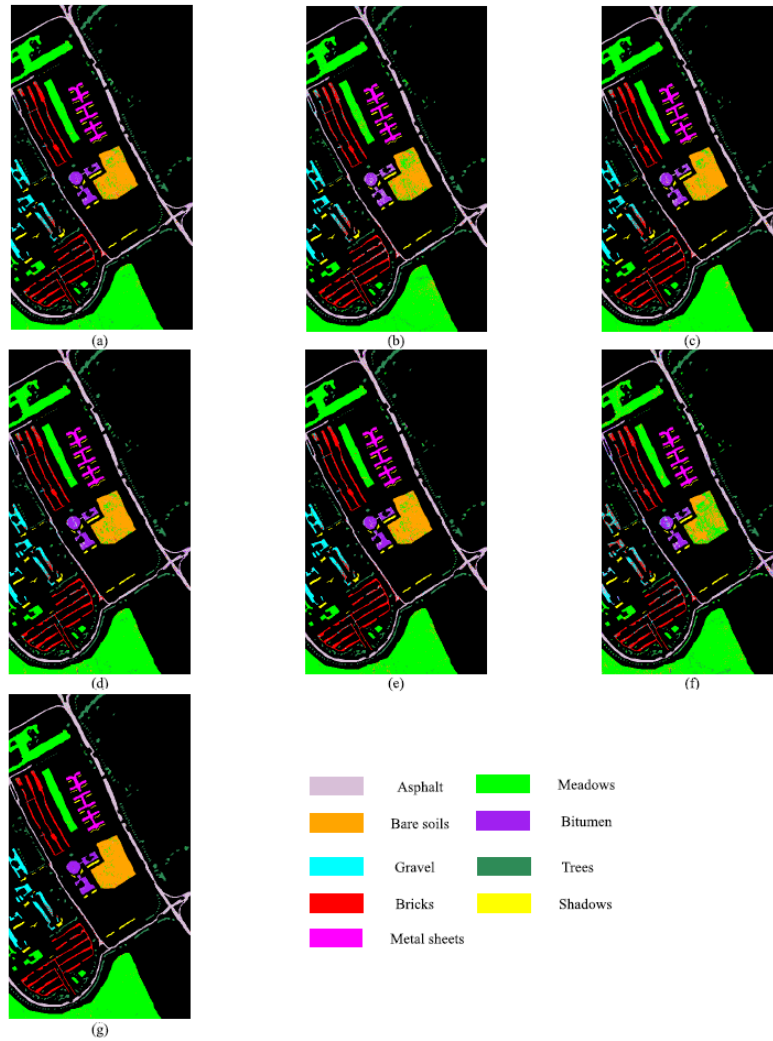


FIGURE 4.18 – Les résultats de classifications obtenues par SVM de Pavia University en utilisant les différentes approches de RD : (a) ORIGIN, (b) SSDR, (c) SSHM, (d) BSICA, (e) SESVM, (f) FEWT, (g) ACBS

LuMI) [Martínez-Usó Martínez-Usó et al., 2007], l'affinité de propagation (PA) [Qian et al., 2009] et la pondération des caractéristiques d'entropie de similarité (SESVM) [Qi et al., 2014]. Les valeurs les plus élevées dans le tableau 4.11 indiquent la haute stabilité dans les sous-ensembles de bandes sélectionnés. La méthode proposée offre une meilleure stabilité dans les bandes sélectionnées par rapport aux variations des pixels d'apprentissage dans les mesures de l'indice de bande (*band index measures*) et de la valeur de bandes (*band values measures*). En fait, la méthode proposée ACBS sélectionne les bandes proches qui sont valides pour les IHSs, dont les fréquences spectrales ont un comportement spectral similaire.

Tableau 4.11 – Mesures de stabilité de bandes sur dix sous-ensembles de bandes sélectionnées pour différentes méthodes de sélection de bandes ($p = 15$).

Méthodes	Indian Pines			Pavia University		
	JI	KSI	ISI	JI	KSI	ISI
ACBS	0.7526	0.8218	0.9344	0.7219	0.8547	0.9745
WaLuMI	0.4513	0.5362	0.7090	0.3914	0.5928	0.8724
WaLuDi	0.3457	0.4911	0.7984	0.5624	0.6412	0.8451
PA	0.4243	0.5124	0.7922	0.6891	0.7944	0.8972
SESVM	0.6214	0.7418	0.8452	0.7214	0.8845	0.8827

4.6 Interprétation des bandes spectrales sélectionnées

Les bandes sélectionnées avec l’approche proposée ACBS et les méthodes de sélection de bandes existantes considérées, c’est-à-dire WaluMI, WaLuDI, PA et SESVM sont rapportées dans la fig. 4.19. Pour les deux images, c’est à dire Indian Pines et Pavia University, 15 bandes spectrales ont été sélectionnées sont présentées dans la figure 4.19(a) et (b). Le grand nombre de bandes spectrales sélectionnées pour une région particulière du spectre est indiqué dans la plus grande largeur de la barre horizontale [Camps-Valls et al., 2010, Damodaran et al., 2017]. L’interprétation visuelle de la figure 4.19 indique l’efficacité et le potentiel de la méthode proposée ACBS pour la sélection de bandes spectrales pertinentes pour l’ensemble de la région du spectre afin d’ajuster divers objets naturels et artificiels dans les deux images Indian Pines et Pavia University. Cependant, les méthodes de sélection de bandes existantes sélectionnent un ensemble de bandes spectrales sur une région particulière du spectre. De plus, selon l’étude comparative, nous pouvons constater que les méthodes de sélection de bandes, qui couvrent uniformément toute la région du spectre, produisent de meilleurs résultats de classification.

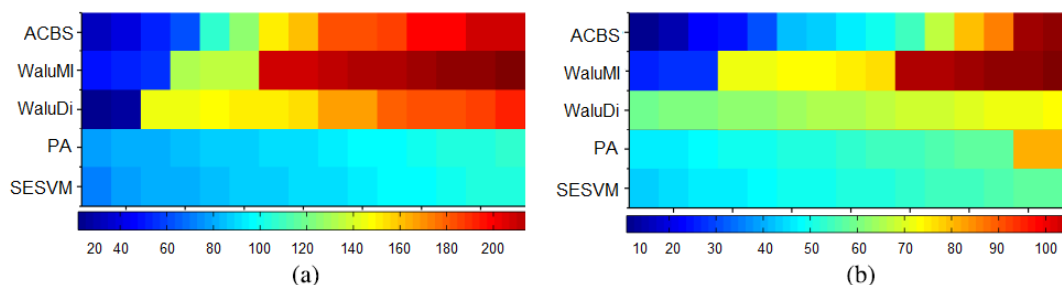


FIGURE 4.19 – Bande spectrales sélectionnées avec les différentes méthodes de sélection de bandes. (a) Indian Pines ($p = 15$), (b) Pavia University ($p = 15$). La barre de couleur indique l’indice d’une bande spectrale.

4.7 Etude de l'impact de la variation du paramètre K dans l'algorithme K-means

Dans cette section, nous avons analysé l'influence de la variation du paramètre K dans l'algorithme de clustering k-means. La figure 4.20 rapporte la classification obtenue afin de démontrer l'impact de la variation du paramètre K sur la précision de la sélection des bandes spectrales pertinentes. Pour l'IHS d'Indian Pines, nous fixons $K = 6, 9, 14, 16, 17$ et 18 clusters et $K = 4, 6, 8, 9, 10$ et 11 clusters pour l'image de Pavia University. Le nombre de bandes sélectionnées p est fixé à 15 bandes spectrales ($p = 15$). La même figure 4.20 montre la précision de classification obtenue pour les deux images. Selon les résultats obtenus, nous pouvons constater que lorsque le paramètre K dans le K-means diminue par rapport au nombre de classes de la vérité au sol, (16 classes pour Indian Pines et 9 classes pour Pavia University) le taux de classification diminue. On peut également constater que la meilleure précision de classification est obtenue avec le nombre de bandes spectrales égal à 15 pour les deux images.

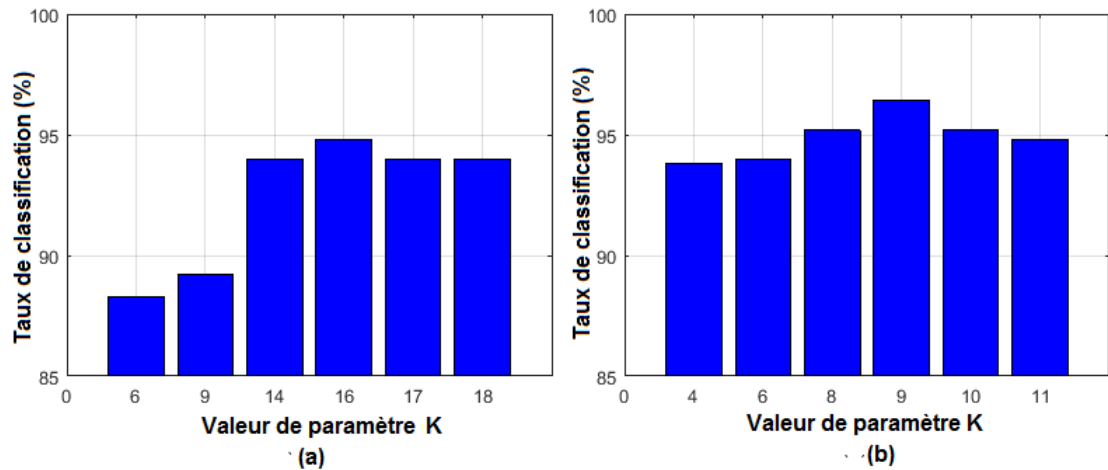


FIGURE 4.20 – L'impact de la variation du paramètre K dans l'algorithme k-means pour 15 bandes spectrales sélectionnées. (a) Indian Pines ($K = 6, 9, 14, 16, 17$ et 18). (b) Pavia University ($K = 4, 6, 8, 9, 10$ et 11)

4.8 Synthèse et discussions

Dans ce travail, nous avons introduit une approche de sélection adaptative de bandes spectrales basée sur la technique CBS et les techniques d'extraction des connaissances. La première phase consiste à projeter l'IHS originale en utilisant la technique TLPP. En outre, un nouvel algorithme d'extraction des règles à partir des images projetées a été développé en utilisant les techniques K-means et SVM. K-means vise à partitionner les données projetées pour avoir des clusters regroupant les points de données similaires. En se basant sur ces clusters, le classifieur SVM a été implémenté pour générer des règles visant de déterminer les composantes et

les bandes spectrales pertinentes permettant de discriminer les classes d'IHS. Dans ce cadre, des pondérations sont affectées aux différentes caractéristiques et bandes spectrales sous forme des règles. Par conséquent, ces règles ont été exploitées afin de sélectionner automatiquement et de manière adaptative les bandes spectrales moins redondantes, discriminantes, distinctives et informatives pour la représentation et la discrimination des classes de l'IHS. Par conséquent, les bandes sélectionnées peuvent améliorer le taux de classification et l'interprétation sémantique d'IHS de façon générale avec une corrélation inter-bandes très faible. Les résultats expérimentaux ont montré que l'approche proposée de sélection adaptative de bandes spectrales a des bonnes performances que d'autres techniques classiques de la RD d'IHS.

5 Conclusion

Afin de démontrer l'efficacité des approches proposées dans ce manuscrit de thèse, des expérimentations ont été abordées dans ce chapitre. Par conséquent, différents jeux de données ont été utilisés pour évaluer les deux contributions développées. La première approche a montré des résultats satisfaisants par rapport aux approches de RD existantes. Cependant, les données hyperspectrales réduites obtenues par cette approche ne sont pas informatives et elles ne préservent pas l'information contextuelle. De ce fait, nous avons essayé de valider la deuxième approche de sélection adaptative de bandes spectrales visant à sélectionner à la fois les bandes informatives et discriminantes en préservant l'information pertinente contextuelle. Les résultats obtenus ont démontré l'efficacité de l'approche proposée en la comparant avec différentes approches de la RD.

Conclusions générales et perspectives

La réduction de dimensionnalité en imagerie hyperspectrale est l'un des défis majeurs de la communauté scientifique de la télédétection. En effet, ces données massives issues par les capteurs hyperspectraux sont difficiles à exploiter et à analyser vu la forte corrélation inter-bandes spectrales, le nombre très élevé de bandes spectrales et aussi la redondance des informations. Dans le contexte de classification, ces bandes spectrales sont considérées comme étant trop riches au niveau spectral, ce qui entraîne une bonne discrimination des objets et des matériaux d'une scène hyperspectrale. Cependant, le nombre limité des échantillons d'apprentissages, i.e., les classes étiquetées et le grand nombre de bandes spectrales détériorent les performances de classifieurs, notamment le taux de classification n'est pas toujours optimal. Par conséquent, les approches de RD en IHS tentent à résoudre ces différents problèmes soit en transformant les données hyperspectrales dans un sous espace réduit tout en exploitant toutes les bandes spectrales originales ou en sélectionnant un sous-ensemble de bandes spectrales en utilisant un critère d'optimisation. Les approches de projection visent à chercher une meilleure projection de l'image originale, c'est à dire de trouver une représentation fidèle aux données hyperspectrales originales. Néanmoins, ces approches exploitent toutes les bandes spectrales de l'IHS pour aboutir à une projection contenant des nouvelles informations significatives, mais le changement de l'espace original provoque une perte et une transformation de l'information spectro-spatiale originale. Par conséquent, des informations utiles peuvent être perdues suite à cette transformation de données hyperspectrales originales. L'atout de ces approches que les informations significatives sont obtenues après une transformation de toutes les bandes originales de l'image. Cependant, les bandes spectrales tentent uniquement à sélectionner quelques bandes spectrales de l'image originale. Bien que les approches de sélection de bandes spectrales préservent le sens physique de données hyperspectrales originales, mais elles souffrent de plusieurs contraintes. En effet, les bandes sélectionnées dans le cas supervisé sont redondantes et exigent une grande base d'apprentissages étiquetée c'est à dire une base d'apprentissage représentative, mais comme mentionnée précédemment, le nombre d'échantillons dans le contexte hyperspectral est très limité. De même dans les approches de sélection non supervisée, les bandes sélectionnées sont uniquement discriminantes et distinctives mais elles ne sont pas informatives. Par conséquent, ces approches ne préservent pas l'information pertinente désirée selon un contexte donné. Ceux-ci exigent le développement des approches permettant de combler ces problèmes liés à la RD et à la technologie hyperspectrale. Dans ce contexte, émergent les différents problèmes étudiés et traités dans ce manuscrit de la thèse :

- La réduction de dimensionnalité en imagerie hyperspectrale en intégrant l'information spectro-spatiale ;
- La modélisation de l'incertitude et l'imperfection dans la classification et la réduction de dimensions ;
- L'extraction des connaissances sémantiques à partir des approches de projection
- la sélection adaptative de bandes spectrales pertinentes basée sur des connaissances sémantiques : sélection de bandes discriminantes, distinctives, peu redondantes et informatives selon le contexte.

1 Synthèse

Suite aux différents problèmes relevés dans ce travail, la première contribution a été portée sur l'intégration de l'information spectro-spatiale dans la RD afin d'améliorer les performances des classifieurs et d'aboutir à une classification d'IHS satisfaisante et optimale. En effet, nous avons essayé de proposer une nouvelle approche de réduction spectro-spatiale qui vise à préserver l'information spatiale dans la projection par LPP par le biais des tenseurs et de l'algèbre multi-linéaire. L'approche de LPP consiste généralement à préserver la distance entre les pixels de l'image originale dans l'espace de projection linéaire. Cependant, cette approche ne tient pas en compte de l'arrangement spatial des pixels vu à la vectorisation de l'image originale d'entrée. En d'autre terme, les informations de voisinage à savoir la structure spatiale de l'image n'est pas préservée. Par conséquent, nous avons essayé dans la première contribution de proposer une approche nommée TLPP (Tensor Locality Preserving Projection) permettant d'exploiter les tenseurs afin de prendre en considération conjointement l'information spectrale et spatiale. En effet, l'image originale a été considérée comme un tenseur de mode-3 (cube 3-D) au lieu d'une image vectorisée $2 - D$. Ceci préserve les dimensions spatiales et la dimension spectrale dans la projection avec TLPP et améliore le taux de précision de la classification d'IHS. De plus, l'approche de sélection de bandes spectrales CBS a été utilisée afin de sélectionner les bandes spectrales discriminantes et distinctives avec un minimum de corrélation et dépendance entre les bandes des données hyperspectrales originales. Dans une deuxième contribution, nous avons proposé de réduire l'incertitude et d'imprécision entachant la RD et les classifieurs. Pour ce faire, une approche de fusion évidentielle basée sur la théorie de Dempster-Shafer (DST) a été développée pour fusionner des images classifiées issues des approches de RD à savoir TLPP et CBS. Les résultats obtenus par cette approche de RD spectro-spatiale [Sellami and Farah, 2016] sont meilleurs et satisfaisants par rapport à d'autres approches de RD et ils montrent que l'approche proposée permet d'intégrer l'information spectro-spatiale dans la RD, préserve l'information pertinente tout en améliorant la précision de la classification d'IHS.

La troisième contribution de ce travail consiste à exploiter des connaissances sémantiques de bas-niveau issues des caractéristiques extraites par des approches de

projection. En fait, ces connaissances ont été exploitées dans le processus de sélection de bandes spectrales afin de fournir des connaissances sémantique de haut niveau permettant d'enrichir et d'alimenter la technique de sélection CBS. Pour cela, nous avons proposé d'étendre le modèle hybride de réduction spectro-spatiale, en exploitant les techniques d'extraction de connaissances afin de rendre l'approche proposée adaptative en terme de sélection de bandes pertinentes et ne nécessite qu'un minimum de connaissances *a priori* tout en profitant de ces connaissances sémantiques issues des approches de projection. Dans ce cadre, en se basant sur les caractéristiques obtenues par TLPP, les deux techniques k-means et SVM ont été utilisés pour extraire des connaissances qui peuvent être utilisées pour enrichir la technique de sélection de bandes spectrales CBS. L'algorithme k-means a été implémenté afin de regrouper les nouveaux points de données extraites par la technique TLPP. En effet, les différents clusters obtenus correspondent à des points de données similaires permettant de discriminer les différentes classes. En se basant sur ces clusters obtenus par k-means, le classifieur SVM a été utilisé pour générer des connaissances sémantique de bas-niveau sous forme des règles décrivant les contributions des caractéristiques et de bandes spectrales pour la discrimination des classes de l'image originale d'entrée. Par conséquent, il s'agit dans cette contribution de proposer une approche de RD permettant d'extraire de connaissances sémantique issue des approches de projection.

La dernière contribution de ce travail consiste alors à choisir l'ensemble de bandes pertinentes qui sont à la fois discriminantes, distinctives et informative avec un minimum de redondance en profitant des connaissances sémantiques qui décrivent les contributions de bandes spectrales pour la discrimination de classes de l'image originale. De ce fait, les connaissances extraites avec k-means et SVM représentent des pondérations relatives aux contributions des composantes extraites avec TLPP. Par l'exploitation de ces règles, les bandes spectrales pertinentes et informatives peuvent être détectées. Par conséquent, l'approche de sélection adaptative de bandes spectrales permet de sélectionner de façon automatique le bandes informatives, discriminantes de différentes classes de l'imagerie hyperspectrale d'entrée en profitant des connaissances sémantiques issues de TLPP et de la l'approche de sélection de bandes CBS. Enfin, les résultats obtenus en utilisant cette approche de RD adaptative montrent l'efficacité de l'exploitation des connaissances sémantiques pour la préservation de l'information pertinentes et l'amélioration de l'interprétation sémantique en IHS.

2 Perspectives

Plusieurs axes de recherches peuvent être envisagés de notre travail sur le plan méthodologique et applicatif.

Sur le plan méthodologique nous comptons d'étendre le modèle de réduction adaptative en utilisant des techniques récentes d'interprétation et de classification, notamment l'apprentissage profond (*Deep Learning*). En effet, ce modèle d'appren-

tissage profond permet supporter la redondance et l'hétérogénéité des informations issues de différentes bandes spectrales en fournissant des connaissances relatives à l'interprétation sémantique d'une scène hyperspectrale. En d'autres termes, ces algorithmes peuvent exploiter conjointement des bandes spectrales informatives et discriminantes en préservant l'information spectrale et spatiale pertinente en fournissant des interprétations relatives à la composition d'une scène hyperspectrale, les relations spatiales entre les classes et d'une carte thématique de l'IHS. De plus, nous projetons aussi dans les futurs travaux d'intégrer des approches de projection non linéaires. L'objectif principal est d'élaborer un modèle adaptatif fournissant des connaissances riches extraites à partir de deux familles des approches de projection, notamment linéaires et non linéaires, vu que les approches linéaires permettant de préserver la non linéarité de données hyperspectrales, notamment les informations de voisinage entre les pixels de l'image. Dans ce cadre, la base de connaissances sera enrichie par des nouvelles connaissances permettant bien évidemment le processus de sélection de bandes qui sera alimenté par ces nouvelles connaissances, ce qui améliore les résultats de sélection de bandes spectrales pertinentes. Dans un autre axe de recherche, nous visons à modéliser l'incertitude et l'imprécision entachant les données hyperspectrales ainsi les approches de RD en exploitant d'autres modèles à savoir les théories de possibilités ou les techniques de la logique floue qui permettent de fournir une interprétation sémantique de la scène hyperspectrale en réduisant les problèmes liés à l'incertitude et l'ambiguïté de l'information spectro-spatiale. Sur le plan applicatif, nous comptons à faire d'autres tests pour améliorer la performance de notre application. En effet, nous pensons à appliquer notre méthodologie sur d'autres types d'images hyperspectrales, notamment sur des images hyperspectrales médicales. D'autres domaines d'applications peuvent être envisagé dans les futurs travaux comme la détection de cibles.

Liste des Publications

Papier soumis

- **SELLAMI A**, FARAH IR, and SOLAIMAN B. Hyperspectral Imagery Semantic Interpretation based on Adaptive Constrained Band Selection and Knowledge Extraction Techniques, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. (En cours de révision)

Articles de revue avec comité de lecture

- **SELLAMI A**, FARAH IR, and SOLAIMAN B. An Adaptive Dimensionality Reduction Approach for Hyperspectral Imagery Semantic Interpretation. *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 2017, vol. 11, no. 2, p. 176-183.
- **SELLAMI A**, and FARAH IR. A Spatial Hypergraph Based Semi-Supervised Band Selection Method for Hyperspectral Imagery Semantic Interpretation. *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 2016, vol. 10, no. 10, p. 1798-1805.
- **SELLAMI A**, and FARAH IR. High-level hyperspectral image classification based on spectro-spatial dimensionality reduction. *Spatial Statistics*, 2016, vol. 16, p. 103-117.

Communications dans une conférence avec acte

- **SELLAMI A**, HAMDY R, and FARAH IR. An Adaptive Semantic Dimensionality Reduction Approach for Hyperspectral Imagery Classification, *14th ACS/IEEE International Conference on Computer Systems and Applications AICCSA 2017, October 30th to November 3rd, 2017* (Accepté).
- **SELLAMI A**, and FARAH IR. Interprétation sémantique d'images hyperspectrales basée sur la réduction adaptative de dimensionnalité. *PHARES 2017 : 7^{eme} atelier Phénomène Aléatoires pour la Représentation des signaux et d'images*, 22-24 Mai 2017. Hammamet, Tunisie.
- **SELLAMI A**, and FARAH IR. Sélection adaptative des bandes spectrales basée sur les techniques de projection pour la classification en imagerie hyperspectrale. *RFMI 2015 : 5^{eme} édition des ateliers Représentation des Formes et des Mouvements à partir des données Images*, 11-12 Novembre 2015, Hammamet.
- **SELLAMI A** and FARAH IR. Classification of hyperspectral images based on progressive bands selection and eigenmaps techniques. *TAIMA 2015 : Traitement et Analyse de l'Information : Méthodes et Applications*. 11-16 may 2015, Hammamet, Tunisia.

- **SELLAMI A**, FARAH IR, and SOLAIMAN B. Interpretation of hyperspectral imagery based on hybrid dimensionality reduction methods. *IPAS 2014 : International Image Processing and Systems Conference*, 05-07 november 2014, Hammamet, Tunisia, p. 1-7.

Bibliographie

- [Ahmad et al., 2012] Ahmad, M., Lee, S., Haq, I. U., and Mushtaq, Q. (2012). Hyperspectral remote sensing : Dimensional reduction and end member extraction. *International journal of soft computing and engineering*, 2(2) :170–175. (Cited on page 1.)
- [Ahn and Horenstein, 2013] Ahn, S. C. and Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3) :1203–1227. (Cited on page 51.)
- [Asl et al., 2014] Asl, M. G., Mobasheri, M. R., and Mojaradi, B. (2014). Unsupervised feature selection using geometrical measures in prototype space for hyperspectral imagery. *IEEE transactions on geoscience and remote sensing*, 52(7) :3774–3787. (Cited on page 49.)
- [Bai et al., 2015] Bai, X., Guo, Z., Wang, Y., Zhang, Z., and Zhou, J. (2015). Semi-supervised hyperspectral band selection via spectral–spatial hypergraph model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6) :2774–2783. (Cited on pages 74, 89 and 92.)
- [Bajcsy and Groves, 2004] Bajcsy, P. and Groves, P. (2004). Methodology for hyperspectral band selection. *Photogrammetric Engineering & Remote Sensing*, 70(7) :793–802. (Cited on page 36.)
- [Bajwa et al., 2004] Bajwa, S., Bajcsy, P., Groves, P., and Tian, L. (2004). Hyperspectral image data mining for band selection in agricultural applications. *Transactions-american Society of Agricultural Engineers*, 47(3) :895–908. (Cited on page 44.)
- [Bao et al., 2016] Bao, R., Xia, J., Dalla Mura, M., Du, P., Chanussot, J., and Ren, J. (2016). Combining morphological attribute profiles via an ensemble method for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 13(3) :359–363. (Cited on pages 29 and 52.)
- [Belkin and Niyogi, 2003] Belkin, M. and Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6) :1373–1396. (Cited on page 41.)
- [Benediktsson and Ghamisi, 2015] Benediktsson, J. A. and Ghamisi, P. (2015). *Spectral-Spatial Classification of Hyperspectral Remote Sensing Images*. Artech House. (Cited on pages 29, 31 and 32.)
- [Benediktsson et al., 2003] Benediktsson, J. A., Pesaresi, M., and Amason, K. (2003). Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9) :1940–1949. (Cited on page 29.)
- [Benediktsson et al., 1990] Benediktsson, J. A., Swain, P. H., and Ersoy, O. K. (1990). Neural network approaches versus statistical methods in classification of

- multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 28(4) :540–552. (Cited on page 26.)
- [Bjorgan and Randeberg, 2015] Bjorgan, A. and Randeberg, L. L. (2015). Real-time noise removal for line-scanning hyperspectral devices using a minimum noise fraction-based approach. *Sensors*, 15(2) :3362–3378. (Cited on pages 39 and 52.)
- [Borengasser et al., 2007] Borengasser, M., Hungate, W. S., and Watkins, R. (2007). *Hyperspectral remote sensing : principles and applications*. Crc Press. (Cited on page 17.)
- [Burgers et al., 2009] Burgers, K., Fessehatsion, Y., Rahmani, S., Seo, J., and Wittman, T. (2009). A comparative analysis of dimension reduction algorithms on hyperspectral data. *LAMDA Research Group*, pages 1–23. (Cited on page 2.)
- [Burges, 1998] Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2) :121–167. (Cited on page 66.)
- [Camastra, 2003] Camastra, F. (2003). Data dimensionality estimation methods : a survey. *Pattern recognition*, 36(12) :2945–2954. (Cited on page 49.)
- [Camps-Valls and Bruzzone, 2005] Camps-Valls, G. and Bruzzone, L. (2005). Kernel-based methods for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 43(6) :1351–1362. (Cited on page 26.)
- [Camps-Valls et al., 2010] Camps-Valls, G., Mooij, J., and Scholkopf, B. (2010). Remote sensing feature selection by kernel dependence measures. *IEEE Geoscience and Remote Sensing Letters*, 7(3) :587–591. (Cited on page 98.)
- [Canter, 1997] Canter, F. (1997). Evaluating the uncertainty of area estimates derived from fuuy land-cover classification. *Photogrammetric Engineering & Remote Sensing*, 63(4) :403–414. (Cited on page 32.)
- [Chang, 2003] Chang, C.-I. (2003). *Hyperspectral imaging : techniques for spectral detection and classification*, volume 1. Springer Science & Business Media. (Cited on pages 43 and 53.)
- [Chang and Du, 2004] Chang, C.-I. and Du, Q. (2004). Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *IEEE Transactions on geoscience and remote sensing*, 42(3) :608–619. (Cited on pages 49 and 51.)
- [Chang et al., 1999] Chang, C.-I., Du, Q., Sun, T.-L., and Althouse, M. L. (1999). A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 37(6) :2631–2641. (Cited on page 45.)
- [Chang and Liu, 2014] Chang, C.-I. and Liu, K.-H. (2014). Progressive band selection of spectral unmixing for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(4) :2002–2017. (Cited on pages 45 and 49.)
- [Chang and Wang, 2006] Chang, C.-I. and Wang, S. (2006). Constrained band selection for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6) :1575–1585. (Cited on pages 3, 6, 45, 53 and 63.)

- [Chaudhary et al., 2016] Chaudhary, M. K., Tripathi, P., Bhatt, A., Saxena, A., et al. (2016). A comparative performance analysis of feature extraction techniques for hyperspectral image classification. *International Journal of Software Engineering and Its Applications*, 10(12) :179–188. (Cited on page 1.)
- [Chen et al., 2013] Chen, J., Richard, C., and Honeine, P. (2013). Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model. *IEEE Transactions on Signal Processing*, 61(2) :480–492. (Cited on page 16.)
- [Chen et al., 2010] Chen, L., Huang, R., and Huang, W. (2010). Graph-based semi-supervised weighted band selection for classification of hyperspectral data. In *Audio Language and Image Processing (ICALIP), 2010 International Conference on*, pages 1123–1126. IEEE. (Cited on pages 47 and 53.)
- [Chen et al., 2016] Chen, Y., Jiang, H., Li, C., Jia, X., and Ghamisi, P. (2016). Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10) :6232–6251. (Cited on pages 16 and 26.)
- [Clark, 2016] Clark, M. L. (2016). Mapping land cover with hyperspectral and multispectral satellites using machine learning and spectral mixture analysis. In *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, pages 513–516. IEEE. (Cited on page 1.)
- [Cox and Cox, 2000] Cox, T. F. and Cox, M. A. (2000). *Multidimensional scaling*. CRC press. (Cited on pages 41 and 52.)
- [Damodaran et al., 2017] Damodaran, B. B., Courty, N., and Lefèvre, S. (2017). Sparse hilbert schmidt independence criterion and surrogate-kernel-based feature selection for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(4) :2385–2398. (Cited on page 98.)
- [Datta et al., 2014] Datta, A., Ghosh, S., and Ghosh, A. (2014). Band elimination of hyperspectral imagery using partitioned band image correlation and capacity discrimination. *International Journal of Remote Sensing*, 35(2) :554–577. (Cited on page 45.)
- [De Backer et al., 2005] De Backer, S., Kempeneers, P., Debruyn, W., and Scheunders, P. (2005). A band selection technique for spectral classification. *IEEE geoscience and remote sensing letters*, 2(3) :319–323. (Cited on pages 43 and 53.)
- [De Lathauwer, 1997] De Lathauwer, L. (1997). *Signal processing based on multilinear algebra*. Katholieke Universiteit Leuven. (Cited on page 59.)
- [De Silva and Tenenbaum, 2003] De Silva, V. and Tenenbaum, J. B. (2003). Global versus local methods in nonlinear dimensionality reduction. *Advances in neural information processing systems*, pages 721–728. (Cited on pages 41 and 52.)
- [Diederich, 2008] Diederich, J. (2008). *Rule extraction from support vector machines*, volume 80. Springer Science & Business Media. (Cited on page 72.)
- [Du et al., 2003] Du, H., Qi, H., Wang, X., Ramanath, R., and Snyder, W. E. (2003). Band selection using independent component analysis for hyperspectral image

- processing. In *Applied Imagery Pattern Recognition Workshop, 2003. Proceedings. 32nd*, pages 93–98. IEEE. (Cited on pages 74 and 89.)
- [Du and Yang, 2008] Du, Q. and Yang, H. (2008). Similarity-based unsupervised band selection for hyperspectral image analysis. *IEEE Geoscience and Remote Sensing Letters*, 5(4) :564–568. (Cited on pages 45 and 53.)
- [Du and Younan, 2008] Du, Q. and Younan, N. (2008). Dimensionality reduction and linear discriminant analysis for hyperspectral image classification. In *Knowledge-Based Intelligent Information and Engineering Systems*, pages 392–399. Springer. (Cited on pages 40 and 52.)
- [Ehlers et al., 2003] Ehlers, M., Gähler, M., and Janowsky, R. (2003). Automated analysis of ultra high resolution remote sensing data for biotope type mapping : new possibilities and challenges. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57(5) :315–326. (Cited on page 16.)
- [Elavarasi et al., 2011] Elavarasi, S. A., Akilandeswari, J., and Sathiyabhama, B. (2011). A survey on partition clustering algorithms. *International Journal of Enterprise Computing and Business Systems*, 1(1). (Cited on page 28.)
- [Fang et al., 2014] Fang, Y., Li, H., Ma, Y., Liang, K., Hu, Y., Zhang, S., and Wang, H. (2014). Dimensionality reduction of hyperspectral images based on robust spatial information using locally linear embedding. *IEEE Geoscience and Remote Sensing Letters*, 11(10) :1712–1716. (Cited on pages 3, 41 and 52.)
- [Farah et al., 2008a] Farah, I. R., Boulila, W., Ettabaa, K. S., and Ahmed, M. B. (2008a). Multiapproach system based on fusion of multispectral images for land-cover classification. *IEEE Transactions on Geoscience and Remote Sensing*, 46(12) :4153–4161. (Cited on page 67.)
- [Farah et al., 2008b] Farah, I. R., Boulila, W., Ettabaa, K. S., Solaiman, B., and Ahmed, M. B. (2008b). Interpretation of multisensor remote sensing images : Multiapproach fusion of uncertain information. *IEEE Transactions on Geoscience and Remote Sensing*, 46(12) :4142–4152. (Cited on page 67.)
- [Farrell and Mersereau, 2005] Farrell, M. D. and Mersereau, R. M. (2005). On the impact of pca dimension reduction for hyperspectral detection of difficult targets. *IEEE Geoscience and Remote Sensing Letters*, 2(2) :192–195. (Cited on page 39.)
- [Fauvel et al., 2006] Fauvel, M., Chanussot, J., and Benediktsson, J. A. (2006). Kernel principal component analysis for feature reduction in hyperspectral images analysis. In *Signal Processing Symposium, 2006. NORISIG 2006. Proceedings of the 7th Nordic*, pages 238–241. IEEE. (Cited on pages 41 and 52.)
- [Feng et al., 2015] Feng, J., Jiao, L., Liu, F., Sun, T., and Zhang, X. (2015). Mutual-information-based semi-supervised hyperspectral band selection with high discrimination, high information, and low redundancy. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5) :2956–2969. (Cited on pages 43, 47 and 53.)
- [Feng et al., 2014] Feng, J., Jiao, L., Zhang, X., and Sun, T. (2014). Hyperspectral band selection based on trivariate mutual information and clonal selection. *IEEE*

- Transactions on Geoscience and Remote Sensing*, 52(7) :4092–4105. (Cited on pages 47 and 53.)
- [Feng et al., 2011] Feng, J., Jiao, L., Zhang, X., and Yang, D. (2011). Bag-of-visual-words based on clonal selection algorithm for sar image classification. *IEEE Geoscience and Remote Sensing Letters*, 8(4) :691–695. (Cited on page 47.)
- [Fisher and Chang, 2010] Fisher, K. and Chang, C.-I. (2010). Progressive band selection for satellite hyperspectral data compression and transmission. *J. Appl. Remote Sens.*, 4(1) :041770–1. (Cited on page 49.)
- [Fontani et al., 2013] Fontani, M., Bianchi, T., De Rosa, A., Piva, A., and Barni, M. (2013). A framework for decision fusion in image forensics based on dempster-shafer theory of evidence. *IEEE Transactions on Information Forensics and Security*, 8(4) :593–607. (Cited on page 67.)
- [Frey and Dueck, 2007] Frey, B. J. and Dueck, D. (2007). Clustering by passing messages between data points. *science*, 315(5814) :972–976. (Cited on page 28.)
- [Fu et al., 2014] Fu, Y., Jia, X., Huang, W., and Wang, J. (2014). A comparative analysis of mutual information based feature selection for hyperspectral image classification. In *Signal and Information Processing (ChinaSIP), 2014 IEEE China Summit & International Conference on*, pages 148–152. IEEE. (Cited on pages 43 and 53.)
- [Fukunaga, 2013] Fukunaga, K. (2013). *Introduction to statistical pattern recognition*. Academic press. (Cited on page 28.)
- [Ghamisi et al., 2014] Ghamisi, P., Benediktsson, J. A., and Ulfarsson, M. O. (2014). Spectral-spatial classification of hyperspectral images based on hidden markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 52(5) :2565–2574. (Cited on page 28.)
- [Ghamisi et al., 2015] Ghamisi, P., Dalla Mura, M., and Benediktsson, J. A. (2015). A survey on spectral-spatial classification techniques based on attribute profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5) :2335–2353. (Cited on page 32.)
- [Gillis et al., 2015] Gillis, N., Kuang, D., and Park, H. (2015). Hierarchical clustering of hyperspectral images using rank-two nonnegative matrix factorization. *IEEE Transactions on Geoscience and Remote Sensing*, 53(4) :2066–2078. (Cited on page 27.)
- [Golipour et al., 2016] Golipour, M., Ghassemian, H., and Mirzapour, F. (2016). Integrating hierarchical segmentation maps with mrf prior for classification of hyperspectral images in a bayesian framework. *IEEE Transactions on Geoscience and Remote Sensing*, 54(2) :805–816. (Cited on page 29.)
- [Groenen and Jajuga, 2001] Groenen, P. J. and Jajuga, K. (2001). Fuzzy clustering with squared minkowski distances. *Fuzzy Sets and Systems*, 120(2) :227–237. (Cited on page 28.)

- [Gu et al., 2014] Gu, Y., Gao, G., Zuo, D., and You, D. (2014). Model selection and classification with multiple kernel learning for hyperspectral images via sparsity. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6) :2119–2130. (Cited on page 26.)
- [Gu et al., 2016] Gu, Y., Liu, T., Jia, X., Benediktsson, J. A., and Chanussot, J. (2016). Nonlinear multiple kernel learning with multiple-structure-element extended morphological profiles for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6) :3235–3247. (Cited on page 29.)
- [Guerra et al., 2015] Guerra, R., Santos, L., López, S., and Sarmiento, R. (2015). A new fast algorithm for linearly unmixing hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(12) :6752–6765. (Cited on page 16.)
- [Gulgezen et al., 2009] Gulgezen, G., Cataltepe, Z., and Yu, L. (2009). Stable and accurate feature selection. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 455–468. Springer. (Cited on page 96.)
- [Harsanyi and Chang, 1994] Harsanyi, J. C. and Chang, C.-I. (1994). Hyperspectral image classification and dimensionality reduction : An orthogonal subspace projection approach. *IEEE Transactions on geoscience and remote sensing*, 32(4) :779–785. (Cited on page 36.)
- [He et al., 2005] He, X., Cai, D., and Niyogi, P. (2005). Laplacian score for feature selection. In *NIPS*, volume 186, page 189. (Cited on pages 45 and 53.)
- [Hemissi et al., 2013] Hemissi, S., Farah, I. R., Ettabaa, K. S., and Solaiman, B. (2013). Multi-spectro-temporal analysis of hyperspectral imagery based on 3-d spectral modeling and multilinear algebra. *IEEE Transactions on Geoscience and Remote Sensing*, 51(1) :199–216. (Cited on page 64.)
- [Hossain et al., 2014] Hossain, M. A., Jia, X., and Pickering, M. (2014). Subspace detection using a mutual information measure for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 11(2) :424–428. (Cited on pages 43 and 53.)
- [Huang et al., 2013] Huang, R., Yang, L., and Lv, Z. (2013). Unlabeled sample reduction in semi-supervised graph-based band selection for hyperspectral image classification. In *Image and Graphics (ICIG), 2013 Seventh International Conference on*, pages 414–417. IEEE. (Cited on pages 47 and 53.)
- [Hughes, 1968] Hughes, G. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE transactions on information theory*, 14(1) :55–63. (Cited on page 2.)
- [Hung et al., 2011] Hung, C.-C., Kulkarni, S., and Kuo, B.-C. (2011). A new weighted fuzzy c-means clustering algorithm for remotely sensed image classification. *IEEE Journal of Selected Topics in Signal Processing*, 5(3) :543–553. (Cited on page 28.)

- [Ifarraguerri and Prairie, 2004] Ifarraguerri, A. and Prairie, M. W. (2004). Visual method for spectral band selection. *IEEE Geoscience and Remote Sensing Letters*, 1(2) :101–106. (Cited on page 43.)
- [Imani and Ghassemian, 2015] Imani, M. and Ghassemian, H. (2015). Feature extraction using weighted training samples. *IEEE Geoscience and Remote Sensing Letters*, 12(7) :1387–1391. (Cited on page 89.)
- [Jia et al., 2012] Jia, S., Ji, Z., Qian, Y., and Shen, L. (2012). Unsupervised band selection for hyperspectral imagery classification without manual band removal. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2) :531–543. (Cited on page 45.)
- [Jimenez and Landgrebe, 1998] Jimenez, L. O. and Landgrebe, D. A. (1998). Supervised classification in high-dimensional space : geometrical, statistical, and asymptotical properties of multivariate data. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 28(1) :39–54. (Cited on pages 31 and 41.)
- [Jolliffe, 2002] Jolliffe, I. (2002). *Principal component analysis*. Wiley Online Library. (Cited on pages 3 and 39.)
- [Kalakech et al., 2011] Kalakech, M., Biela, P., Macaire, L., and Hamad, D. (2011). Constraint scores for semi-supervised feature selection : A comparative study. *Pattern Recognition Letters*, 32(5) :656–665. (Cited on page 47.)
- [Kanungo et al., 2002] Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y. (2002). An efficient k-means clustering algorithm : Analysis and implementation. *IEEE transactions on pattern analysis and machine intelligence*, 24(7) :881–892. (Cited on pages 28 and 71.)
- [Keinosuke, 1990] Keinosuke, F. (1990). Introduction to statistical pattern recognition. *Academic Press Inc.* (Cited on page 31.)
- [Keshava and Mustard, 2002] Keshava, N. and Mustard, J. F. (2002). Spectral unmixing. *IEEE signal processing magazine*, 19(1) :44–57. (Cited on page 16.)
- [Khoder et al., 2012] Khoder, J., Younes, R., and Ouezdou, F. (2012). Stability of dimensionality reduction methods applied on artificial hyperspectral images. *Computer Vision and Graphics*, pages 465–474. (Cited on page 58.)
- [Kim and Finkel, 2003] Kim, D. H. and Finkel, L. H. (2003). Hyperspectral image processing using locally linear embedding. In *Neural Engineering, 2003. Conference Proceedings. First International IEEE EMBS Conference on*, pages 316–319. IEEE. (Cited on pages 41 and 52.)
- [Kim and Rattakorn, 2011] Kim, S. B. and Rattakorn, P. (2011). Unsupervised feature selection using weighted principal components. *Expert systems with applications*, 38(5) :5704–5710. (Cited on pages 45 and 53.)
- [Kröse et al., 1993] Kröse, B., Krose, B., van der Smagt, P., and Smagt, P. (1993). An introduction to neural networks. *The Netherlands : Univ. Amsterdam Press*. (Cited on page 26.)

- [Kuncheva, 2007] Kuncheva, L. I. (2007). A stability index for feature selection. In *Artificial intelligence and applications*, pages 421–427. (Cited on page 96.)
- [Kuo et al., 2014] Kuo, B.-C., Ho, H.-H., Li, C.-H., Hung, C.-C., and Taur, J.-S. (2014). A kernel-based feature selection method for svm with rbf kernel for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(1) :317–326. (Cited on page 26.)
- [Kuybeda et al., 2007] Kuybeda, O., Malah, D., and Barzohar, M. (2007). Rank estimation and redundancy reduction of high-dimensional noisy signals with preservation of rare vectors. *IEEE Transactions on Signal Processing*, 55(12) :5579–5592. (Cited on page 41.)
- [Landgrebe, 1999] Landgrebe, D. (1999). Information extraction principles and methods for multispectral and hyperspectral image data. *Information processing for remote sensing*, 82 :3–38. (Cited on page 20.)
- [Landgrebe, 2005] Landgrebe, D. A. (2005). *Signal theory methods in multispectral remote sensing*, volume 29. John Wiley & Sons. (Cited on page 31.)
- [Landis and Koch, 1977] Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, pages 159–174. (Cited on page 85.)
- [Larar et al., 2014] Larar, A. M., Suzuki, M., and Wang, J. (2014). Multispectral, hyperspectral, and ultraspectral remote sensing technology, techniques and applications v. In *Proc. of SPIE Vol*, volume 9263, pages 926301–1. (Cited on page 16.)
- [Le Moan et al., 2011] Le Moan, S., Mansouri, A., Voisin, Y., and Hardeberg, J. Y. (2011). A constrained band selection method based on information measures for spectral image color visualization. *IEEE Transactions on Geoscience and Remote Sensing*, 49(12) :5104–5115. (Cited on page 45.)
- [Lelong et al., 1998] Lelong, C. C., Pinet, P. C., and Poilvé, H. (1998). Hyperspectral imaging and stress mapping in agriculture : a case study on wheat in beauce (france). *Remote sensing of environment*, 66(2) :179–191. (Cited on page 1.)
- [Letexier et al., 2008] Letexier, D., Bourennane, S., and Blanc-Talon, J. (2008). Nonorthogonal tensor matricization for hyperspectral image filtering. *IEEE Geoscience and Remote Sensing Letters*, 5(1) :3–7. (Cited on page 59.)
- [Li et al., 2014] Li, C., Yin, J., and Zhao, J. (2014). Using improved ica method for hyperspectral data classification. *Arabian Journal for Science & Engineering (Springer Science & Business Media BV)*, 39(1). (Cited on pages 40 and 52.)
- [Li et al., 2012] Li, W., Prasad, S., Fowler, J. E., and Bruce, L. M. (2012). Locality-preserving dimensionality reduction and classification for hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4) :1185–1198. (Cited on pages 40 and 52.)
- [Lu and Weng, 2007] Lu, D. and Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International journal of Remote sensing*, 28(5) :823–870. (Cited on page 32.)

- [Ma et al., 2010] Ma, L., Crawford, M. M., and Tian, J. (2010). Local manifold learning-based k -nearest-neighbor for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 48(11) :4099–4109. (Cited on pages 16 and 26.)
- [Manolakis et al., 2003] Manolakis, D., Marden, D., and Shaw, G. A. (2003). Hyperspectral image processing for automatic target detection applications. *Lincoln laboratory journal*, 14(1) :79–116. (Cited on page 15.)
- [Martínez-Usó Martínez-Usó et al., 2007] Martínez-Usó Martínez-Usó, A., Pla, F., Sotoca, J. M., and García-Sevilla, P. (2007). Clustering-based hyperspectral band selection using information measures. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12) :4158–4171. (Cited on pages 45, 91 and 97.)
- [Mayer et al., 2000] Mayer, R., Priest, R., Steliman, C., Hazel, G., and Schaum, A. (2000). Detection of camouflaged targets in cluttered backgrounds using fusion of near simultaneous spectral and polarimetric imaging. Technical report, DTIC Document. (Cited on page 1.)
- [Mitra et al., 2002] Mitra, P., Murthy, C., and Pal, S. K. (2002). Unsupervised feature selection using feature similarity. *IEEE transactions on pattern analysis and machine intelligence*, 24(3) :301–312. (Cited on pages 45 and 53.)
- [Mojaradi et al., 2009] Mojaradi, B., Abrishami-Moghaddam, H., Zoej, M. J. V., and Duin, R. P. (2009). Dimensionality reduction of hyperspectral data via spectral feature extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7) :2091–2105. (Cited on page 49.)
- [Muti and Bourennane, 2007] Muti, D. and Bourennane, S. (2007). Survey on tensor signal algebraic filtering. *Signal Processing*, 87(2) :237–249. (Cited on page 59.)
- [Özdemir et al., 2015] Özdemir, O. B., Soydan, H., Çetin, Y. Y., and Düzgün, H. Ş. (2015). Signature based vegetation detection on hyperspectral images. In *Signal Processing and Communications Applications Conference (SIU), 2015 23th*, pages 2501–2504. IEEE. (Cited on page 1.)
- [Pal and Foody, 2010] Pal, M. and Foody, G. M. (2010). Feature selection for classification of hyperspectral data by svm. *IEEE Transactions on Geoscience and Remote Sensing*, 48(5) :2297–2307. (Cited on pages 43 and 44.)
- [Pan et al., 2016] Pan, B., Shi, Z., Zhang, N., and Xie, S. (2016). Hyperspectral image classification based on nonlinear spectral-spatial network. *IEEE Geoscience and Remote Sensing Letters*, 13(12) :1782–1786. (Cited on page 30.)
- [Pascual et al., 2010] Pascual, D., Pla, F., and Sánchez, J. S. (2010). Cluster validation using information stability measures. *Pattern Recognition Letters*, 31(6) :454–461. (Cited on page 96.)
- [Peng et al., 2005] Peng, H., Long, F., and Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8) :1226–1238. (Cited on pages 2, 43 and 47.)

- [Pinto et al., 2016] Pinto, T., Sousa, T. M., Praça, I., Vale, Z., and Morais, H. (2016). Support vector machines for decision support in electricity markets strategic bidding. *Neurocomputing*, 172 :438–445. (Cited on page 72.)
- [Plaza et al., 2009] Plaza, A., Benediktsson, J. A., Boardman, J. W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., et al. (2009). Recent advances in techniques for hyperspectral image processing. *Remote sensing of environment*, 113 :S110–S122. (Cited on pages 2 and 36.)
- [Plaza et al., 2005] Plaza, A., Martinez, P., Plaza, J., and Perez, R. (2005). Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations. *IEEE Transactions on Geoscience and remote sensing*, 43(3) :466–479. (Cited on page 29.)
- [Pons, 2014] Pons, J. M. F. (2014). *Robust target detection for Hyperspectral Imaging*. PhD thesis, Supélec. (Cited on pages xv and 19.)
- [Pudil et al., 1994] Pudil, P., Novovičová, J., and Kittler, J. (1994). Floating search methods in feature selection. *Pattern recognition letters*, 15(11) :1119–1125. (Cited on page 43.)
- [Qi et al., 2014] Qi, B., Zhao, C., and Yin, G. (2014). Feature weighting algorithms for classification of hyperspectral images using a support vector machine. *Applied optics*, 53(13) :2839–2846. (Cited on pages 89 and 97.)
- [Qian et al., 2009] Qian, Y., Yao, F., and Jia, S. (2009). Band selection for hyperspectral imagery using affinity propagation. *IET Computer Vision*, 3(4) :213–222. (Cited on pages 2, 31, 45, 53, 91 and 97.)
- [Quesada-Barriuso et al., 2014] Quesada-Barriuso, P., Arguello, F., and Heras, D. B. (2014). Spectral-spatial classification of hyperspectral images using wavelets and extended morphological profiles. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(4) :1177–1185. (Cited on page 29.)
- [Quinzán et al., 2009] Quinzán, I., Sotoca, J. M., and Pla, F. (2009). Clustering-based feature selection in semi-supervised problems. In *Intelligent Systems Design and Applications, 2009. ISDA'09. Ninth International Conference on*, pages 535–540. IEEE. (Cited on page 46.)
- [Real and Vargas, 1996] Real, R. and Vargas, J. M. (1996). The probabilistic basis of jaccard's index of similarity. *Systematic biology*, 45(3) :380–385. (Cited on page 96.)
- [Renard and Bourennane, 2008] Renard, N. and Bourennane, S. (2008). Improvement of target detection methods by multiway filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 46(8) :2407–2417. (Cited on page 59.)
- [Renard and Bourennane, 2009] Renard, N. and Bourennane, S. (2009). Dimensionality reduction based on tensor modeling for classification methods. *IEEE Transactions on Geoscience and Remote Sensing*, 47(4) :1123–1131. (Cited on pages 59 and 91.)

- [Richards and Richards, 1999] Richards, J. A. and Richards, J. (1999). *Remote sensing digital image analysis*, volume 3. Springer. (Cited on page 25.)
- [Riedmann and Milton, 2003] Riedmann, M. and Milton, E. (2003). Supervised band selection for optimal use of data from airborne hyperspectral sensors. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings. 2003 IEEE International*, volume 3, pages 1770–1772. IEEE. (Cited on page 44.)
- [Roweis and Saul, 2000] Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500) :2323–2326. (Cited on page 40.)
- [Sabins, 1999] Sabins, F. F. (1999). Remote sensing for mineral exploration. *Ore Geology Reviews*, 14(3) :157–183. (Cited on page 1.)
- [Saraloğlu et al., 2016] Saraloğlu, E., Görmüş, E. T., and Güngör, O. (2016). Mineral exploration with hyperspectral image fusion. In *Signal Processing and Communication Application Conference (SIU), 2016 24th*, pages 1281–1284. IEEE. (Cited on page 1.)
- [Schweizer and Moura, 2001] Schweizer, S. M. and Moura, J. M. (2001). Efficient detection in hyperspectral imagery. *IEEE Transactions on Image Processing*, 10(4) :584–597. (Cited on page 62.)
- [Sellami and Farah, 2016] Sellami, A. and Farah, I. R. (2016). High-level hyperspectral image classification based on spectro-spatial dimensionality reduction. *Spatial Statistics*, 16 :103–117. (Cited on pages 69, 89, 92 and 102.)
- [Serpico and Bruzzone, 2001] Serpico, S. B. and Bruzzone, L. (2001). A new search algorithm for feature selection in hyperspectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7) :1360–1367. (Cited on page 36.)
- [Shen et al., 2013] Shen, L., Zhu, Z., Jia, S., Zhu, J., and Sun, Y. (2013). Discriminative gabor feature selection for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 10(1) :29–33. (Cited on pages 43 and 47.)
- [Sidiropoulos et al., 2000] Sidiropoulos, N. D., Giannakis, G. B., and Bro, R. (2000). Blind parafac receivers for ds-cdma systems. *IEEE Transactions on Signal Processing*, 48(3) :810–823. (Cited on page 59.)
- [Smith et al., 2002] Smith, M.-L., Ollinger, S. V., Martin, M. E., Aber, J. D., Hallett, R. A., and Goodale, C. L. (2002). Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. *Ecological applications*, 12(5) :1286–1302. (Cited on page 1.)
- [Sun et al., 2014] Sun, K., Geng, X., Ji, L., and Lu, Y. (2014). A new band selection method for hyperspectral image based on data quality. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6) :2697–2703. (Cited on page 45.)
- [Sun et al., 2015a] Sun, S., Zhong, P., Xiao, H., and Wang, R. (2015a). An mrf model-based active learning framework for the spectral-spatial classification of

- hyperspectral imagery. *IEEE Journal of Selected Topics in Signal Processing*, 9(6) :1074–1088. (Cited on page 29.)
- [Sun et al., 2015b] Sun, W., Zhang, L., Du, B., Li, W., and Lai, Y. M. (2015b). Band selection using improved sparse subspace clustering for hyperspectral imagery classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6) :2784–2797. (Cited on page 49.)
- [Suykens et al., 2002] Suykens, J. A., Van Gestel, T., and De Brabanter, J. (2002). *Least squares support vector machines*. World Scientific. (Cited on page 26.)
- [Tan et al., 2014] Tan, K., Li, E., Du, Q., and Du, P. (2014). Hyperspectral image classification using band selection and morphological profiles. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(1) :40–48. (Cited on pages 45 and 53.)
- [Tarabalka et al., 2009] Tarabalka, Y., Benediktsson, J. A., and Chanussot, J. (2009). Spectral–spatial classification of hyperspectral imagery based on partitional clustering techniques. *IEEE Transactions on Geoscience and Remote Sensing*, 47(8) :2973–2987. (Cited on page 28.)
- [Tarabalka et al., 2010a] Tarabalka, Y., Benediktsson, J. A., Chanussot, J., and Tilton, J. C. (2010a). Multiple spectral–spatial classification approach for hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 48(11) :4122–4132. (Cited on page 29.)
- [Tarabalka et al., 2010b] Tarabalka, Y., Fauvel, M., Chanussot, J., and Benediktsson, J. A. (2010b). Svm-and mrf-based method for accurate classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 7(4) :736–740. (Cited on page 29.)
- [Tarabalka et al., 2012] Tarabalka, Y., Tilton, J. C., Benediktsson, J. A., and Chanussot, J. (2012). A marker-based approach for the automated selection of a single segmentation from a hierarchical set of image segmentations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(1) :262–272. (Cited on page 28.)
- [Teke and Sakarya, 2013] Teke, M. and Sakarya, U. (2013). Classification on hyperspectral images using enhanced fisher discriminant criterion. In *Signal Processing and Communications Applications Conference (SIU), 2013 21st*, pages 1–4. IEEE. (Cited on pages 3, 16 and 26.)
- [Tolle et al., 2003] Tolle, C. R., McJunkin, T. R., and Gorsich, D. J. (2003). Suboptimal minimum cluster volume cover-based method for measuring fractal dimension. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(1) :32–41. (Cited on page 49.)
- [Vasilescu and Terzopoulos, 2005] Vasilescu, M. A. O. and Terzopoulos, D. (2005). Multilinear independent components analysis. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 547–553. IEEE. (Cited on page 59.)

- [Wang and Chang, 2006] Wang, J. and Chang, C.-I. (2006). Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE transactions on geoscience and remote sensing*, 44(6) :1586–1600. (Cited on page 40.)
- [Wang et al., 2012] Wang, M., Gao, K., Wang, L.-j., and Miu, X.-h. (2012). A novel hyperspectral classification method based on c5.0 decision tree of multiple combined classifiers. In *Computational and Information Sciences (ICCIS), 2012 Fourth International Conference on*, pages 373–376. IEEE. (Cited on pages 16 and 26.)
- [Wang, 2015] Wang, X. (2015). Learning from big data with uncertainty—editorial. *Journal of Intelligent & Fuzzy Systems*, 28(5) :2329–2330. (Cited on page 32.)
- [Wang et al., 2017] Wang, Z., Du, B., Zhang, L., Zhang, L., and Jia, X. (2017). A novel semisupervised active-learning algorithm for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*. (Cited on page 30.)
- [Wang and He, 2011] Wang, Z. and He, B. (2011). Locality perserving projections algorithm for hyperspectral image dimensionality reduction. In *Geoinformatics, 2011 19th International Conference on*, pages 1–4. IEEE. (Cited on pages 6 and 64.)
- [Woodcock and Strahler, 1987] Woodcock, C. E. and Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote sensing of Environment*, 21(3) :311–332. (Cited on page 32.)
- [Yan et al., 2016] Yan, C., Bai, X., Ren, P., Bai, L., Tang, W., and Zhou, J. (2016). Band weighting via maximizing interclass distance for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 13(7) :922–925. (Cited on pages 43 and 53.)
- [Yu et al., 2016] Yu, H., Gao, L., Li, J., and Zhang, B. (2016). Spectral-spatial classification based on subspace support vector machine and markov random field. In *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, pages 2783–2786. IEEE. (Cited on page 29.)
- [Yuan et al., 2013] Yuan, H., Lu, Y., Yang, L., Luo, H., and Tang, Y. Y. (2013). Spectral-spatial linear discriminant analysis for hyperspectral image classification. In *Cybernetics (CYBCONF), 2013 IEEE International Conference on*, pages 144–149. IEEE. (Cited on page 40.)
- [Zhang et al., 2007] Zhang, L., Zhong, Y., Huang, B., Gong, J., and Li, P. (2007). Dimensionality reduction based on clonal selection for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12) :4172–4186. (Cited on page 43.)
- [Zhao et al., 2008] Zhao, J., Lu, K., and He, X. (2008). Locality sensitive semi-supervised feature selection. *Neurocomputing*, 71(10) :1842–1849. (Cited on page 46.)

- [Zhao and Liu, 2007] Zhao, Z. and Liu, H. (2007). Semi-supervised feature selection via spectral analysis. In *Proceedings of the 2007 SIAM International Conference on Data Mining*, pages 641–646. SIAM. (Cited on pages 46 and 47.)
- [Zhong et al., 2017] Zhong, P., Gong, Z., Li, S., and Schönlieb, C.-B. (2017). Learning to diversify deep belief networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*. (Cited on page 30.)
- [Zhou et al., 2017] Zhou, X., Li, S., Tang, F., Qin, K., Hu, S., and Liu, S. (2017). Deep learning with grouped features for spatial spectral classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 14(1) :97–101. (Cited on page 30.)
- [Zhu et al., 2015] Zhu, Z., Jia, S., He, S., Sun, Y., Ji, Z., and Shen, L. (2015). Three-dimensional gabor feature extraction for hyperspectral imagery classification using a memetic framework. *Information Sciences*, 298 :274–287. (Cited on page 32.)

Résumé

L'imagerie hyperspectrale permet d'acquérir des informations spectrales riches d'une scène dans plusieurs centaines, voire milliers de bandes spectrales étroites et contiguës. Cependant, avec le nombre élevé de bandes spectrales, la forte corrélation inter-bandes spectrales et la redondance de l'information spectro-spatiale, l'interprétation de ces données hyperspectrales massives est l'un des défis majeurs pour la communauté scientifique de la télédétection. Dans ce contexte, le grand défi posé est la réduction du nombre de bandes spectrales inutiles, c'est-à-dire de réduire la redondance et la forte corrélation de bandes spectrales tout en préservant l'information pertinente. Par conséquent, des approches de projection visent à transformer les données hyperspectrales dans un sous-espace réduit en combinant toutes les bandes spectrales originales. En outre, des approches de sélection de bandes tentent à chercher un sous-ensemble de bandes spectrales pertinentes. Dans cette thèse, nous nous intéressons d'abord à la classification d'imagerie hyperspectrale en essayant d'intégrer l'information spectro-spatiale dans la réduction de dimensions pour améliorer la performance de la classification et s'affranchir de la perte de l'information spatiale dans les approches de projection.

De ce fait, nous proposons un modèle hybride permettant de préserver l'information spectro-spatiale en exploitant les tenseurs dans l'approche de projection préservant la localité (TLPP) et d'utiliser l'approche de sélection non supervisée de bandes spectrales discriminantes à base de contraintes (CBS). Pour modéliser l'incertitude et l'imperfection entachant ces approches de réduction et les classifieurs, nous proposons une approche évidentielle basée sur la théorie de Dempster-Shafer (DST). Dans un second temps, nous essayons d'étendre le modèle hybride en exploitant des connaissances sémantiques extraites à travers les caractéristiques obtenues par l'approche proposée auparavant TLPP pour enrichir la sélection non supervisée CBS. En effet, l'approche proposée permet de sélectionner des bandes spectrales pertinentes qui sont à la fois informatives, discriminantes, distinctives et peu redondantes. En outre, cette approche sélectionne les bandes discriminantes et distinctives en utilisant la technique de CBS en injectant la sémantique extraite par les techniques d'extraction de connaissances afin de sélectionner d'une manière automatique et adaptative le sous-ensemble optimal de bandes spectrales pertinentes. La performance de notre approche est évaluée en utilisant plusieurs jeux de données hyperspectrales réelles.

Mots clef : Réduction de dimension, Apprentissage automatique, Analyse des données, Imagerie hyperspectrale, Algèbre multi-linéaire, Sélection de bandes, Extraction des caractéristiques, Interprétation sémantique

Abstract

Hyperspectral imagery allows to acquire a rich spectral information of a scene in several hundred or even thousands of narrow and contiguous spectral bands. However, with the high number of spectral bands, the strong inter-bands spectral correlation and the redundancy of spectro-spatial information, the interpretation of these massive hyperspectral data is one of the major challenges for the remote sensing scientific community. In this context, the major challenge is to reduce the number of unnecessary spectral bands, that is, to reduce the redundancy and high correlation of spectral bands while preserving the relevant information. Therefore, projection approaches aim to transform the hyperspectral data into a reduced subspace by combining all original spectral bands. In addition, band selection approaches attempt to find a subset of relevant spectral bands. In this thesis, firstly we focus on hyperspectral images classification attempting to integrate the spectro-spatial information into dimension reduction in order to improve the classification performance and to overcome the loss of spatial information in projection approaches.

Therefore, we propose a hybrid model to preserve the spectro-spatial information exploiting the tensor model in the locality preserving projection approach (TLPP) and to use the constraint band selection (CBS) as unsupervised approach to select the discriminant spectral bands. To model the uncertainty and imperfection of these reduction approaches and classifiers, we propose an evidential approach based on the Dempster-Shafer Theory (DST). In the second step, we try to extend the hybrid model by exploiting the semantic knowledge extracted through the features obtained by the previously proposed approach TLPP to enrich the CBS technique. Indeed, the proposed approach makes it possible to select a relevant spectral bands which are at the same time informative, discriminant, distinctive and not very redundant. In fact, this approach selects the discriminant and distinctive spectral bands using the CBS technique injecting the extracted rules obtained with knowledge extraction techniques to automatically and adaptively select the optimal subset of relevant spectral bands. The performance of our approach is evaluated using several real hyperspectral data.

Keywords: Dimensionality reduction, machine learning, Data analytics, Hyperspectral imagery, Multi-linear algebra, Band selection, Feature extraction, Semantic interpretation