
	ENSERG	TELECOM	
	Ecole Nationale Supérieure d'Electronique et de Radioélectricité de Grenoble	ENSIMAG	

PROJET numéro 87

ANNEE UNIVERSITAIRE **2007/2008**

PROJET DE TROISIEME ANNEE

RAPPORT FINAL

JUIN 2008

TITRE DU SUJET :

**Classification d'images hyperspectrales pour la caractérisation
du milieu urbain par une approche multirésolution**

Proposé par l'Entreprise :

Gipsa-lab - ENSIEG

Adresse :

**961 Rue de la Houille Blanche – B.P. 46
38402 – St. Martin D'Hères CEDEX**

Nom et Prénom du Responsable :

Mr FAUVEL Mathieu

Nom et Prénom du Tuteur INPG :

Mr CHANUSSOT Jocelyn

Noms, prénoms et option des Etudiants :

**Mr MERCADE MEDINA Sergi
Option MULTIMEDIA**

<i>Ecole Nationale Supérieure d'Electronique et de Radioélectricité de Grenoble</i> Grenoble INP – Minatec, 3 Parvis Louis Néel - B.P. 257 38016 GRENOBLE Cedex 1 – France Tél : 33 (0)4 56 52 91 00 - Fax : 33 (0)4 56 52 91 03	<i>Télécom - ENSIMAG</i> 681 rue de la Passerelle – Domaine Universitaire – BP 72 38402 ST MARTIN D'HERES Cedex – France Tél : 33 (0)4.76.82.72.22 – Fax : 33 (0)4.76.82.72.50
---	---

TABLE DES MATIÈRES

Abstract	3
1. Introduction.....	4
2. Le laboratoire Gipsa-lab	6
2.1. Département Images et Signal (DIS)	7
2.2. L'équipe Sigma-phi.....	8
3. Cahier des charges	13
4. Machines à Vecteurs Support	14
4.1. Le classifieur linéaire.....	14
4.2. Le classifieur non-linéaire	16
4.3. SVM Multi-classe.....	17
4.4. Sélection des paramètres	18
5. Transformée en ondelettes.....	19
5.1. Définitions mathématiques	19
5.2. Bases d'ondelettes orthogonales.....	20
5.3. Ondelettes et bancs de filtres.....	22
5.4. Bases d'ondelettes séparables.....	23
6. Analyse spatiale multirésolution	27
6.1. Analyse en Composantes Principales.....	27
6.2. Structure multirésolution	30
6.3. Effets de bords	31
7. Classification spectrale.....	33
7.1. Noyau spectral	33
7.2. Résultats expérimentaux	33
8. Classification spatiale	39
8.1. Noyau spatial	39

8.2. Résultats expérimentaux	42
9. Classification conjointe spectrale - spatiale	46
9.1. Noyau spatio-spectral	46
9.2. Résultats expérimentaux	49
10. Plan de travail (Diagramme de Gantt).....	52
11. Conclusion	56
12. Évaluation du coût du projet.....	58
Légende des figures.....	59
Glossaire	61
Références bibliographiques.....	62
Bibliographie complémentaire.....	64

ABSTRACT

The classification of optical urban remote-sensing images is addressed. Support Vector Machines (SVM) are proposed to classify hyperspectral images. An introduction to SVM is given in this report in order to help understand how they classify data according to the spectral information. Some kernel functions which are used to improve classification accuracy are presented as well. Then the use of spatial information through multiresolution decomposition is detailed.

The objective of this report is to propose a methodology including the spatial information in the classification process and trying to evaluate and improve the accuracy of this classification. Spatial information is extracted from a wavelet analysis of the image.

Finally experimental results are presented for each classification method: spectral, spatial and combining both spatial and spectral, and kernel parameters are selected in order to optimize the classification. After including the spatial information, classification accuracy has been improved.

CHAPITRE 1 INTRODUCTION

Dans l'imagerie satellitaire, les récents capteurs spatiaux permettent d'acquérir des images avec une grande résolution spatiale et spectrale. Des images de résolution spatiale de 0,75 mètres par pixel ont été obtenues avec les satellites IKONOS ou QUICKBIRD. Parallèlement des images hyperspectrales contenant des dizaines de bandes spectrales sont acquises par des capteurs aéroportés. Cette haute résolution spectrale et spatiale permet d'identifier sur des images, des zones précises comme des réseaux routiers, des zones de forêt, des zones industrielles, etc. dans des milieux urbains, ce qui motive à les classifier automatiquement.

Les images hyperspectrales ou multispectrales ont un intérêt par rapport aux images panchromatiques ou à trois composantes spectrales (RGB). Elles nous permettent d'avoir une description des caractéristiques physiques des objets de la scène grâce à la représentation sur plusieurs bandes spectrales (Figure 1). Grâce à l'amplitude de bandes qu'elles couvrent (du bleu jusqu'à proche de l'infrarouge : 400 – 800 nm) on peut faire ressortir un objet différent, mais possédant la même réponse spectrale dans le visible que ses voisins.

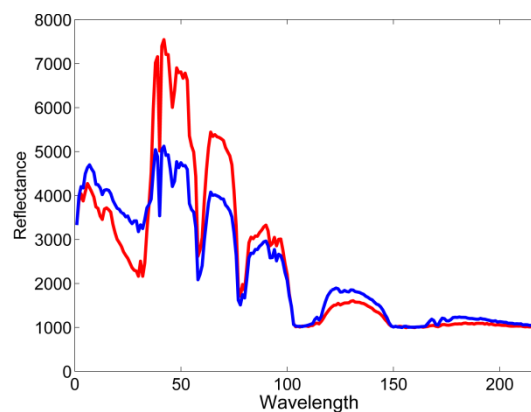


Figure 1 : Facteur de réflexion de deux pixels différents selon la longueur d'onde. Pour certaines longueurs d'onde des deux pixels a un niveau plus haut que l'autre.

Les capteurs hyperspectraux découpent une bande du spectre électromagnétique proche du visible en plusieurs tranches. Les capteurs font une analyse spectrométrique de cette bande et fournissent l'information dans plusieurs "couches" dans l'image. On peut voir une image hyperspectrale sous forme d'hypercube (Figure 2) où la taille de l'image est représentée par les deux premières dimensions et les bandes différentes sur la troisième dimension.

Pour l'analyse des images hyperspectrales, deux types d'approches sont généralement utilisées : les approches spectrales ou spatiales.

Dans la littérature, l'utilisation des caractéristiques spatiales a été peu abordée, alors que les caractéristiques spectrales ont été largement exploitées. Cela est dû à la mauvaise résolution spatiale des données hyperspectrales classiques. Cependant, les nouvelles images à haute résolution spatiale devraient permettre d'améliorer la classification. L'objectif de ce stage est

d'analyser les outils existants de classification et de développer des nouvelles méthodes prenant compte l'information spectrale ainsi que l'information spatiale.

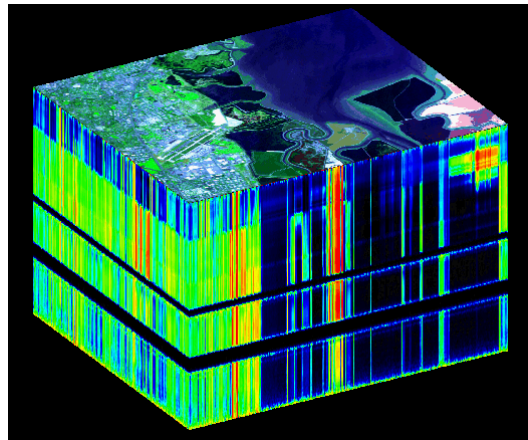


Figure 2 : Représentation d'une image hyperspectral sous forme d'hypercube. Les deux premières dimensions correspondent à la taille de l'image, alors que la troisième correspond aux bandes spectrales

Pour développer ces outils nous proposons d'analyser un modèle multirésolution d'image pour la définition de l'information spatiale.

La classification de ces données hyperspectrales se fait à l'aide des machines à vecteurs support (SVM). Les SVM sont des classificateurs à apprentissage statistique supervisé : il faut fournir un ensemble de données d'entraînement pour construire le classifieur. Face à d'autres méthodes de classification, les SVM ont un comportement plus performant pour un ensemble d'entraînement très réduit. Mais l'intérêt vient aussi de la possibilité d'agir sur la façon comme sont classifiées les données. Dans le quatrième chapitre de ce rapport, une description plus détaillée des SVM est présentée. Notamment, les SVM non-linéaire y seront étudiés.

L'analyse de l'information spatiale est étudiée dans le chapitre 6 à partir d'une analyse multirésolution de l'image. La représentation de l'image à différentes échelles, à l'aide d'une décomposition en ondelettes, va nous fournir des coefficients représentant l'information spatiale pour chaque échelle.

Dans les chapitres suivants, les résultats de classification sont présentés, pour la classification spectrale, spatiale et combinée.

Finalement, le déroulement du stage est présenté dans le chapitre 10 sous forme de diagramme de Gantt, ainsi que la comparaison avec la plan de travail prévu au pré-rapport.

CHAPITRE 2 LE LABORATOIRE GIPSA-LAB

Gipsa-lab s'investit dans des recherches fondamentales sur la parole, la perception, la cognition, le cerveau, le diagnostic et la commande des systèmes. Il développe des applications dans les secteurs de l'interaction multimodale, des télécommunications, de l'énergie, de l'environnement, des systèmes embarqués, de la mécatronique-robotique, de la santé, des transports, etc.

Des laboratoires et une équipe se sont regroupés pour former Gipsa-lab au 1er janvier 2007:

- Laboratoire d'automatique de Grenoble (LAG, UMR CNRS-INPG-Université Joseph Fourier)
- Institut de la communication parlée (ICP, UMR CNRS-INPG-Université Stendhal)
- Laboratoire des images et des signaux (LIS, UMR CNRS-INPG-UJF)
- Equipe de biomécanique (UJF)

Des recherches fondamentales et appliquées

Gipsa-lab développe des recherches au sein de trois départements:

- Automatique,
- Images et signal,
- Parole et cognition

Leurs domaines se situent au carrefour des théories générales en traitement de l'information et de la théorie du contrôle et de la commande, et des expérimentations et observations provenant du monde physique, biologique, cognitif et artefactuel.

L'interaction avec les mathématiques, la physique, les sciences de l'homme, l'automatique, la cognition et les communication

Le laboratoire s'appuie sur le socle des théories générales en traitement de l'information et contrôle/commande pour le développement d'algorithmes originaux, issus de fondements mathématiques et validés par des implémentations matérielles et logicielles efficaces.

Il se confronte à des mesures, des données, des observations, des modèles, des théories provenant du "monde réel, naturel ou artefactuel": physique, biologique et cognitif.

Il maintient un lien constant avec des applications de terrain dans des domaines multiples: interaction, communication, réseau, santé, environnement, systèmes embarqués, robotique, systèmes industriels, etc.

Parfaire la chaîne de l'observation par le développement de méthodes avancées

La démarche scientifique s'appuie sur le développement de méthodes avancées en automatique, traitement de la parole, signal et images. Ces méthodes intègrent des notions de

réseau, de contrôle-commande, de modélisation, d'analyse. Le souci est toujours présente de parfaire la chaîne de l'observation jusqu'à la conduite du processus, à l'interprétation ou à l'enrichissement de l'information multi-composante, multi-modale, multi-observation pour un diagnostic fiable et robuste.

Un ensemble de 300 personnes

A Gipsa-lab, la part de chercheurs est de 30% de l'ensemble chercheurs et enseignants-chercheurs. De même, la part des permanents CNRS est de 35% de l'ensemble des 137 permanents .

Le laboratoire accueille des chercheurs invités, post-doctorants et ATER, soit une vingtaine de personnes.

De nombreux futurs doctorants postulent à Gipsa-lab, qui encadre actuellement près de 140 thèses. Toutes les offres sont rémunérées et encadrées par des enseignants chercheurs et chercheurs, dont 44 titulaires d'une HDR.

Enfin, une soixantaine de stagiaires de master vient chaque printemps grossir les rangs du laboratoire.

Le personnel accompagnant la recherche est titulaire en totalité, soit 42 personnes réparties dans les services communs, les équipes de recherche ou les services des départements

2.1 . DÉPARTEMENT IMAGES ET SIGNAL (DIS)

Le département est structuré en équipes de recherche développant une forte collaboration et des projets transversaux:

- Signal images physique
- Géométrie, perception, images, gestes
- Communication, signal et sécurité
- Surveillance, signal et diagnostic
- Systèmes commandés en réseaux

Le traitement de l'information qu'elle soit Signal ou Images passe par des processus d'observation, d'acquisition, de traitement, d'analyse et d'interprétation, elle-même pouvant reboucler sur l'observation.

Ce sont ces éléments qui constituent les axes de recherche prioritaires au département Images et signal que ce soit dans un but de modéliser, de comprendre, de communiquer ou de sécuriser.

L'information elle-même passe par différents modes de représentation soit à l'aide de transformations (spatiales, fréquentielles, temps-fréquences, multi-échelles) soit de modes d'analyse (passage d'une image à l'analyse de son contenu en termes d'objets ou de formes)

soit de modes d'interprétation (analyse statistique, analyse de changements spatio-temporels).

Parmi les mots clés articulatoires on trouve

- Signal-Images-Physique
- Géométrie-Perception-Images-Gestes
- Adéquation-Algorithmes-Architecture
- Communication-Sécurité de contenu
- Signal-Diagnostic-Surveillance.

Face à la multiplicité des observations (multi-capteurs, multi-modalités, multi-composantes, multi voire hyperspectrales), il est important de maîtriser la pertinence et la redondance de chaque source avec le développement de méthodes de séparation de sources ainsi que de procédés de fusion.

Ce sont sur des aspects multi-modalités (audio et vidéo) ainsi que sur des aspects surveillance et diagnostic que le département Images et Signal élabore des coopérations avec les 2 autres départements du laboratoire en faisant émerger des projets ou des équipes communes.

2.2. L'ÉQUIPE SIGMA-PHY

L'équipe SIGMA-PHY (Signal image physique) est une fusion de l'ancienne équipe Signaux et Images Naturels (SIN) et d'une grande partie de l'équipe Non Linéaire (NL) du Laboratoire des Images et des signaux. Ces deux équipes ont toujours montré des intérêts communs autour de thèmes tels que la séparation de sources, l'analyse en composantes indépendantes, la déconvolution par exemple. Des projets de recherche communs, des masters, des thèses en co-direction existent entre ces deux groupes depuis de nombreuses années. Cette synergie s'est traduite par un ensemble de publications et par la création de l'équipe SIGMAPHY animée conjointement par P.O. Amblard et J.I. Mars.

Axes de recherche

Les activités du groupe SIGMAPHY se déclinent autour de 5 thèmes pour développer et acquérir de la connaissance par des travaux tant théoriques qu'appliqués autour de la théorie du signal et de l'image en liaison avec une connaissance de la physique des milieux étudiés.

1. Signal et Physique Non Linéaire
2. Séparation de Sources
3. Multidimensionnalité
4. Signal et Propagation des ondes
5. Géosciences et Télédétection

Cette équipe est reconnue par l'Observatoire des Sciences de l'Univers de Grenoble (OSUG).

Axe 1 : Signal et Physique Non Linéaire

L'idée forte au cœur des recherches menées ici est la fécondité de l'interaction entre sciences physiques et sciences de l'information. Les travaux se déclinent ainsi en trois sous thèmes:

1.1. *Fluctuations et systèmes (bio)physiques non linéaires*

Les travaux consistent en la compréhension des interactions fluctuations-systèmes non linéaires physiques et biophysiques, avec comme perspective le développement de nouveaux moyens de traitement de l'information en prenant en compte les fluctuations ambiantes. Exemple :

- Comportement du transfert de l'information dans des systèmes non linéaires.
- Utilisation des quantifieurs aléatoires pour des tâches de détection.
- Développement et compréhension de réseaux neuronaux réels, en architecture contrôlée et couplage parfait avec les capteurs.

1.2. *Information et physique*

Les travaux sont de nature théorique et vont se concentrer sur l'étude d'inégalités entropiques apparaissant en mécanique quantique. Des études plus pratiques concernent l'analyse d'un signal dans le plan Fisher-Entropie, en termes de classification et de caractérisation avec des applications au traitement de signaux biomédicaux.

1.3. *Invariances d'échelle et signaux fractals*

Les invariances d'échelle étudiées ici concernent deux types différents de lois de puissance : les lois temporelles et les lois de probabilité. Ces travaux concernent le traitement des signaux aléatoires impulsifs dont la modélisation probabiliste repose sur les lois alpha-stable. Dans le domaine temporel, les invariances d'échelle concernent l'invariance des propriétés statistiques par transformation de l'axe des temps (dilatation dans sa version la plus simple). Les travaux touchent l'utilisation de systèmes de fonctions itérés (IFS) aléatoires pour la modélisation des signaux fractals.

Axe 2 : Séparation de sources

Dans le contexte de signaux multidimensionnels, les travaux de ce thème de recherche se déclinent en deux sous thèmes : le problème de séparation de sources et de déconvolution. Nos travaux se positionnent sur plusieurs points théoriques et méthodologiques difficiles et importants (mélanges convolutifs, non linéaires bruités, Etude de critères d'indépendance, entropique, quadratique ou basé sur les fonctions caractéristiques, Etude des estimateurs et de leurs propriétés).

2.1. *Séparation de sources*

Dans un cadre semi-aveugle, nous nous intéressons à l'exploitation des propriétés de non-stationarité et de cyclostationarité (communications, signaux vibratoires), et de parcimonie (signaux de parole, signaux sismiques) ou de support borné des signaux. D'un point de vue

applicatif, un projet phare de ce thème est l'extraction non-invasive de signaux biomédicaux (extraction de l'ECG du fœtus, et extraction de l'EMG du diaphragme). Dans le domaine de la parole, nous affinons les approches fréquentielles ou temps-fréquence en exploitant des observations multimodales audio-visuelles. Des applications en traitement d'images sont également réalisées à l'aide des méthodes d'ACI pour la localisation fines des aires impliquées dans des tâches cognitives en IMR fonctionnelle, pour l'analyse d'images tatouées et pour le traitement d'images hyperspectrales en télédétection

2.2. Déconvolution

Dans le domaine de la déconvolution, nous développons des méthodes basées notamment sur les statistiques d'ordre supérieur. Nous proposons une approche très novatrice s'appuyant sur la parcimonie de la réflectivité et de blancheur dans le contexte de la déconvolution aveugle. A l'instar des méthodes de débruitage, nous pourrions prendre en compte la dimension spatiale de l'acquisition (antenne de capteurs et /ou capteurs multicomposantes). L'ajout de cette diversité spatiale pourra permettre de limiter l'influence du bruit additif important.

Axe 3 : Multidimensionnalité

Le thème de recherche intitulé « Multidimensionnalité » a pour objectif la mise en place théorique et pratique d'outils et de méthodes basées sur les aspects « multi-dimension, multi capteur », que l'on retrouve dans de nombreuses applications intéressant le groupe. Les perspectives de ce thème de recherche se déclinent principalement en deux sous thèmes.

3.1. Modèles hypercomplexes et multilinéaires

L'utilisation de modèles hypercomplexes et multilinéaires (modélisation quaternionique et biquaternionique) est valide en séparation de sources ainsi qu'en traitement d'antenne où elle permet une prise en compte de la polarisation des ondes dans les traitements. Nous étudions un modèle plus général, basé sur les algèbres de Clifford, permettant de décrire les signaux vectoriels pour obtenir une analyse fine de l'information contenue dans la dimension vectorielle des échantillons (polarisation pour les signaux sismologiques, sismiques ou électromagnétiques, bandes de fréquences pour les images multi- et hyper-spectrales, etc.). Une piste, plus prospective est l'utilisation de la théorie des groupes permettant de développer des outils de traitement des signaux vectoriels, notamment en s'appuyant, comme point de départ, sur les résultats d'analyse harmonique non-commutative. Parmi d'autres avantages, cette théorie permet la définition de la convolution entre fonctions à valeurs sur des groupes.

3.2. Généralisation des filtrages matriciels et multidimensionnels

Ce sous thème aborde la généralisation de techniques matricielles en traitement des données multidimensionnelles. Des résultats ont été obtenus lors du quadriennal autour des thèmes acoustique sous-marine et traitement des signaux sismiques. Il sera nécessaire de continuer cet effort avec notamment des validations des dernières techniques multidimensionnelles développées sur exemples réels. Ce sous-thème est en partie explicité dans le thème 4 et 5.

Axe 4 : Signal et Propagation des ondes

En acoustique sous-marine, de nombreux problèmes subsistent encore pour les systèmes acoustiques travaillant dans des environnements côtiers petits-fonds ou les phénomènes de propagation sont encore plus complexes. Nous proposons de poursuivre nos études sur les plusieurs points notamment :

Les ondes d'Ultra Basse Fréquence (UBF). Ces ondes de fréquence inférieure à 100 Hz, propagées dans la mer et/ou dans les fonds marins, constituent un moyen potentiel pour la détection et la localisation de cibles en ASM et pour la connaissance du milieu. Notre objectif est l'établissement de protocoles d'acquisition et de traitement pour la détection et la caractérisation des sources et la connaissance du milieu.

La détection, l'identification et la caractérisation d'objets enfouis par l'utilisation des ondes de surface. Nous avons montré que l'analyse de ce type d'ondes est une alternative intéressante au SONAR acoustique. La faisabilité est en partie démontrée. Nous montrons que l'apport de l'utilisation des traitements des caractéristiques multidimensionnelles des récepteurs et de la source est importante pour améliorer le pouvoir de détection.

Axe 5 : Géosciences et télédétection

5.1. Géosciences

Autour de ce thème nous retrouverons beaucoup d'interactions avec les thèmes « Séparation de sources » et « Multidimensionnalité ». Notre objectif sera de poursuivre la mise en œuvre des techniques de caractérisation d'ondes et de séparation d'ondes tenant compte explicitement du caractère multicomposante des données et de développer des modèles originaux pour les signaux vectoriels et le développement d'outils de traitement spécifiquement dédiés à ces signaux. Notamment l'utilisation des matrices polynômiales pourra permettre de tenir compte de ces différents retards sur les capteurs. La conjugaison de l'approche hypercomplexe et polynômiale permet d'améliorer encore les performances des algorithmes de traitement (de type séparation de sources ou déconvolution) des signaux vectoriels ou polarisés.

5.2. Imagerie et Télédétection Haute Résolution

Les activités sont guidées par les besoins nouveaux en termes théoriques et méthodologiques que soulèvent l'apparition des nouveaux capteurs avec des capacités et des résolutions sans cesse accrues: résolutions spatiales, fréquentielles et temporelles.

Dans le domaine visible, le principal verrou concerne le traitement des images hyperspectrales. Avec des techniques de séparation de sources, on peut extraire les signatures spectrales des différents objets présents. Dans le domaine du radar, l'avènement d'une nouvelle génération de capteurs à haute résolution spatiale va révolutionner les applications envisagées. D'une façon générale, la fusion d'information, tant au niveau données qu'au niveau méthodes devrait jouer un rôle central dans l'activité du groupe autour de ce thème.

Les deux domaines d'applications de prédilection seront maintenus :

- L'analyse du milieu urbain.
- L'étude des glaciers

CHAPITRE 3 CAHIER DES CHARGES

Les récents progrès des capteurs hyperspectraux aéroportés permettent l'acquisition d'images à haute résolution spatiale (1 mètre par pixel) et spectrale (plusieurs dizaines de bandes). En milieu urbain, il est possible d'identifier les réseaux routiers, les zones résidentielles et industrielles ainsi que les espaces verts et les différentes espèces végétales présentes. La rapidité ainsi que la précision de l'analyse sont deux critères importants pour la gestion des risques urbains, l'étude du développement démographique et la gestion de l'environnement.

Les Machines à Support Vecteur (SVM) ont donné des résultats encourageants pour analyser ce type de données [11, 8]. En s'appuyant sur les caractéristiques géométriques dans l'espace spectral des données, les SVM séparent les différentes classes par des hyperplans. Les paramètres des hyperplans sont estimés par maximisation d'une fonction convexe, ce qui confère à la solution optimalité et unicité. Cependant, la caractérisation des classes se base uniquement sur l'information spectrale (longueur d'onde) sans tenir compte de l'information spatiale (texture, forme ...). Or l'utilisation de l'information spatiale est primordiale pour l'analyse du milieu urbain. Quelques travaux étudient l'utilisation conjointe des informations pour une classification par SVM [6, 7]. Cette combinaison améliore la précision de la classification finale.



Figure 3 : Une image IKONOS

L'objectif du stage sera de développer, à partir des SVM, une méthode de classification prenant en compte les informations spatiale et spectrale. Pour cela, le processus de formation/discrétisation de l'image devra être pris en compte [9] :

Image originale → Flou optique → Discrétisation → Bruit additionnel → Image observée.

L'étude de cette chaîne d'acquisition devrait fournir une bonne modélisation des relations inter-pixels de l'image. Ces relations devront être prise en compte dans le processus de classification des SVM. La méthode sera évaluée sur des données hyperspectrales aériennes réelles.

CHAPITRE 4

MACHINES À VECTEURS SUPPORT

Ce chapitre est dédié aux Machines à Vecteurs Support (dorénavant SVM, de l'anglais Support Vector Machines). Les SVM sont des classifieurs basés sur la théorie de l'apprentissage statistique supervisé, c'est-à-dire que pour la classification il faut fournir, préalablement, un ensemble d'entraînement à l'SVM. Le principe des SVM est de séparer les données par un hyperplan. Une description plus détaillée est donnée dans la suite.

Les SVM présentent des résultats meilleurs par comparaison avec d'autres méthodes comme le « maximum likelihood » ou les réseaux neuronaux [11]. Dans le contexte de classification d'image hyperspectrale, des études ont montré l'efficacité des SVM sans nécessiter une réduction préalable de l'ensemble de caractéristiques [11].

Les SVM ont été appliqués avec succès sur plusieurs problèmes pratiques : l'imagerie biomédicale, la compression d'image ou la reconnaissance d'objets 3D. L'intérêt des SVM par rapport aux autres approches est dû, principalement, à quatre raisons :

1. Méthode non-linéaire.
2. Le problème d'apprentissage se réduit à la résolution d'un problème d'optimisation quadratique convexe (question très traitée dans la littérature).
3. Robustesse à la dimension des données
4. Utilisation d'un ensemble d'entraînement de petite taille

La suite de ce chapitre présente les deux types de SVM : les classifieurs linéaires et les non-linéaires, ainsi que la problématique de classification lorsque nous avons plusieurs classes. Finalement, la méthode pour sélectionner les paramètres du SVM est présentée.

4.1. LE CLASSIFIEUR LINÉAIRE

Les SVM font partie de la famille d'algorithmes de classification supervisé : étant donné un ensemble d'échantillons d'entraînement, chacun appartenant à une classe, le SVM cherche la fonction qui assigne chaque échantillon à sa classe correspondante. L'objectif de la théorie d'apprentissage statistique est de chercher une certaine fonction qui classe de façon satisfaisante les échantillons qui n'ont pas été utilisés pour l'entraînement, c'est-à-dire, que l'erreur de classification soit minimal.

Pour le cas de la classification linéaire, on considère un ensemble d'entraînement de N vecteurs dans un espace de caractéristiques de d dimensions $x_i \in \mathbb{R}^d$ ($i = 1, \dots, N$) avec $y_i \in \{-1, +1\}$ associé à chaque vecteur x_i . Si les deux classes sont linéairement séparables, on peut trouver un hyperplan (une surface linéaire) définie par le vecteur $w \in \mathbb{R}^d$ (vecteur normal au hyperplan) et un biais $b \in \mathbb{R}$ capable de séparer les deux classes sans erreur. Voir Figure 4.

On définit, une fonction de décision linéaire, f , associée à l'hyperplan :

$$f(x) = \langle w, x \rangle + b \quad (1)$$

Pour décider à quelle classe appartient un échantillon, il suffit d'évaluer le signe de cette fonction $\text{sgn}[f(x)]$.

L'objectif, donc, est d'estimer w et b pour que

$$y_i(\langle w, x_i \rangle + b) \geq 1, \quad i = 1, \dots, N \quad (2)$$

Plusieurs solutions sont en général possibles, un critère permettant de choisir la meilleure est la marge géométrique, qui est la distance entre deux classes et vaut $2/\|w\|$. La grandeur de cette marge est proportionnelle à la capacité de généralisation des SVM. Le plus large est la marge, plus performant sera le classifieur : l'erreur de classification sera minimal.

Ainsi, il faut maximiser $2/\|w\|$, ou ce qui est équivalent, minimiser $\|w\|/2$. Cela nous amène au problème d'optimisation quadratique suivant :

$$\begin{aligned} &\text{minimiser} && \left[\frac{\|w\|^2}{2} \right] && (3) \\ &\text{sous la contrainte} && y_i(\langle w, x_i \rangle + b) > 1, \quad i = 1, \dots, N \end{aligned}$$

Cette optimisation peut se transformer, grâce aux multipliers de Lagrange, en cette autre formulation [11] :

$$\begin{aligned} &\text{maximiser} && \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle && (4) \\ &\text{sous la contrainte} && \sum_{i=1}^N \alpha_i y_i = 0 \text{ et } \alpha_i \geq 0, \quad i = 1, \dots, N \end{aligned}$$

La résolution de ce problème d'optimisation se fait par méthodes de programmation quadratique. Des détails sur ce sujet peuvent se trouver dans [1] et [3].

Finalement, la fonction de discrimination associée à l'hyperplan optimal dévient une équation dépendant des multipliers de Lagrange et des échantillons d'entraînement :

$$f(x) = \sum_{i \in S} \alpha_i y_i \langle x_i, x \rangle + b \quad (5)$$

Où S est le sous-ensemble d'échantillons d'entraînement correspondant aux multipliers de Lagrange α_i 's non nulls.

Les contraintes supposent que les données sont linéairement séparables, mais pour des applications réelles, cela n'est pas toujours vrai. Pour traiter les données non-linéairement séparables, l'introduction d'une certaine marge d'erreur est nécessaire, ce qui change l'équation (2) :

$$\begin{aligned}
 y_i(\langle w, x_i \rangle + b) &\geq 1 - \xi_i, & i = 1, \dots, N \\
 \xi_i &\geq 0, & i = 1, \dots, N
 \end{aligned}
 \tag{6}$$

Où les ξ_i sont les variables introduites pour neutraliser l'effet de la non-séparabilité des données. Le nouveau problème d'optimisation change légèrement et devient :

$$\begin{aligned}
 \text{minimiser} & \quad \left[\frac{\|w\|^2}{2} + C \sum_{i=1}^N \xi_i \right] \\
 \text{sous la contrainte} & \quad y_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i, \quad \xi_i \geq 0
 \end{aligned}
 \tag{7}$$

Où la constante C contrôle le taux de pénalité et devient un paramètre du classifieur. De nouveau, ce problème d'optimisation se résout par programmation quadratique après la transformation par méthodes de Lagrange.

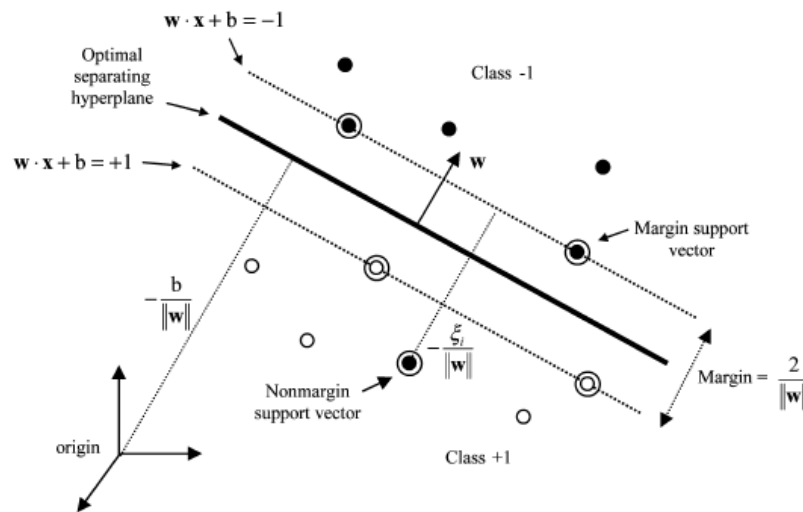


Figure 4 : Hyperplan optimal de séparation dans un SVM pour le cas linéairement non-séparables

4.2. LE CLASSIFIEUR NON-LINÉAIRE

Pour améliorer les performances de classification présentés ci-dessus il faut faire appel aux SVM non-linéaires. La base des SVM non-linéaires est l'utilisation des fonctions noyaux afin de projeter les données d'entrée sur un plus grand espace des caractéristiques pour pouvoir séparer les nouvelles variables par des hyperplans. Le choix du noyau ce fait à priori et a de l'influence sur les résultats de classification.

On peut penser la transformation comme une fonction $\Phi(x)$, non-linéaire qui projette les données sur un espace $\mathbb{R}^{d'}$ (avec $d' > d$). Alors, la formulation faite pour l'espace à dimension d , doit se faire pour l'espace à dimension d' , c'est-à-dire, que le produit scalaire $\langle x_i, x_j \rangle$ se transforme en $\langle \Phi(x_i), \Phi(x_j) \rangle$. A ce point-là, le problème principal consiste en le calcul de la fonction $\Phi(x)$, qui peut supposer un coût calculatoire très grand et qui parfois est incalculable.

L'utilisation des fonctions noyau résout ce problème, car elles s'expriment selon :

$$K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle \quad (8)$$

Ainsi, le problème est simplifié, car on évite le calcul du produit scalaire dans l'espace transformé $\langle \Phi(x_i), \Phi(x_j) \rangle$. Le problème d'optimisation s'exprime

$$\begin{aligned} \text{maximiser} \quad & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \text{sous la contrainte} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \text{ et } 0 \leq \alpha_i \leq C \end{aligned} \quad (9)$$

Le résultat final est la fonction discriminante $f(x)$ exprimée en fonction des données originales dans l'espace de caractéristiques de dimension d

$$f(x) = \sum_{i \in S} \alpha_i y_i K(x_i, x) + b \quad (10)$$

Pour plus de détails sur les fonctions noyau se trouvent dans [12].

Deux des fonctions noyau ont été étudiées dans l'annexe de [5] : le noyau polynomial et le noyau RBF. Les deux fonctions décrivant les noyaux sont :

- Noyau polynomial :

$$K(x, y) = (\langle x, y \rangle + q)^p \quad (11)$$

- Noyau RBF :

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (12)$$

Des exemples d'applications de ces noyaux peuvent se trouver dans [5].

4.3. SVM MULTI-CLASSE

Les classifieurs SVM permettent classifier des ensembles de données d'une façon binaire, c'est-à-dire, un problème à deux classes. Pourtant, en applications de télédétection, plusieurs classes sont envisagées, ce qui nous amène à considérer plusieurs solutions pour classifier les échantillons dans les classes différentes. Pour un problème de classification proposé pour m classes différentes, deux solutions sont présentées, principalement [11] et [5]:

- Un contre le reste : La stratégie "un contre le reste" représente la plus communément utilisée en termes de classification par SVM. Son fonctionnement consiste en m classifieurs binaires mis en parallèle qui s'appliquent un à chaque classe. La classification finale est le résultat de prendre la valeur de sortie maximale des m classifieurs antérieurs.
- Un contre un : cette stratégie concerne $\frac{m(m-1)}{2}$ classifieurs binaires qui sont appliqués à chaque pair de classes. Chaque échantillon est assigné à la classe qui a le maximum de

votes. Le vote pour une classe donnée s'attribue quand un classifieur assigne un échantillon à cette classe.

Pour cette dernière stratégie, le nombre de classifieurs est notamment plus grand et est seulement recommandable pour des problèmes avec un grand nombre d'échantillons.

4.4. SÉLECTION DES PARAMÈTRES

Quelques paramètres des SVM doivent être réglés à fin d'obtenir des résultats de classification satisfaisants en tant que marge d'erreur et temps de calcul. Selon le type de SVM choisi, des paramètres différents vont pouvoir se régler. Ainsi, si on travaille avec un noyau polynomial, on pourra jouer avec l'exposant q ou avec le terme b , alors que si on choisit le noyau RBF, pour différents valeurs de σ , on obtiendra des résultats de classification différents.

Un autre paramètre dont tenir compte est la valeur de C . Pour sélectionner sa valeur on peut se servir des connaissances à priori, puisque si l'ensemble de données est linéairement séparable dans l'espace de caractéristiques, on peut donner une valeur à C très grande pour ne pas avoir des erreurs d'entraînement.

Une méthode pour sélectionner les paramètres de façon automatique est la validation croisée [4]. Cette méthode consiste en une division de l'ensemble de données en deux sous-ensembles de façon à utiliser le premier de ces sous-ensembles comme ensemble d'entraînement et d'utiliser le second ensemble pour tester les résultats obtenus. Dans les itérations successives, les sous-ensembles changent à fin d'utiliser toutes les données à la fois comme sous-ensemble d'entraînement et de test. Des résultats statistiques sur la précision des résultats sont présentés et grâce à eux, on peut sélectionner les paramètres à régler.

CHAPITRE 5

TRANSFORMÉE EN ONDELETTES

Dans ce chapitre on présente la transformée en ondelettes comme une approche pour représenter une image sous plusieurs résolutions spatiales. Avec la transformée en ondelettes on peut obtenir une décomposition de l'image à plusieurs échelles dans lesquelles plusieurs niveaux de détails sont représentés. Cela a pour but l'analyse multirésolution de l'image à fin de séparer les zones plus homogènes dans l'image ou les objets qui s'étalent sur une certaine direction.

Ce chapitre présente, d'abord, les définitions mathématiques de la transformée en ondelettes. Ensuite, on va voir les bases d'ondelettes orthogonales qui nous permettent de décomposer un signal en plusieurs coefficients, appelés coefficients d'ondelettes. Dans une troisième partie du chapitre, on va centrer notre attention sur les bancs de filtres qui permettent de calculer d'une façon rapide et simple les coefficients de la décomposition en ondelettes.

Cet approche sur la transformée en ondelettes se fait sur une seule dimension, alors que les signaux qu'on veut traiter sont des images en deux dimensions. Cela nous amène à présenter les bases orthogonales séparables qui vont nous permettre décomposer une image 2D en plusieurs niveaux de résolution par une extension de la formulation présentée pour le cas d'une seule dimension.

5.1. DÉFINITIONS MATHÉMATIQUES

Pour analyser les composantes de durées différentes dans un signal (à une ou plusieurs dimensions), il est nécessaire d'utiliser des fonctions à support temporel (ou spatial) de taille variable. Pour cela, la transformée en ondelettes décompose les signaux sur une famille d'ondelettes translatées et dilatées. Une ondelette est une fonction $\psi \in L^2(\mathbb{R})$, où L^2 est l'ensemble des fonctions d'énergie finie telles que $\int |f(t)|^2 dt < +\infty$, de moyenne nulle :

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (13)$$

L'ondelette ψ est normalisée à $\|\psi\| = 1$, et centrée au voisinage de $t = 0$. On obtient une base d'ondelettes temps-fréquence en dilatant l'ondelette ψ par un facteur s et en translatant par u :

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \quad (14)$$

L'ondelette ψ est appelée ondelette mère et il s'agit, normalement d'une fonction de carré sommable [2]. On peut montrer facilement que cette base reste de norme 1 [10]. La transformée en ondelettes du signal $f \in L^2(\mathbb{R})$ au temps u et à l'échelle s est

$$Wf(u, s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-u}{s} \right) dt \quad (15)$$

où W est l'initiale de *Wavelet*, c'est-à-dire "ondelette" en anglais .

5.2. BASES D'ONDELETTES ORTHOGONALES

On peut construire des ondelettes ψ qui génèrent des bases orthonormées de $L^2(\mathbb{R})$ par translation et dilatation, comme on a vu précédemment :

$$\psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi \left(\frac{t-2^j n}{2^j} \right) \text{ avec } (j, n) \in \mathbb{Z}^2 \quad (16)$$

Les ondelettes orthogonales dilatées de 2^j reproduisent les variations d'un signal à la résolution 2^{-j} . Pour trouver une base d'ondelettes orthogonales, on doit étudier, d'abord, les approximations multirésolutions.

La modification de la résolution d'un signal permet de ne traiter que les détails intéressants pour une tâche donnée. On peut construire une pyramide multirésolution permettant de traiter l'image à basse résolution puis augmenter cette résolution pour extraire des informations pour chaque résolution.

L'approximation d'une fonction f à la résolution 2^{-j} est déterminée par une grille discrète d'échantillons fournissant des moyennes locales de f sur des voisinages de taille proportionnelle à 2^j . Plus formellement, l'approximation d'une fonction à la résolution 2^{-j} est définie comme sa projection orthogonale sur un espace $V_j \subset L^2(\mathbb{R})$. L'espace V_j regroupe toutes les approximations possibles à l'échelle 2^{-j} . Pour éviter toute ambiguïté, on définit le paramètre d'échelle 2^j comme l'inverse de la résolution 2^{-j} .

Fonction d'échelle

Pour calculer la projection orthogonale $P_{V_j} f$ sur V_j , étant cette projection l'approximation de f à la résolution 2^{-j} , il faut une base orthonormée de V_j . On obtient cette base à partir d'une fonction ϕ appelée *fonction d'échelle* [10].

La projection orthogonale de f sur V_j s'obtient par décomposition sur la base orthonormée de fonctions d'échelle :

$$P_{V_j} f = \sum_{n=-\infty}^{+\infty} \langle f, \phi_{j,n} \rangle \phi_{j,n} \quad (17)$$

Les produits scalaires

$$a_j[n] = \langle f, \phi_{j,n} \rangle \quad (18)$$

fournissent une approximation discrète de f à l'échelle 2^j .

Filtres miroirs conjugués

Une approximation multirésolution est entièrement déterminée par sa fonction d'échelle ϕ qui engendre une base orthonormée pour chaque espace V_j . À la fois, toute fonction d'échelle est déterminée par un filtre discret appelé *filtre miroir conjugué*. Selon [10], la fonction d'échelle et son filtre miroir conjugué suivent une relation exprimée par l'équation d'échelle :

$$\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} h[n]\phi(t-n) \quad (19)$$

avec

$$h[n] = \left\langle \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right), \phi(t-n) \right\rangle \quad (20)$$

Cette équation d'échelle exprime la dilatation de ϕ par 2 en fonction de ses translations entières. Selon [10], la fonction de transfert des filtres miroirs conjugués doit vérifier

$$\forall \omega \in \mathbb{R} \quad , \quad |\hat{h}(\omega)|^2 + |\hat{h}(\omega + \pi)|^2 = 2 \quad (21)$$

où $\hat{h}(\omega)$ est la transformée de Fourier de $h[n]$.

Ondelettes orthogonales

Les ondelettes orthogonales contiennent les détails nécessaires à l'augmentation de la résolution de l'approximation d'un signal. Les approximations de f aux échelles 2^j et 2^{j-1} sont données respectivement par la projection orthogonale de f sur V_j et sur V_{j-1} . Nous savons que V_j est inclus dans V_{j-1} . Soit W_j le complémentaire orthogonal de V_j dans V_{j-1} :

$$V_{j-1} = V_j \oplus W_j \quad (22)$$

La projection orthogonale de f sur V_{j-1} peut se décomposer comme la somme des projections orthogonales sur V_j et sur W_j :

$$P_{V_{j-1}}f = P_{V_j}f + P_{W_j}f \quad (23)$$

Le complément $P_{W_j}f$ fournit les "détails" de f qui existent à l'échelle 2^{j-1} et qui sont absents à l'échelle plus grossière 2^j . Dans [10] on montre que l'on peut construire une base orthonormée de W_j par translation et dilatation d'une ondelette ψ .

Soit ϕ une fonction d'échelle et h son filtre miroir conjugué. Soit ψ la fonction dont la transformée de Fourier vaut

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}}\hat{g}\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) \quad (24)$$

avec

$$\hat{g}(\omega) = e^{-i\omega} \hat{h}^*(\omega + \pi) \quad (25)$$

On pose

$$\psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t - 2^j n}{2^j}\right) \quad (26)$$

Pour toute échelle 2^j , $\{\psi_{j,n}\}_{n \in \mathbb{Z}}$ forme une base orthonormée de \mathbf{W}_j . Sur l'ensemble des échelles, $\{\psi_{j,n}\}_{(j,n) \in \mathbb{Z}^2}$ forme une base orthonormée de $L^2(\mathbb{R})$.

À partir de la fonction d'échelle ϕ on obtient l'approximation de l'image à la résolution immédiatement inférieur, alors qu'à partir de la fonction ψ on obtient les détails qu'on a perdu lors de la représentation à la nouvelle résolution.

5.3. ONDELETTES ET BANCS DE FILTRES

Les coefficients de la décomposition d'un signal sur une base orthonormée d'ondelettes se calculent par un algorithme rapide qui cascade des convolutions discrètes avec h et g dont les sorties sont sous-échantillonnées.

Une transformée en ondelettes décompose successivement chaque approximation $P_{V_j} f$ en une approximation plus grossière $P_{V_{j+1}} f$ et en coefficients d'ondelettes correspondant à $P_{W_{j+1}} f$. À la reconstruction, chaque $P_{V_j} f$ est calculée à partir de $P_{V_{j+1}} f$ et de $P_{W_{j+1}} f$.

Comme $\{\phi_{j,n}\}_{n \in \mathbb{Z}}$ et $\{\psi_{j,n}\}_{n \in \mathbb{Z}}$ sont des bases orthonormées de \mathbf{V}_j et \mathbf{W}_j la projection sur ces espaces est caractérisée par

$$a_j[n] = \langle f, \phi_{j,n} \rangle \quad \text{et} \quad d_j[n] = \langle f, \psi_{j,n} \rangle \quad (27)$$

où les coefficients $a_j[n]$ fournissent l'approximation discrète de f à l'échelle 2^j et les coefficients $d_j[n]$ représentent les détails perdus lors de la décomposition.

On peut noter la projection précédente comme un produit de convolution discret :

$$a_{j+1}[p] = \sum_{n=-\infty}^{+\infty} h[n - 2p] a_j[n] = a_j \star \bar{h}[2p] \quad (28)$$

$$d_{j+1}[p] = \sum_{n=-\infty}^{+\infty} g[n - 2p] a_j[n] = a_j \star \bar{g}[2p] \quad (29)$$

Lors de la reconstruction, on a

$$\begin{aligned}
 a_j[p] &= \sum_{n=-\infty}^{+\infty} h[p - 2n]a_{j+1}[n] + \sum_{n=-\infty}^{+\infty} g[p - 2n]d_{j+1}[n] \\
 &= \check{a}_{j+1} \star h[n] + \check{d}_{j+1} \star g[n]
 \end{aligned}
 \tag{30}$$

où $\bar{h}[p] = h[-p]$ et $\check{a}[n] = \begin{cases} a[n] & \text{pour } p = 2n \\ 0 & \text{pour } p = 2n + 1 \end{cases}$

Les coefficients a_{j+1} et d_{j+1} se calculent en prenant un échantillon sur deux de la convolution de a_j avec \bar{h} et \bar{g} respectivement, comme le montre la Figure 5. Le filtre \bar{h} enlève les hautes fréquences de la suite des produits scalaires a_j , alors que \bar{g} est un filtre passe-haut qui récupère les hautes fréquences ainsi supprimées. La reconstruction est une interpolation qui insère des zéros dans les suites a_{j+1} et d_{j+1} pour doubler leur longueur, puis un filtre, comme le montre la Figure 5.

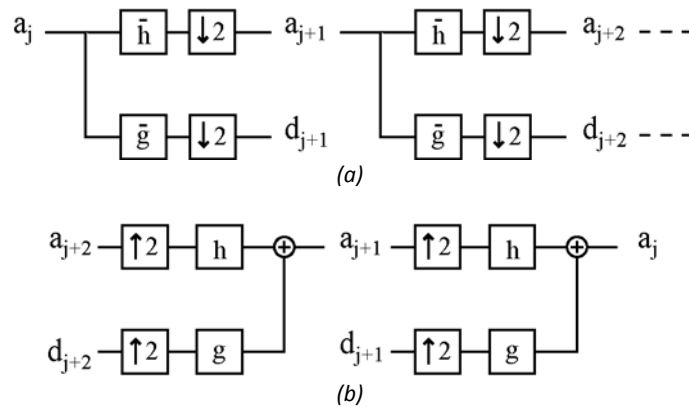


Figure 5 : (a) : Une transformée en ondelettes rapide se calcule par cascades de filtrages par \bar{h} et \bar{g} suivis de sous-échantillonnage d'un facteur 2. (b) : La transformée en ondelettes inverse reconstruit \mathbf{a}_j en insérant des zéros entre les échantillons de \mathbf{a}_{j+1} et \mathbf{d}_{j+1} , puis en les filtrant et en prenant la somme des sorties

5.4. BASES D'ONDELETTES SÉPARABLES

Jusqu'à présent on n'a considéré que les signaux à une dimension. Dans cette section on va traiter les ondelettes séparables pour faire la décomposition des signaux à deux dimensions : les images.

On peut associer une base séparable d'ondelettes orthonormées de $L^2(\mathbb{R}^2)$ à toute base d'ondelettes orthonormées $\{\psi_{j,n}\}_{(j,n) \in \mathbb{Z}^2}$ de $L^2(\mathbb{R})$:

$$\{\psi_{j_1, n_1}(x_1) \psi_{j_2, n_2}(x_2)\}_{(j_1, j_2, n_1, n_2) \in \mathbb{Z}^4}
 \tag{31}$$

où n_1 et n_2 sont les directions d'étalement et j_1 et j_2 , les échelles de chaque ondelette.

Les fonctions $\psi_{j_1, n_1}(x_1) \psi_{j_2, n_2}(x_2)$ mélangent l'information à deux échelles différentes 2^{j_1} et 2^{j_2} le long de x_1 et de x_2 , ce que l'on veut éviter. On va considérer, donc $j_1 = j_2 = j$.

On peut appliquer le même formalisme appliqué précédemment en deux dimensions : la notation de multirésolution se formalise avec des projections orthogonales sur des espaces de tailles variables. L'approximation d'une image $f(x_1, x_2)$ à la résolution 2^{-j} est définie comme la projection de f sur un espace V_j^2 inclus dans $L^2(\mathbb{R}^2)$. L'espace V_j^2 est l'ensemble de toutes les approximations à la résolution 2^{-j} . La taille de V_j^2 diminue quand la résolution diminue.

On considère le cas particulier des multirésolutions séparables. Soit $\{V_j\}_{j \in \mathbb{Z}}$ une multirésolution de $L^2(\mathbb{R})$. Une multirésolution séparable en dimension 2 est constituée des produits tensoriels d'espaces

$$V_j^2 = V_j \otimes V_j \quad (32)$$

La base orthonormée de V_j^2 engendrée par la fonction d'échelle ϕ s'exprime comme

$$\left\{ \phi_{j,n}^2 = \phi_{j,n_1}(x_1) \phi_{j,n_2}(x_2) = \frac{1}{2^j} \phi\left(\frac{x_1 - 2^j n_1}{2^j}\right) \phi\left(\frac{x_2 - 2^j n_2}{2^j}\right) \right\}_{n \in \mathbb{Z}^2} \quad (33)$$

Elle s'obtient par changement d'échelle de 2^j sur la fonction d'échelle séparable $\phi^2(x) = \phi(x_1) \phi(x_2)$, et par translation sur une grille de pas 2^j .

Bases d'ondelettes en dimension 2

Une base orthonormée d'ondelettes séparables de $L^2(\mathbb{R}^2)$ est construite à partir de produits séparables d'une fonction échelle ϕ et d'une ondelette ψ . La fonction d'échelle est associée à une approximation multirésolution $\{V_j\}_{j \in \mathbb{Z}}$ de $L^2(\mathbb{R})$. Soit $\{V_j^2\}_{j \in \mathbb{Z}}$ la multirésolution séparable de $L^2(\mathbb{R}^2)$ définie par $V_j^2 = V_j \otimes V_j$. Soit W_j^2 l'espace de détails égal au complémentaire orthogonal de V_j^2 dans V_{j-1}^2 :

$$V_{j-1}^2 = V_j^2 \oplus W_j^2 \quad (34)$$

On construit une base orthonormée d'ondelettes de $L^2(\mathbb{R}^2)$ à partir d'une base d'ondelettes pour chaque espace de détails W_j^2 . Soit ϕ une fonction d'échelle et ψ l'ondelette correspondante engendrant une base d'ondelettes orthonormée de $L^2(\mathbb{R})$. On définit trois ondelettes :

$$\psi^1(x_1, x_2) = \phi(x_1) \psi(x_2), \quad \psi^2(x_1, x_2) = \psi(x_1) \phi(x_2), \quad \psi^3(x_1, x_2) = \psi(x_1) \psi(x_2)$$

et on pose, pour $1 \leq k \leq 3$

$$\psi_{j,n}^k(x) = \frac{1}{2^j} \psi^k\left(\frac{x_1 - 2^j n_1}{2^j}, \frac{x_2 - 2^j n_2}{2^j}\right) \quad (35)$$

La famille d'ondelettes

$$\{\psi_{j,n}^1, \psi_{j,n}^2, \psi_{j,n}^3\}_{n \in \mathbb{Z}^2} \quad (36)$$

est une base orthonormée de W_j^2 , et

$$\{\psi_{j,n}^1, \psi_{j,n}^2, \psi_{j,n}^3\}_{(j,n) \in \mathbb{Z}^3} \quad (37)$$

est une base orthonormée de $L^2(\mathbb{R}^2)$.

Transformée en ondelettes rapide en dimension 2

L'algorithme de transformée en ondelettes rapide de la section 5.3 est étendu au cas bidimensionnel. À chaque échelle 2^j et pour tout $n = (n_1, n_2)$, on note

$$a_j[n] = \langle f, \phi_{j,n}^2 \rangle \quad \text{et} \quad d_j^k[n] = \langle f, \psi_{j,n}^k \rangle \quad \text{pour } 1 \leq k \leq 3$$

Pour toute paire de filtres monodimensionnels $y[m]$ et $z[m]$ on écrit le filtre produit $yz[n] = y[n_1]z[n_2]$, et $\bar{y}[m] = y[-m]$. Soit $h[m]$ et $g[m]$ les filtres miroirs conjugués associés à l'ondelette ψ .

Les coefficients d'ondelettes à l'échelle 2^{j+1} se calculent à partir de a_j avec des convolutions bidimensionnelles séparables et des sous-échantillonnages, comme dans la section 5.3 :

$$a_{j+1}[n] = a_j \star \bar{h}\bar{h}[2n] \quad (38)$$

$$d_{j+1}^1[n] = a_j \star \bar{h}\bar{g}[2n] \quad (39)$$

$$d_{j+1}^2[n] = a_j \star \bar{g}\bar{h}[2n] \quad (40)$$

$$d_{j+1}^3[n] = a_j \star \bar{g}\bar{g}[2n] \quad (41)$$

Une convolution séparable bidimensionnelle peut se factoriser en convolutions monodimensionnelles le long des lignes et des colonnes de l'image. Les lignes de a_j sont d'abord convoluées avec \bar{h} et \bar{g} puis sous-échantillonnées d'un facteur 2. Les colonnes de ces deux nouvelles images sont ensuite convoluées avec \bar{h} et \bar{g} respectivement et sous-échantillonnées, ce qui donne les quatre images sous-échantillonnées $a_{j+1}, d_{j+1}^1, d_{j+1}^2, d_{j+1}^3$.

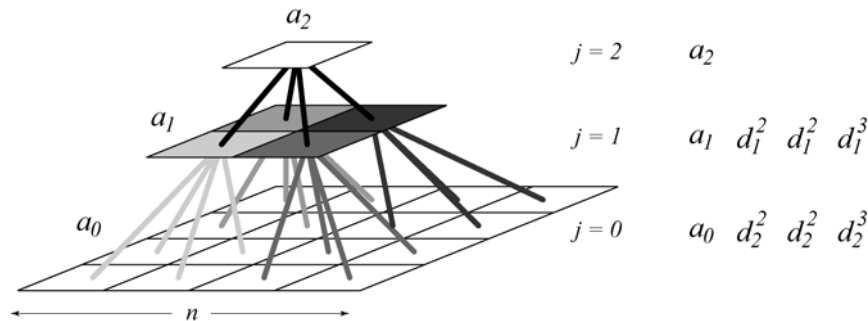


Figure 6 : Schéma multirésolution avec les coefficients d'ondelettes pour 2 résolutions différentes et le vecteur de caractéristiques

On note $\check{y}[n] = \check{y}[n_1, n_2]$ l'image dont la taille est deux fois celle de $y[n]$ obtenue en insérant une ligne de zéros et une colonne de zéros entre les lignes et les colonnes consécutives de

$y[n]$. L'approximation a_j est reconstruite à partir de l'approximation a_{j+1} à l'échelle plus grossière et les coefficients d'ondelettes d_{j+1}^k , avec une formule de convolutions bidimensionnelles que l'on déduit des formules monodimensionnelles de reconstruction (30)

$$a_j[n] = \check{a}_{j+1} \star hh[n] + \check{d}_{j+1}^1 \star hg[n] + \check{d}_{j+1}^2 \star gh[n] + \check{d}_{j+1}^3 \star gg[n] \quad (42)$$

ANALYSE SPATIALE MULTIRÉSOLUTION

Dans ce chapitre on va aborder le traitement fait sur les images pour extraire l'information spatiale. Grâce à la transformée en ondelettes qu'on a vu dans le chapitre précédent, on peut représenter un signal, dans notre cas une image, sous plusieurs résolutions, ce qu'on appelle l'analyse multirésolution.

La plupart des images qu'on va traiter sont des images hyperspectrales, où l'information est répartie sur plusieurs bandes. On peut voir les images hyperspectrales comme une suite d'images 2D empilées pour chaque bande. Pour faire une analyse multirésolution on a besoin d'une seule image 2D. Une possible solution serait de faire une analyse multirésolution de chaque bande, mais son coût algorithmique rend cette solution inappropriée. Il faut penser, alors, à représenter l'image hyperspectrale en une seule image 2D contenant la plupart de son information. La méthode choisie pour cela est l'Analyse en Composantes Principales. Dans une première partie de ce chapitre on va faire une introduction à ce méthode de réduction des données, en donnant sa base mathématique, ainsi que voir des résultats graphiques de l'analyse.

Dans une deuxième partie de ce chapitre, on va aborder l'analyse multirésolution. On va voir la structure des données qu'on va fournir au classifieur, après avoir fait la décomposition en ondelettes.

Finalement, dans ce chapitre on présente les problèmes de la décomposition en ondelettes pour les images finies : l'effet de bords. L'utilisation des filtres discrets pour l'obtention des coefficients d'ondelettes favorise l'apparition de défauts dans la décomposition, spécialement aux bords de l'image. On verra que l'utilisation de certains filtres permettent la réduction de cet effet, mais ils ne peuvent pas l'annuler totalement.

6.1. ANALYSE EN COMPOSANTES PRINCIPALES

L'Analyse en Composantes Principales (ACP) est une technique classique dans l'analyse statistique de données. L'ACP a pour but la décorrélation des variables x_1, \dots, x_2 appartenant à un vecteur donné $x \in \mathbb{R}^n$. Les variables du vecteur projeté $y = P^t x$ sont décorrélées. Cela veut dire que sa matrice de covariance $\Sigma_y = E[y_c y_c^t]$ est diagonale, où y_c est le vecteur y centré. Le calcul de la matrice de covariance peut s'écrire comme :

$$\begin{aligned} \Sigma_y &= E[(y - m_y)(y - m_y)^t] = E[(P^t x - P^t m_x)(P^t x - P^t m_x)^t] \\ &= P^t E[(x - m_x)(x - m_x)^t] P = P^t \Sigma_x P \end{aligned} \quad (43)$$

Σ_x est une matrice réelle, symétrique de dimension finie. Σ_x peut se diagonaliser par une matrice orthogonale M ($M^t = M^{-1}$) : $M^{-1} \Sigma_x M = \Sigma_y$. L'obtention de la matrice orthogonale P se réduit au calcul des valeurs propres (λ) et des vecteurs propres (v) de norme unitaire :

$$\begin{aligned} \lambda v &= \Sigma_x v \\ \|v\|^2 &= 1 \end{aligned} \tag{44}$$

Finalement, on obtient P à partir de l'ensemble de vecteurs propres v de Σ_x , avec un vecteur propre par colonne.

Les valeurs propres obtenus représentent la variance du vecteur y , i.e., $\text{var}(y_i) = \lambda_i$. Les valeurs propres sont stockés en ordre décroissant $\lambda_1 > \lambda_2 \dots > \lambda_n$. La réduction des caractéristiques se fait en tenant compte que la variance la plus grande contribue de façon plus importante sur la représentation. Par conséquent, les variables associés à des valeurs propres les plus hauts doivent être considérées après la réduction de caractéristiques.

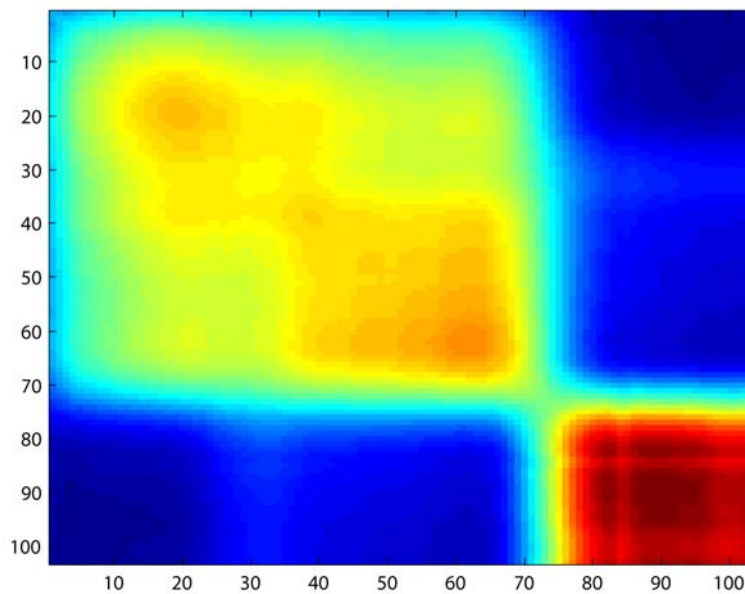


Figure 7 : Matrice de covariance pour l'image hyperspectrale correspondant à l'Université de Pavie (Italie)

Sur la Figure 7 on a la représentation de la matrice de covariance d'une image hyperspectrale de 103 bandes, matrice carrée et symétrique de dimension 103 x 103. Si on diagonalise cette matrice et on obtient les valeurs propres, on peut représenter la contribution de chaque valeur propre pondéré par la somme totale des valeurs propres (voire Figure 8). On observe que la majeure contribution on l'obtient avec les 3 premiers valeurs propres.

Le pourcentage de concentration de la variance pour les trois premiers valeurs propres est du 98,96% :

$$\frac{\sum_{i=1}^3 \lambda_i}{\sum_{i=1}^{103} \lambda_i} = 0,9896 \tag{44}$$

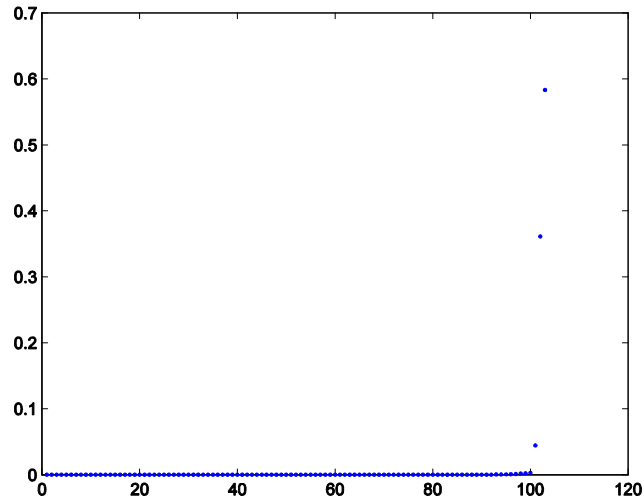


Figure 8 : Contribution de chaque valeur propre pondérée par la somme totale des 103 valeurs propres

La représentation dans les différents composantes principales s'obtient en projetant l'image originale sur chaque axe engendré par la base des vecteurs propres. Pour chaque vecteur, on obtiendra une représentation différente. Sur la Figure 9 on peut observer deux représentations de la même image pour la première et troisième composante principale.



Figure 9 : Projection sur la 1^{ère} et 3^{ème} composante principale

6.2. STRUCTURE MULTIRÉSOLUTION

L'analyse multirésolution d'une image consiste à l'extraction de l'information à plusieurs échelles pour obtenir des éléments qui ressortent à certaines résolutions. À chaque échelle, on obtient des coefficients qui caractérisent l'image au niveau pixel, c'est-à-dire, que chaque pixel est représenté par une suite de coefficients. Au fur et à mesure qu'on prend une échelle plus grande, ce qui correspond à une résolution plus grossière, les pixels appartenant à un voisinage vont se caractériser par les mêmes coefficients.

Les coefficients qui caractérisent chaque échelle s'obtiennent directement de la transformée en ondelettes. Comme vu dans le chapitre précédent, à chaque échelle on obtient 4 coefficients à partir de l'image correspondante à l'échelle immédiatement inférieure : un coefficient (a_{j+1}) correspondant à l'approximation de l'image à une échelle plus grossière et 3 coefficients (d_{j+1}^1 , d_{j+1}^2 et d_{j+1}^3) correspondants aux détails horizontaux, verticaux et diagonaux. Pour la première échelle, les coefficients sont directement la valeur de chaque pixel.

Soit Λ_x le vecteur qui contient les coefficients de la décomposition en ondelettes aux différents niveaux de résolution pour chaque pixel. Pour chaque pixel de l'image on obtient un vecteur dont la longueur est définie selon le niveau de résolution souhaité :

$$\text{longueur}(\Lambda_x) = 1 + 4 \times (\log_2 n - L) \quad (45)$$

où n est la taille de l'image et L est l'exposant de la taille de l'image après la dernière décomposition en ondelettes.

Ce dernier paramètre nous permet d'obtenir une analyse multirésolution plus ou moins détaillé. Par exemple, pour une image carrée de 256 pixels, on peut décomposer l'image en ondelettes jusqu'à obtenir une image de taille 2 x 2 pixels, ce qui va nous donner un vecteur de taille $\text{longueur}(\Lambda_x) = 1 + 4 \times (8 - 2) = 25$.

Sur la Figure 10 on a une représentation du noyau spatial pour une niveau de résolution égal à 2, ce qui correspond à une taille minimale de l'image de 64 pixels. Pour chaque pixel, on a un vecteur de 9 composantes avec la représentation de l'image à 3 échelles différentes et 6 détails de la décomposition.

Il faut remarquer que Λ_x a été construit à partir d'une image 2D, c'est-à-dire après avoir fait l'Analyse en Composantes Principales et avoir projeté les données sur le premier axe principal. En utilisant plusieurs composantes principales, on peut construire un vecteur de caractéristiques Λ_x plus grand :

$$\Lambda_x = [\Lambda_x^{PC_1}, \Lambda_x^{PC_2}, \Lambda_x^{PC_3} \dots] \quad (46)$$

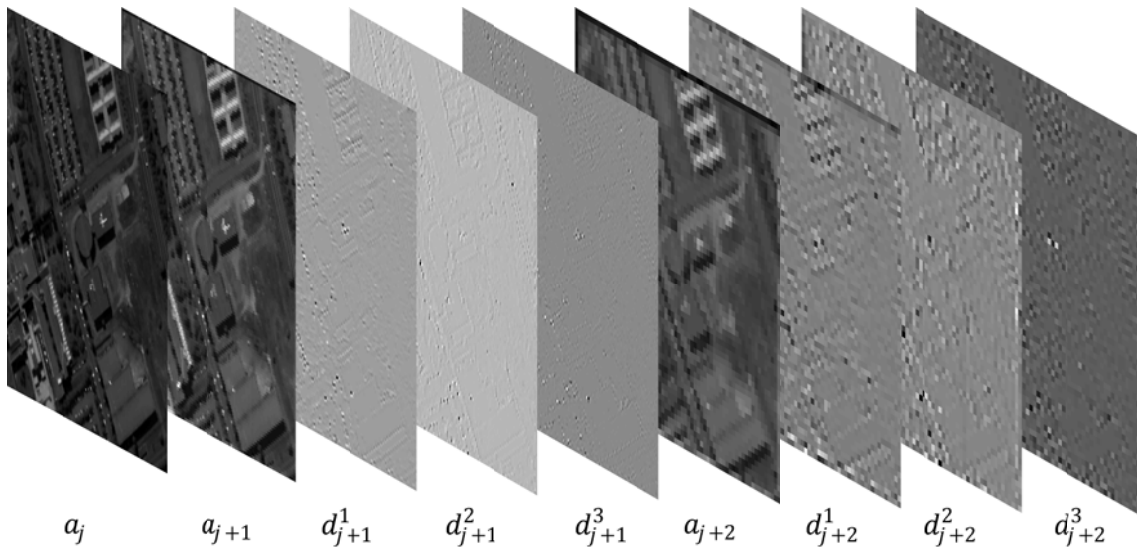


Figure 10 : Structure multirésolution avec coefficients de la décomposition

6.3. EFFET DE BORDS

La décomposition en ondelettes a été présentée d'un point de vue théorique dans le chapitre précédent. Cependant, lorsqu'on l'applique sur des images réelles, on observe tout de suite un phénomène gênant dû à l'utilisation de signaux de durée finie ou dans le cas des images, à support borné.

Étant donné que la transformée en ondelettes se base sur l'utilisation de filtres discrets, après le filtrage des images, des défauts apparaissent sur les bords de l'image. Pour éviter ce phénomène, on doit effectuer un filtrage tenant compte de la nature finie du signal d'entrée. Dans la toolbox utilisée pour l'analyse en ondelettes, un filtre est prévu à cet effet.

On peut observer les effets de bords sur l'image de la Figure 11. L'image correspond à la transformée en ondelettes après l'utilisation d'un filtre qui ne tient pas en compte la nature finie de l'image.

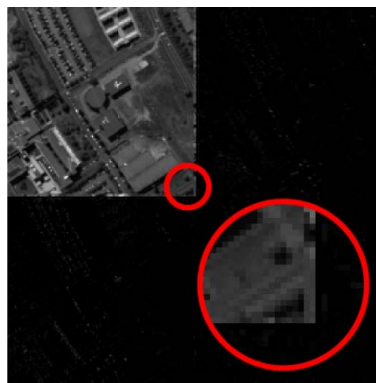


Figure 11 : Transformée en ondelettes à l'échelle 2 avec zoom sur l'effet de bords

Au fur et à mesure qu'on diminue l'échelle de la transformée en ondelettes, l'effet de bords est plus fort, dû au fait qu'on fait plusieurs filtrages en cascade (Figure 12).



Figure 12 : Transformée en ondelettes à l'échelle 3 avec zoom sur l'effet de bords

L'effet de bords peut devenir très gênant lorsqu'on veut obtenir une image à une échelle très petite. Dans ce cas, les coefficients qui caractérisent l'image à cette résolution sont faux, puisque l'erreur est plus grande que l'information utile.

La classification spectrale a été abordée dans une première partie de ce stage avec l'étude et compréhension, d'abord, du code Matlab utilisé pour la classification avec SVM et ensuite avec l'adaptation de ce code, utilisant une toolbox en C, en une autre toolbox exclusivement sous Matlab [13]. Dans ce chapitre il est présenté la procédure de classification de deux images hyperspectrales réelles avec ce nouveau classifieur.

7.1. NOYAU SPECTRAL

Dans le chapitre 4 on a étudié les classifieurs basés sur les SVM. Parmi les noyaux non linéaires qu'on a vu dans la section 4.3, on a choisi le noyau RBF pour classifier les données hyperspectrales, puisqu'il a des meilleures performances face à d'autres noyaux. Soit x l'image à classifier contenant dans chaque pixel un vecteur avec l'information correspondant à chaque bande, on peut écrire le noyau spectrale comme l'application suivante :

$$K_{\sigma_{spect}}^{spect} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0,1] \tag{47}$$
$$(x, y) \rightarrow \exp\left(-\frac{\|x-y\|^2}{2\sigma_{spect}^2}\right)$$

où σ_{spect} est la variance du noyau gaussien et elle agit comme un paramètre à régler à fin d'obtenir des résultats de précision optimaux.

7.2. RÉSULTATS EXPÉRIMENTAUX

L'image de test choisie est une image du satellite ROSIS-03 (*Reflective Optics System Imaging Spectrometer*) correspondant à une zone de l'Université de Pavie, en Italie. La taille de l'image est de 610 fois 340 pixels et le nombre de bandes est de 103. Une représentation de l'image avec seulement 3 bandes correspondant au rouge, vert et bleu est donnée sur la Figure 13.

Pour la classification par SVM, il nous faut un ensemble d'échantillons d'entraînement pour construire l'hyperplan.

Sur la Figure 14 on a la représentation des ensembles d'entraînement et de test. Il faut remarquer les différents couleurs correspondant chacun à une classe différente. Pour le cas de l'image utilisée, on différencie 9 couleurs, c'est-à-dire, 9 classes différentes.



Figure 13 : Image ROSIS de la zone de l'Université de Pavie, en Italie, composée de 3 bandes

Après l'entraînement , on teste le SVM avec l'ensemble de test. On peut observer le résultat sur la Figure 15 et le comparer avec l'ensemble de test de la Figure 14. En général le SVM classe bien l'ensemble de test. On remarque seulement une zone de mauvaise classification en bas de l'image.

Pour déterminer la précision de la classification, on se sert de la matrice de confusion. À partir de la matrice de confusion, on peut extraire trois paramètres utilisés pour déterminer cette précision : la précision globale (*Over-all Accuracy, OA*), la précision moyenne (*Average Accuracy, AA*), la précision par classe (*Class Accuracy, CA*) et le coefficient Kappa (K). Le paramètre OA est le pourcentage de pixels bien classifiés, AA représente la moyenne pour toutes les classes du pourcentage de pixels bien classifiés dans une classe (précision de classe), CA est le pourcentage de pixels bien classifiés pour une classe donnée et le coefficient Kappa est une mesure statistique de la concordance, plus robuste que les deux antérieurs, puisqu'il tient en compte la classification due à la chance. Plus de détails sur les paramètres et la matrice de confusion peuvent se trouver dans l'Annexe B de [5].

Le SVM utilisé pour la classification est basé sur le noyau gaussien. Sur ce noyau, on peut régler comme paramètre la variance de la gaussienne. Après plusieurs essais, on a choisi une valeur

pour la variance égal à l'unité puisqu'on éteignait un valeur maximale du coefficient Kappa. Sur la Figure 15 on présente le résultat de classification pour l'ensemble de test.

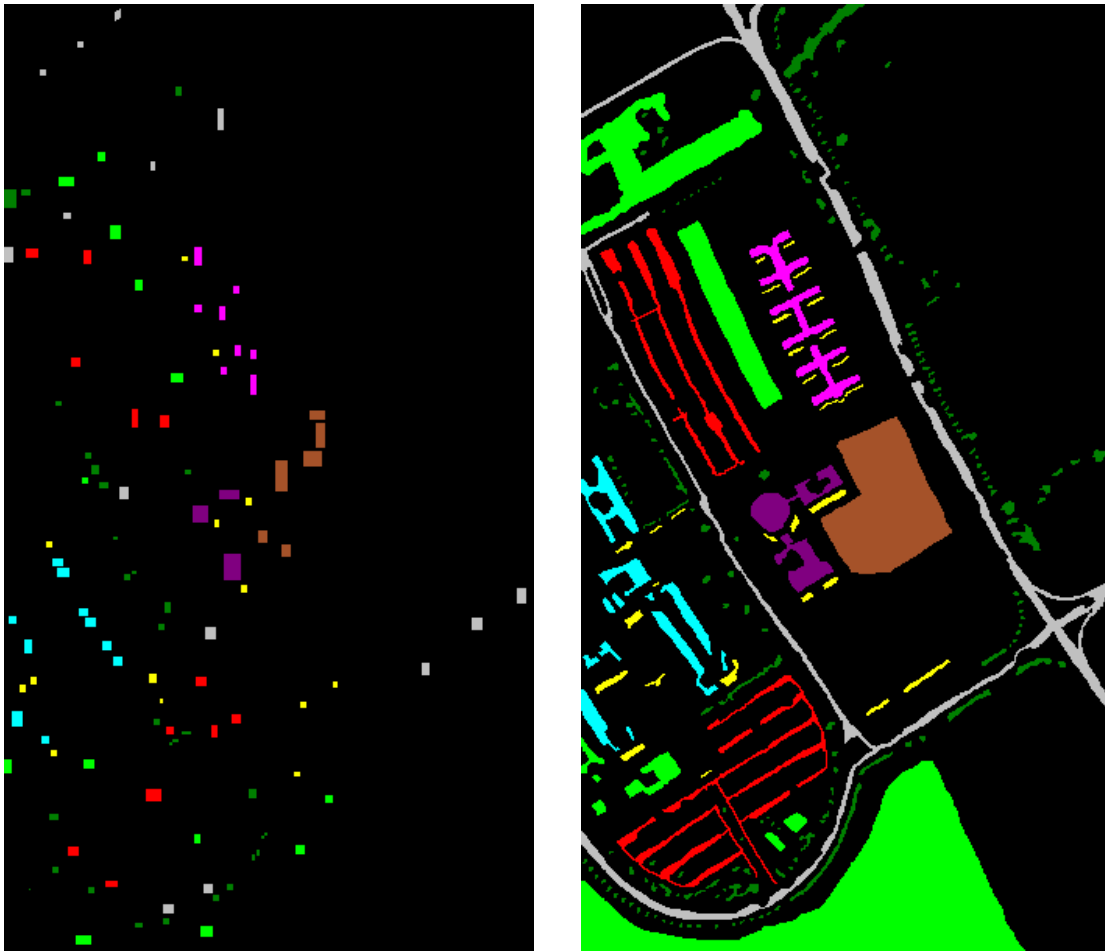


Figure 14 : Ensembles d'entraînement et de test. Les couleurs indiquent les différentes classes : *asphalte, prairie, gravier, arbre, plaque métallique, terrain dépouillé, bitume, brique* et *ombre*

Pour observer comment l'effet des paramètres du noyau ont de l'influence sur la précision de classification, plusieurs essais ont été faits avec 5 valeurs différents pour la variance du noyau gaussien. Ces résultats sont présentés sur la Table 1, ainsi que les résultats de la première simulation, où la variance est la racine carrée de la moitié du nombre de bandes de l'image

Les meilleurs résultats sont atteints pour une variance du noyau égal à l'unité, comme on peut observer sur la Table 1. Quand on utilise une variance plus grande, les valeurs sont moins précis, ainsi qu'une valeur trop petite nous fournit des résultats aussi mauvais.

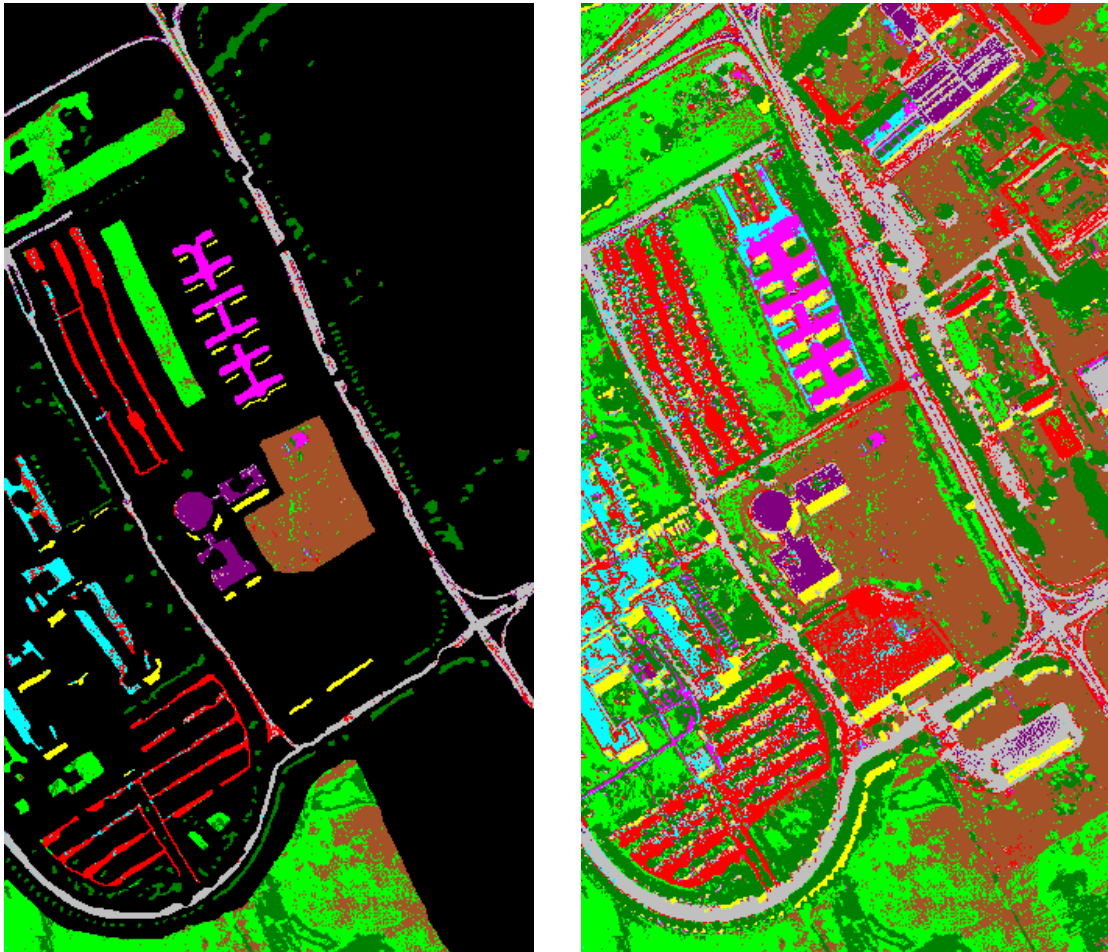


Figure 15 : Résultats de classification sur l'ensemble d'entraînement et sur l'image complète

Selon la classe qu'on veut classifier, on obtient des valeurs optimaux de sigma différents. Par exemple, pour la classe 8, la valeur $\sigma = 1$ est peut appropriée, alors que pour la classe 3 est la plus convenable.

σ	OA	AA	K	CA								
				1	2	3	4	5	6	7	8	9
0.125	72.93	82.31	66.65	75.98	59.52	60.51	99.31	86.02	81.83	89.70	91.53	96.41
0.5	79.67	87.02	74.59	79.94	69.25	67.70	98.27	99.48	92.30	90.00	91.09	95.14
1	80.32	88.29	75.40	81.54	69.08	74.65	98.01	99.11	93.14	91.95	90.93	96.20
2	79.75	88.25	74.71	81.69	67.52	74.42	94.91	99.41	95.07	91.58	91.47	98.20
4	77.00	87.51	71.57	77.30	62.54	73.56	96.15	99.55	94.79	91.80	91.96	99.89
7.17	73.74	86.40	67.96	74.72	55.68	69.60	97.13	99.63	96.40	91.88	92.64	99.89

Table 1 : Résultats de la simulation pour plusieurs valeurs du paramètre σ

Finalement, on teste la classification pour l'image complète (Figure 14). On peut remarquer certaines zones où la classification est très bruitée. Dans les chapitres suivants, on va essayer d'obtenir une classification plus homogène en insérant l'information spatiale.

Un autre test a été effectué avec une autre image hyperspectral obtenue avec le même capteur que l'image précédente, correspondant au centre de la ville de Pavie. La taille de

l'image est de 1096 fois 715 pixels et le nombre de bandes est de 102. Une représentation de l'image avec seulement 3 bandes correspondant au rouge, vert et bleu est donnée sur la Figure 16.



Figure 16 : Image ROSIS du centre de la ville de Pavie, en Italie, composée de 3 bandes

Sur la Figure 17 on a les ensembles d'entraînement et de test. On différencie, aussi 9 couleurs différents pour 9 classes, mais cette fois les classes sont différentes. Le résultat de classification pour l'ensemble de test, ainsi que sur l'image complète sont représentées sur la Figure 18. Si l'on compare avec l'image correspondant à l'Université de Pavie, on observe, d'un point de vue subjectif, des mieux résultats par rapport à l'homogénéité de la classification.

σ	OA	AA	K
0,5	97,95	95,05	97,10
1	98,00	95,62	97,18
2	98,52	96,23	97,90

Table 2 : Résultats de la simulation pour plusieurs valeurs du paramètre σ

Si on fait l'analyse objectif, en extrayant les coefficients de la matrice de confusion, on obtient effectivement des valeurs plus hauts que dans le cas précédent. On doit remarquer que pour cette image, on obtient les valeurs les plus grands du coefficient Kappa pour une variance égale à 2 (Table 2). Cela nous montre que pour chaque image, il faut régler les paramètres du noyau à fin d'obtenir une classification plus précise.

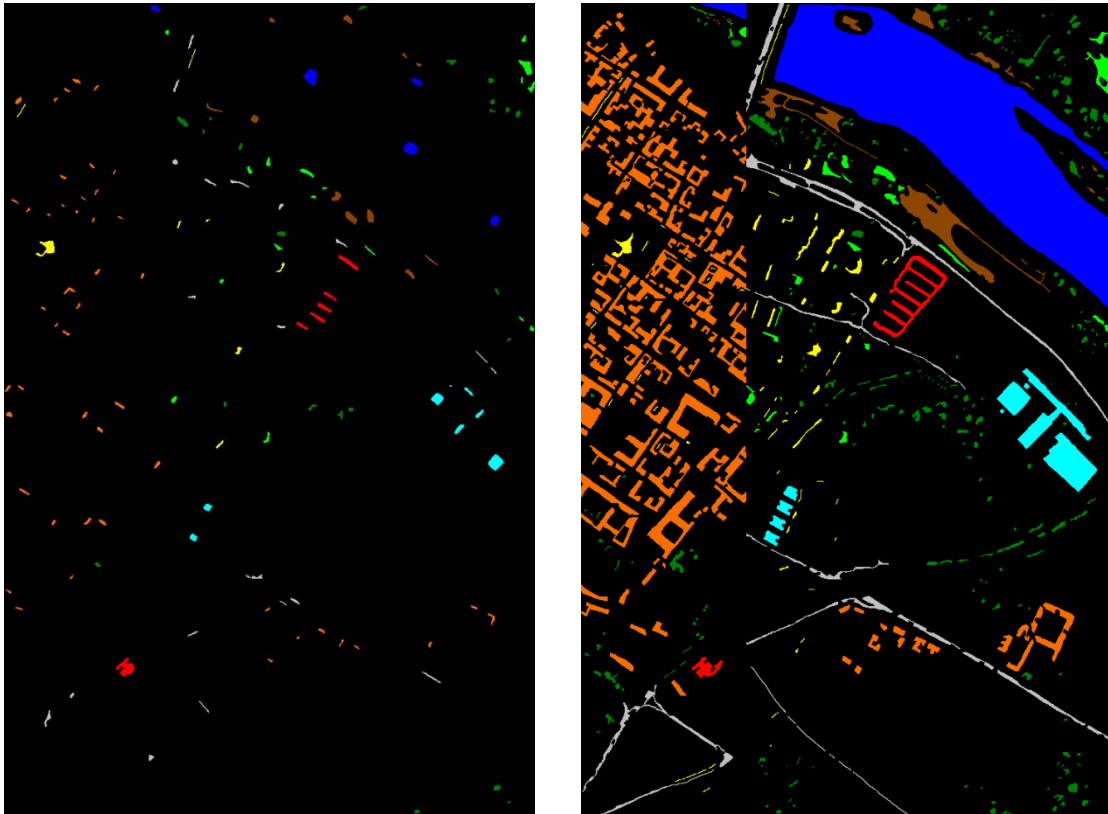


Figure 17 : Ensembles d'entraînement et de test. Les couleurs indiquent les différentes classes : eau, arbres, prairie, brique, terrain, asphalt, bitume, tuile et ombre



Figure 18 : Résultats de classification sur l'ensemble d'entraînement et sur l'image complète

Dans une deuxième partie du stage, la classification tenant compte l'information spatiale a été considérée. Ce chapitre présente, d'abord le noyau utilisé pour la classification spatiale ainsi que les paramètres qu'on peut régler pour obtenir une classification plus précise, et finalement les résultats de classification pour deux images : une première image hyperspectrale déjà utilisée dans les chapitres précédents et une deuxième image panchromatique à haute résolution spatiale.

On doit remarquer que pour la classification spatiale et lors de l'utilisation de la décomposition en ondelettes, l'utilisation d'images carrées et dyadiques (de taille égal à une puissance de 2) est nécessaire. Pour cela, on va découper les images sous deux contraintes : on voudra obtenir l'image de taille la plus grande possible et que cette image contienne toutes les classes.

8.1. NOYAU SPATIAL

Comme on a vu dans la section 6.3, l'information spatiale est présentée sous une structure multirésolution où chaque pixel correspond un vecteur de longueur égale au niveau de résolution utilisé. Ce vecteur contient les coefficients qui caractérisent l'image à plusieurs résolutions. Le noyau inclut, aussi la projection dans les axes principales de l'image hyperspectrale pour la rendre bidimensionnelle.

On a une construction différente du noyau selon si l'image à classifier s'agit d'une image hyperspectrale ou panchromatique. Pour les images hyperspectrales, avant de faire la décomposition en coefficients d'ondelettes, il faut faire une ACP et projeter l'image sur un ou plusieurs axes principaux. On a, d'abord la possibilité de choisir un premier paramètre : le nombre de composantes principales que l'on garde. Après l'ACP, au moment de faire la décomposition en ondelettes, on peut choisir un deuxième paramètre : le niveau de résolution de la décomposition.

Pour les images panchromatiques, ne possédant qu'une bande spectrale, il n'est pas nécessaire d'appliquer l'ACP, ce qui va nous permettre de réduire un paramètre.

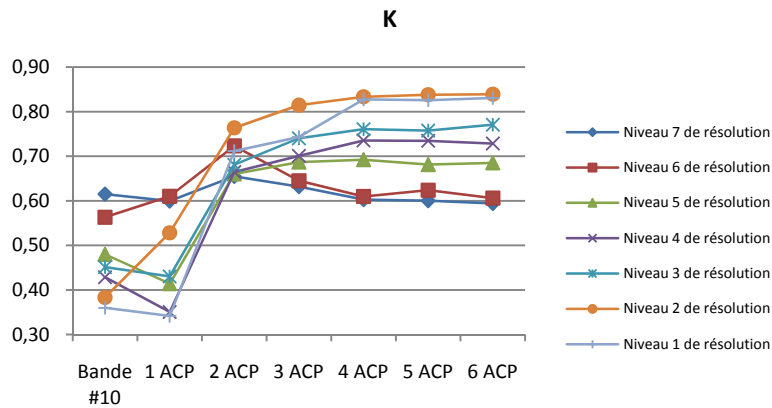
Soit Λ_x le vecteur avec les caractéristiques spatiales de chaque pixel qui dépend de x , l'image à classifier, on peut définir le noyau spatial comme l'application

$$K_{\sigma_{spat}}^{spat} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0,1] \quad (48)$$

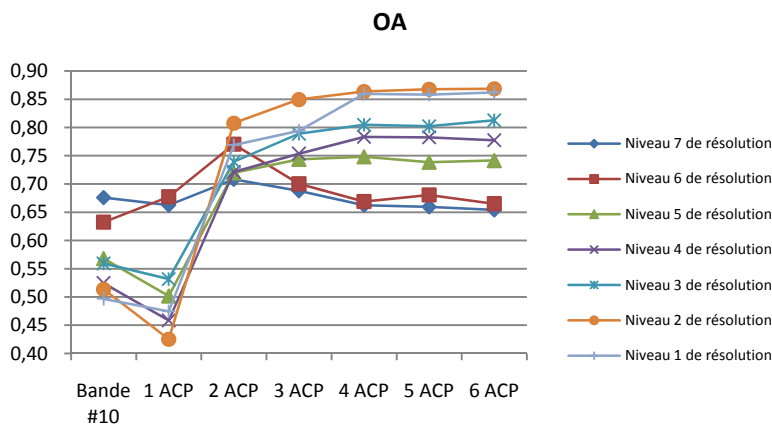
$$(x, y) \rightarrow \exp\left(-\frac{\|\Lambda_x - \Lambda_y\|^2}{2\sigma_{spat}^2}\right)$$

où Λ_x dépend du numéro de composantes principales et du niveau de résolution.

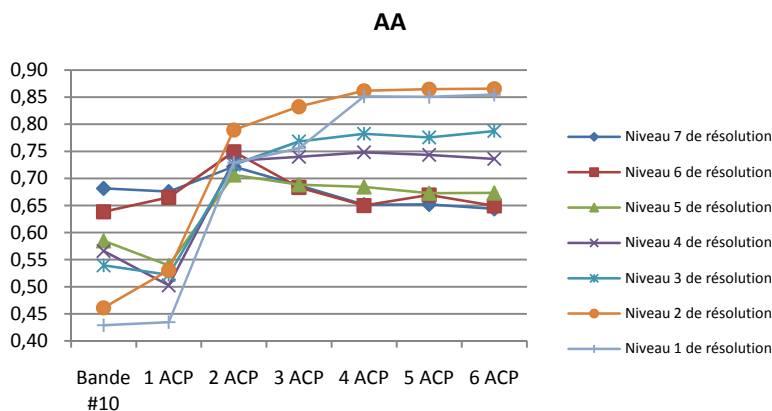
On cherche à optimiser ces 3 paramètres pour l'image hyperspectral de l'Université de Pavie, en Italie. On se base sur les paramètres définies dans les chapitre 4 pour l'analyse de la précision de la classification : la précision globale (OA), la précision moyenne (AA) et le coefficient Kappa (K).



(a)



(b)



(c)

Figure 19 : (a) Réalisations du coefficient Kappa pour différents niveaux de résolution et prenant plusieurs projections sur les axes principales. (b) Idem pour la précision globale. (c) Idem pour la précision moyenne.

Les résultats optimaux (Figure 19) ont été obtenus en prenant les projections de l'image sur les 6 premières axes principales, pour un niveau de résolution égal à 2, ce qui correspond à une seule décomposition en ondelettes et pour une valeur de la variance du noyau gaussien égale à 0,5. Pour cette configuration de paramètres, les valeurs de précision ont été les suivants :

OA	AA	K
86,84	86,56	83,89

Table 3 : Paramètres de précision résultant de la simulation optimale

Un aspect à remarquer sur la Figure 19 est l'inversion des paramètres de précision lorsqu'on augmente le nombre de projections sur les axes principales et on diminue le niveau de résolution : étant donné un nombre de projections élevé, on éteint le maximum de précision pour un bas niveau de résolution, alors que pour un nombre de projections bas, on éteint le maximum de précision pour un haut niveau de résolution. Malgré cet aspect, on cherche à maximiser la précision et on prend le maximum de niveaux de résolution pour un niveau 2 de résolution.

Pour les images panchromatiques un des paramètres disparaît, puisque l'ACP n'est plus nécessaire. Il reste, seulement, optimiser le niveau de résolution et la variance du noyau gaussien.

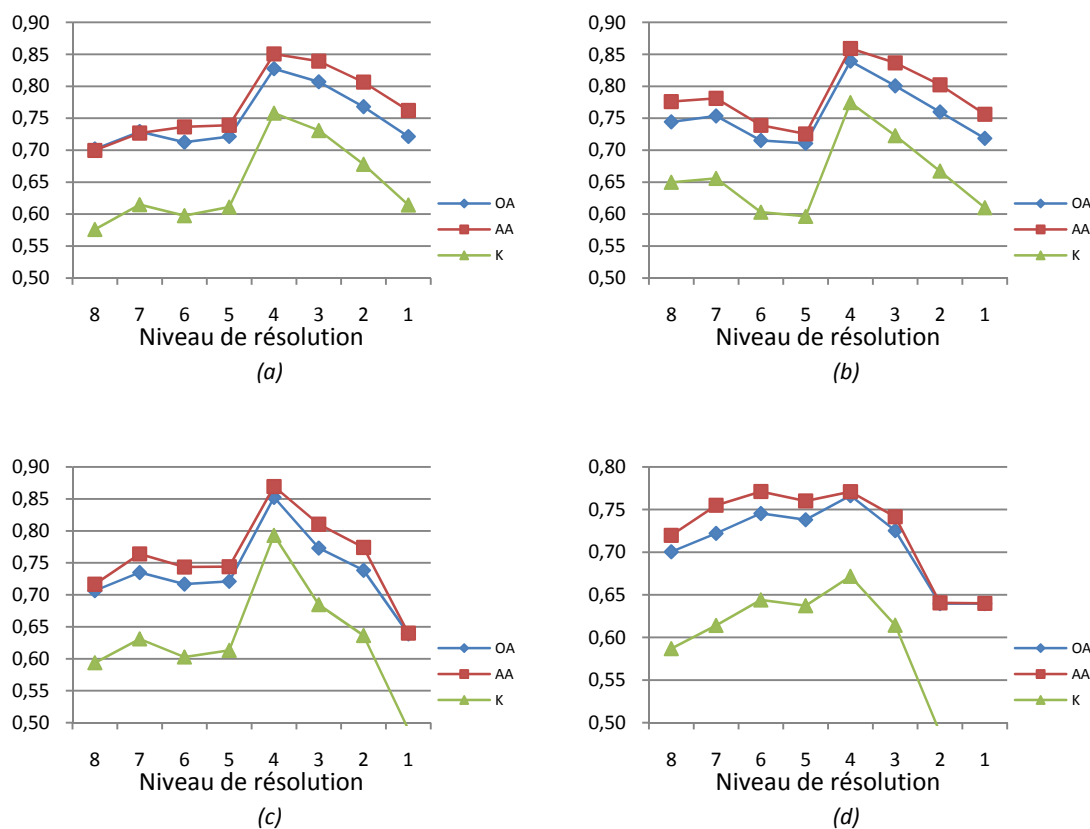


Figure 20 : Réalisations des coefficients de précision selon plusieurs niveaux de résolution et pour $\sigma = 0,5$ (a), $\sigma = 1$ (b), $\sigma = 2$ (c), $\sigma = 4$ (d).

Pour cette image, le niveau de résolution optimal est 4, indépendamment de la valeur de σ . Cependant on a un maximum de précision pour $\sigma = 2$ (Figure 20). Pour cette configuration de paramètres, les valeurs de précision sont les suivantes :

OA	AA	K
85,24	86,95	79,31

Table 4 : Paramètres de précision résultant de la simulation optimale

8.2. RÉSULTATS EXPÉRIMENTAUX

Dans cette section, on présente les résultats de classification de deux images : une image hyperspectrale de l'Université de Pavie, en Italie et une image panchromatique à haute résolution spatiale de la ville de Toulouse. Les paramètres pour la classification ont été réglés à partir des résultats de la section précédente.

L'image hyperspectrale est la même utilisée dans le chapitre précédent, mais découpée en une image carrée et dyadique. L'image originale a une taille de 610 x 340 pixels, l'image carrée de dimension dyadique maximale a une taille de 256 x 256 pixels. La zone de l'image a été choisie de façon à avoir au moins un échantillon appartenant à chaque classe. Une représentation de l'image découpée est donnée sur la Figure 21.



Figure 21 : Image ROSIS de la zone de l'Université de Pavie, en Italie, composée de 3 bandes de taille 256 x 256

De la même façon, les ensembles d'entraînement et de test ont été coupés (Figure 22).

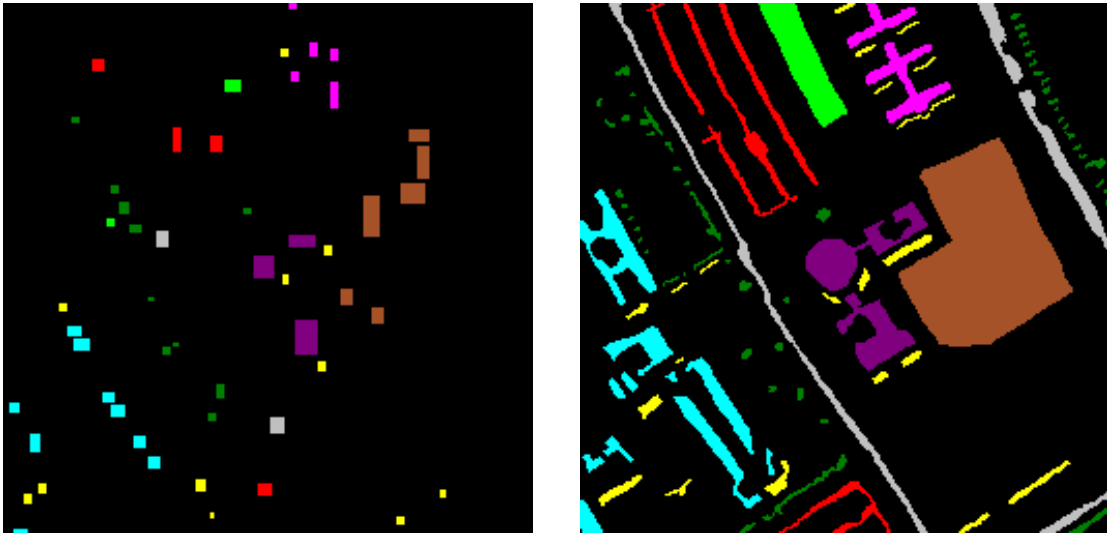


Figure 22 : Ensembles d'entraînement et de test. Les couleurs indiquent les différentes classes : *asphalte, prairie, gravier, arbre, plaque métallique, terrain dépouillé, bitume, brique* et *ombre*

Sur la Figure 23 on peut observer le résultat de classification de l'ensemble d'entraînement et de l'image complète. Si l'on compare ce résultat avec la classification spectrale, on observe que la classification est plus homogène même si elle est moins précise. Le bruit a été supprimé dans certaines zones, alors que dans d'autres zones le bruit est de taille plus grande.

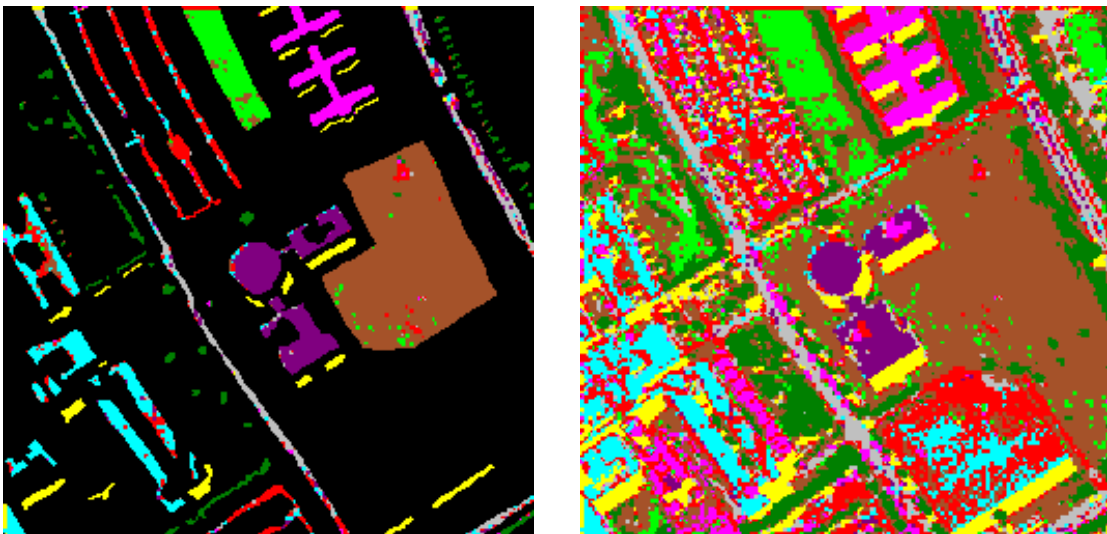


Figure 23 : Résultats de classification sur l'ensemble d'entraînement et sur l'image complète

Pour la classification de l'image panchromatique, une image de la ville de Toulouse du capteur PLEIADES de taille 602 x 540 pixels, on a dû découper une image carrée de taille 512 x 512. Sur la Figure 24 on a une représentation de l'image panchromatique à résolution spatiale 0,75 mètres par pixel.



Figure 24 : Image PLEIADES de la ville de Toulouse. Résolution spatiale : 0,75 m/pixels

Sur la Figure 25 on peut observer les ensembles d'entraînement et de test. L'ensemble d'entraînement est très réduit pour cette image, ce qui rend la classification avec les SVM plus difficile. Le nombre de classes, pour cette image est de 4.

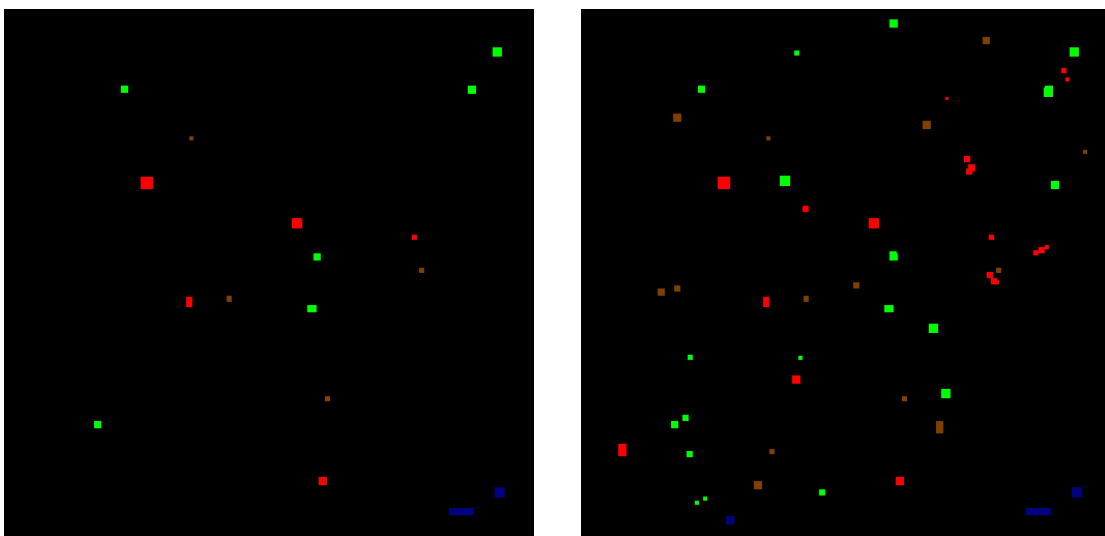


Figure 25 : Ensembles d'entraînement et de test. Les couleurs indiquent les différentes classes : **bâtiments**, **rue**, **terrain dégagé** et **ombre**

Le résultat de classification est présenté à continuation (Figure 26) pour les ensembles d'entraînement et l'image complète. À différence de l'image précédente, pour l'image de Toulouse on a moins d'échantillons d'entraînement, ce qui nous rend difficile évaluer l'homogénéité de la classification. Par contre, sur l'image complète on voit des erreurs de classification notables : la classification de l'intérieur des groupements de maisons comme rue, discontinuité et mauvaise classification des rues, surtout les plus petites.

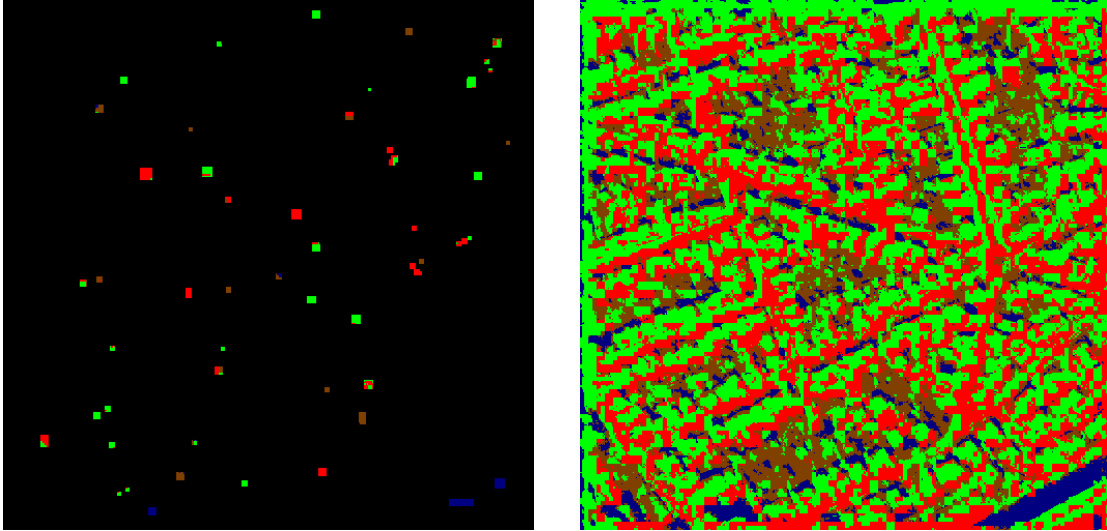


Figure 26 : Résultats de classification sur l'ensemble d'entraînement et sur l'image complète

Dans le chapitre suivant on va combiner les informations spatiales et spectrales pour obtenir une classification plus exacte et précise. En combinant les deux informations, on va essayer de corriger le bruit présent dans les classifications spectrale et spatiale séparées.

CLASSIFICATION CONJOINTE SPECTRALE - SPATIALE

Dans ce chapitre on présente la combinaison des noyaux spectral et spatial étudiés jusqu'à présent de façon séparée pour faire une classification conjointe en tenant compte des caractéristiques spatiales et spectrales des images hyperspectrales.

Dans une première partie de ce chapitre on présente différents types de noyaux conjoints avec ses expressions mathématiques ainsi que les résultats de classification en termes de paramètres de précision. À la fin de ce chapitre les résultats graphiques sont présentés pour la classification d'une image hyperspectrale et une image panchromatique, et ils sont comparés aux résultats de classification tenant compte de l'information seulement spatiale ou spectrale.

9.1. NOYAU SPECTRAL - SPATIAL

Dans le chapitre 7 on a vu la classification des images en tenant compte seulement de l'information spectrale. On a vu le noyau spectrale qui avait la forme

$$K_{\sigma_{spect}}^{spect}: \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0,1] \quad (49)$$

$$(x, y) \rightarrow \exp\left(-\frac{\|x-y\|^2}{2\sigma_{spect}^2}\right)$$

Au même temps, on a défini un autre noyau avec l'information spatiale (cf. chapitre 8), selon la forme

$$K_{\sigma_{spat}}^{spat}: \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0,1] \quad (50)$$

$$(x, y) \rightarrow \exp\left(-\frac{\|\Lambda_x - \Lambda_y\|^2}{2\sigma_{spat}^2}\right)$$

Une première approximation pour obtenir un noyau combinaison de l'information spectrale et spatiale consiste en la concaténation des noyaux spectral et spatial.

Soit l'image x considérée comme une matrice tridimensionnelle où les deux premières dimensions correspondent à la taille de l'image et la troisième correspond aux composantes spectrales et soit la matrice tridimensionnelle Λ_x où les deux premières dimensions correspondent, aussi, à la taille de l'image et la troisième correspond aux coefficients dérivés de l'analyse multirésolution, on peut définir une troisième matrice Φ_x tridimensionnelle en concaténant ces deux premières matrices sur la dimension 3 qui contient pour un pixel donné

$$\Phi_x = [x^1, x^2, \dots, x^n, \quad a_j, a_{j+1}, d_{j+1}^1, d_{j+1}^2, a_{j+2}, d_{j+2}^3, d_{j+1}^1, d_{j+2}^2, d_{j+2}^3, \dots] \quad (51)$$

où n est la résolution spectrale de l'image, x^i les composantes spectrales et a_l et d_l^k les coefficients multirésolutions associés à ce pixel.

On définit le noyau spatio-spectral comme l'application

$$K_\sigma: \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0,1]$$

$$(x, y) \rightarrow \exp\left(-\frac{\|\Phi_x - \Phi_y\|^2}{2\sigma^2}\right) \quad (52)$$

où σ est le paramètre du noyau.

Une autre approche consiste à créer un nouveau noyau. Des règles pour la construction des noyaux peuvent se trouver dans l'Annexe A de [5]. On a utilisé la propriété de linéarité pour construire le nouveau noyau :

si K_1 et K_2 sont deux noyaux, et $\mu_1, \mu_2 \geq 0$, alors $\mu_1 K_1 + \mu_2 K_2$ est un noyau

À partir de cette propriété, on définit le noyau spatio-spectral K selon

$$K_\mu: \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0,1]$$

$$(x, y) \rightarrow \mu K_{\sigma_{spect}}^{spect}(x, y) + (1 - \mu) K_{\sigma_{spat}}^{spat}(x, y) \quad (53)$$

où les paramètres du noyau sont μ , σ_{spect} et σ_{spat} .

On cherche à optimiser ces 3 paramètres pour l'image hyperspectrale de l'Université de Pavie, en Italie. Pour les valeurs de variance σ_{spect} et σ_{spat} on choisit les valeurs déjà calculés dans les chapitres 7 et 8 lors de la classification spectrale et spatiale ($\sigma_{spect} = 1$ et $\sigma_{spat} = 0,5$). Pour le valeur de μ on fait plusieurs simulations de 0 à 1 par intervalles de 0,1. Les résultats sont présentés sous forme graphique sur la Figure 27.

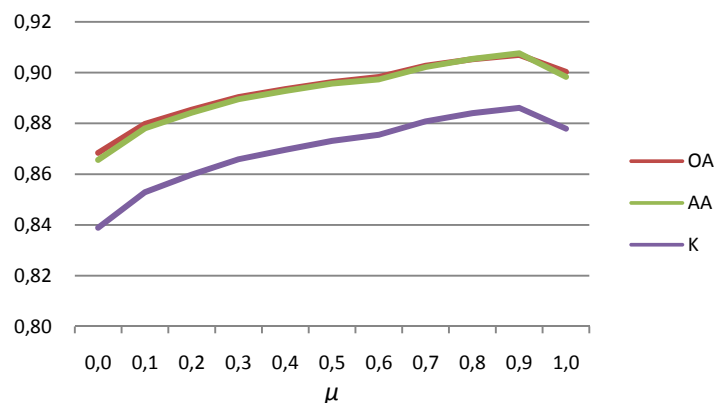


Figure 27 : Paramètres de précision en fonction du valeur de μ pour l'image Université

Le maximum est obtenu pour $\mu = 0,9$ et les valeurs des paramètres de précision sont présentés sur la Table 5.

La valeur de μ obtenu est très proche à 1, c'est-à-dire qu'on va donner plus d'importance au noyau spectral qu'au noyau spatial. Cela est normal du point de vue qu'il s'agit d'une image hyperspectral dont la plupart de l'information on la trouve dans les différentes bandes, alors que sa résolution spatiale, étant de 1,5 m/pixels ne nous fournit autant d'information sur les objets présents dans la scène. Cependant, la valeur de μ égal à 1, ne donne pas les meilleurs résultats. Ceci justifiant l'information spatiale additionnelle.

OA	AA	K
90,70	90,76	88,61

Table 5 : Paramètres de précision résultant de la simulation avec $\mu = 0,9$

Pour l'image panchromatique de la ville de Toulouse on doit optimiser deux paramètres : μ et σ_{spect} , puisque dans le chapitre 7 on n'a pas fait la classification spectrale de cette image. Le paramètre σ_{spat} est figé à 2. Après plusieurs essais, la valeur optimale pour σ_{spect} est de 0,5. Pour le choix de μ on fait plusieurs essais entre 0 et 1 en intervalles de 0,1 (Figure 28).

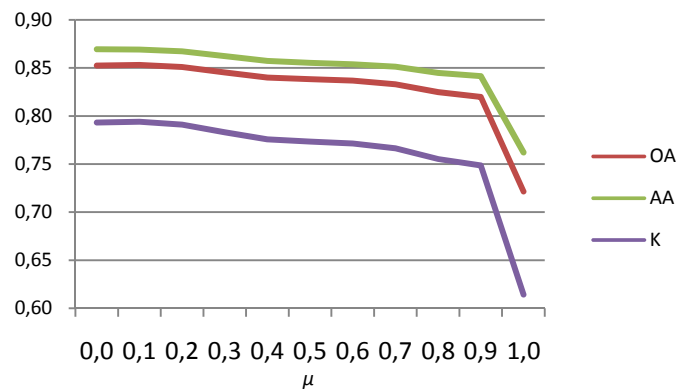


Figure 28 : Paramètres de précision en fonction du valeur de μ pour l'image Toulouse

La valeur de μ pour laquelle on obtient une précision plus haute est de 0,1. Pour cette image on donne plus d'importance au noyau spatial (valeur de μ proche à 0) qu'au noyau spectral dû au fait que l'image est panchromatique (contient une seule composante spectrale) et sa résolution spatiale est de 0,75 m/pixels, plus grande que pour l'image précédente. Sur la Table 6 on a représenté les valeurs des paramètres de précision éteints pour ce valeur de μ .

OA	AA	K
85,31	86,92	79,40

Table 6 : Paramètres de précision résultant de la simulation optimale avec $\mu = 0,1$

9.2. RÉSULTATS EXPÉRIMENTAUX

Après avoir choisi les paramètres qui donnent les meilleurs résultats de classification, les résultats expérimentaux sont présentés sous forme graphique. Dans cette section on va analyser les deux images utilisées pour la classification et on va les comparer avec les classifications purement spectrales et spatiales.

La première image à classer est l'image hyperspectrale du capteur ROSIS de l'Université de Pavie, en Italie. Comme dans le chapitre précédent, l'image originale a été coupée en une image carrée dyadique à fin de pouvoir utiliser le noyau spatial. La représentation de l'image composée de 3 bandes, ainsi que les ensembles d'entraînement et de test on peut les voir respectivement sur les Figures 21 et 22.

Le résultat de classification de l'ensemble de test et de l'image complète en utilisant le noyau combiné est représenté sur la Figure 29. On a inclus, aussi les résultats de classification pour le noyau spectral et spatial à fin de pouvoir comparer qualitativement les résultats de classification.

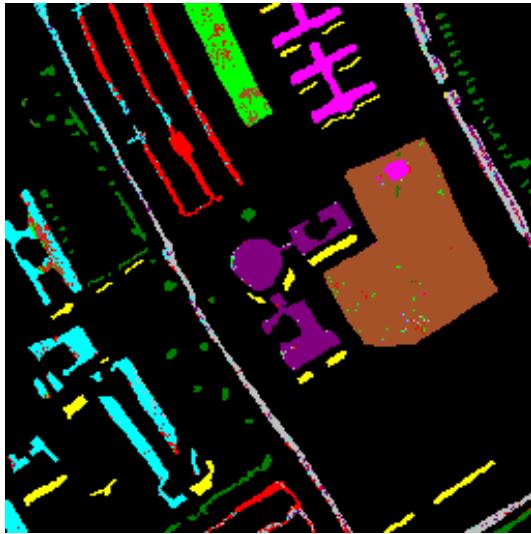
Par rapport à l'homogénéité de la classification, on remarque qu'avec l'utilisation du noyau combiné on réduit du bruit qui apparaissait lors de la classification spectrale grâce à l'introduction de l'information spatiale. Par contre, il y a des zones assez difficiles à classer pour les deux classifieurs spectral et spatial, ce qui se traduit par une mauvaise classification aussi dans le cas combiné. Pour ce qui concerne à l'aspect de l'image résultat de la classification, on observe qu'il est très proche du résultat spectral, dû au fait de la valeur du paramètre μ , qui donne un poids plus important au noyau spectral qu'au noyau spatial.

La deuxième image classifiée est l'image panchromatique de la ville de Toulouse. Une représentation de l'image originale coupée est fournie sur la Figure 24. Les ensembles d'entraînement et de test sont représentés sur la Figure 24 du chapitre précédent.

Sur la Figure 30 on a l'ensemble des résultats de classification sur l'image complète, ainsi que pour l'ensemble de test pour les trois classifications : spectrale, spatiale et combinée.

Il faut remarquer que l'image résultat de la classification avec le noyau combiné est très proche de la classification avec le noyau spatial. Cela est dû à la valeur de μ , proche de 0. Il est difficile de distinguer les différences entre les deux images et c'est pour ça qu'on se base sur les paramètres de précision pour décider quelle classification est plus exacte.

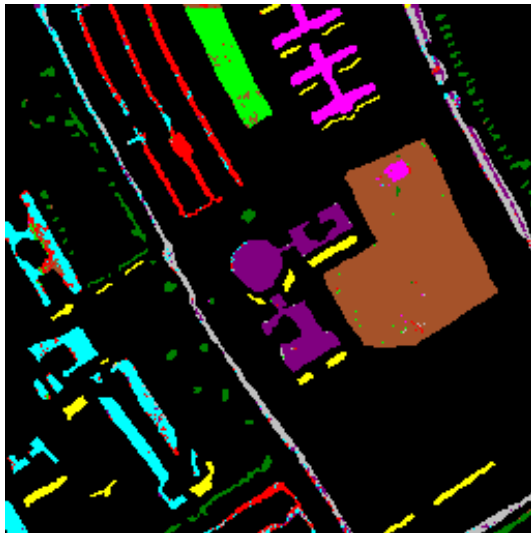
Le fait d'avoir seulement 4 classes qui décrivent toute l'image rend le résultat moins réel et intuitif que pour l'image de l'Université, où on a 9 classes.



(a) Noyau spectral



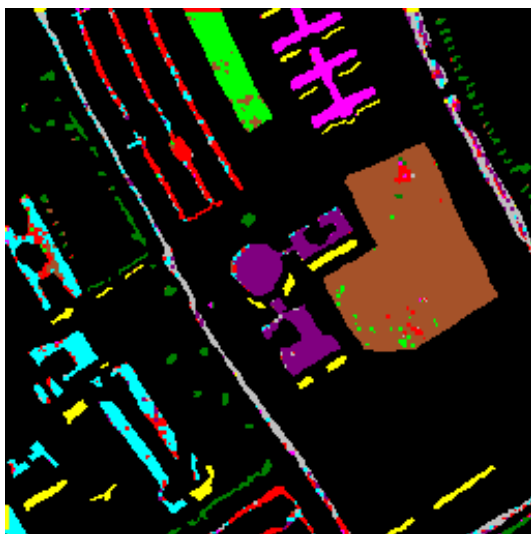
(b) Noyau spectral



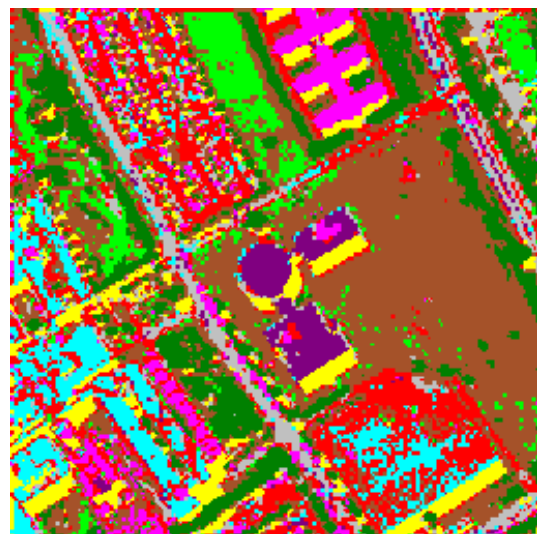
(c) Noyau spatio-spectral



(d) Noyau spatio-spectral

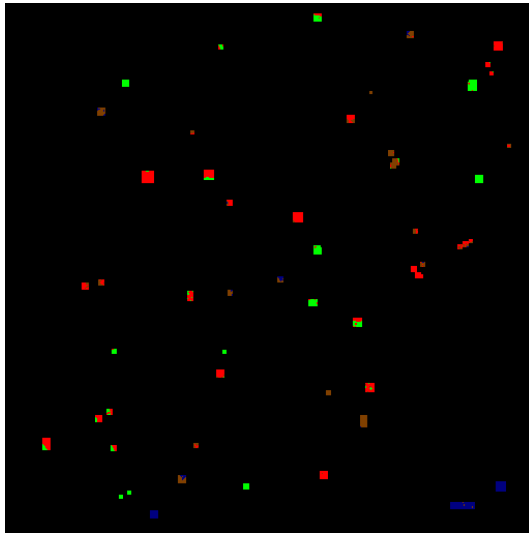


(e) Noyau spatial

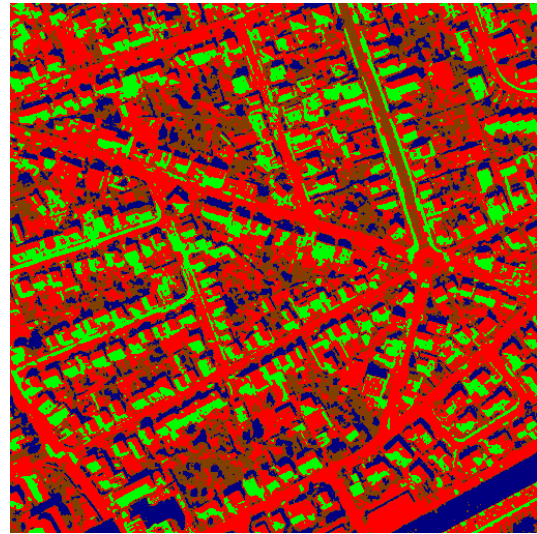


(f) Noyau spatial

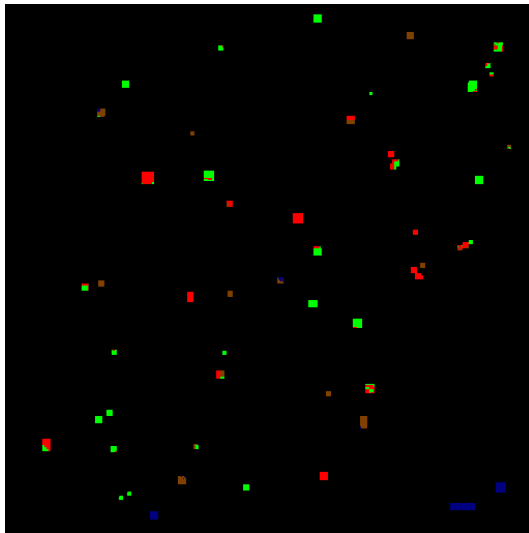
Figure 29 : Classification pour l'ensemble d'entraînement et l'image complète d'Université



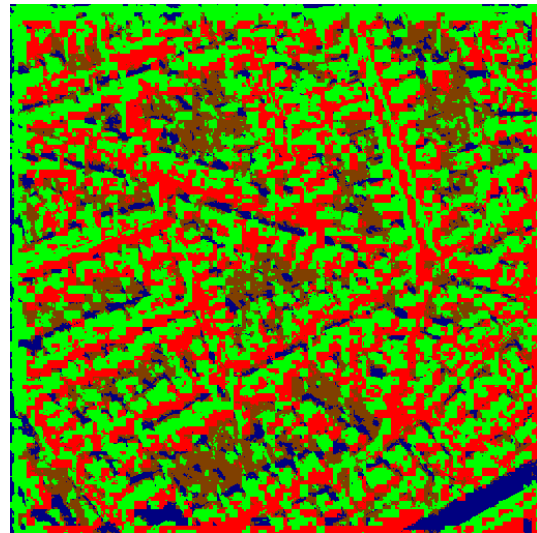
(a) Noyau spectral



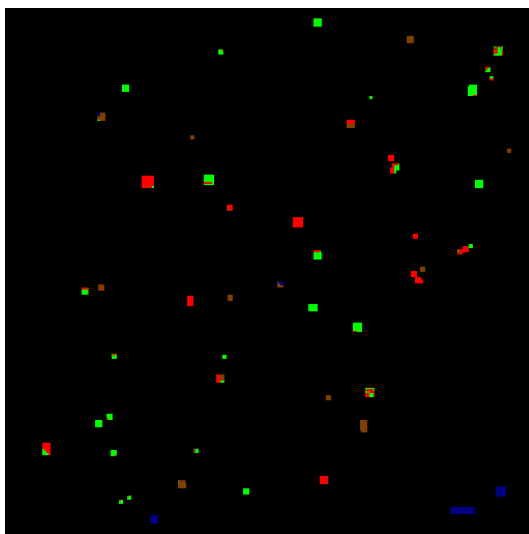
(b) Noyau spectral



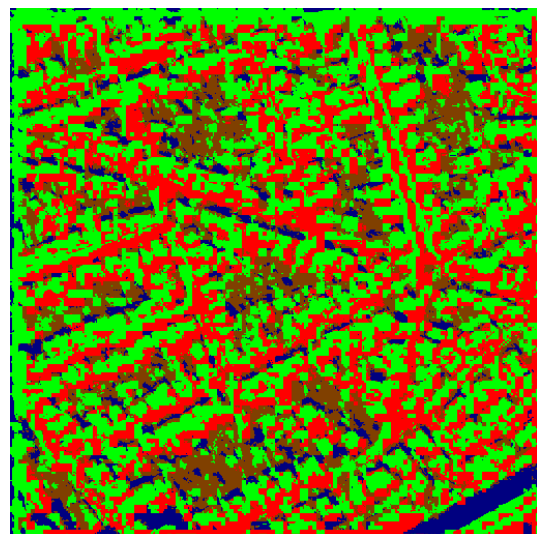
(c) Noyau spatio-spectral



(d) Noyau spatio-spectral



(e) Noyau spatial



(f) Noyau spatial

Figure 30 : Classification pour l'ensemble d'entraînement et l'image complète de Toulouse

CHAPITRE 10

PLAN DE TRAVAIL (DIAGRAMME DE GANTT)

Dans ce chapitre une description du déroulement du stage est présentée sous forme de diagramme de Gantt. Le diagramme est représentée à la fin du chapitre et les données correspondantes sont affichées sur la Table 7. À la suite, une description des tâches présentes au diagramme est fournie, ainsi qu'une comparaison avec le plan de travail prévisionnel.

Liste des tâches	Date de début	Durée (jours)	Date de la fin
Lecture de la bibliographie (1ère partie)	04/02	30	05/03
Résumé et rédaction de la partie théorique du rapport	25/02	14	10/03
Familiarisation avec le code source (SVM)	25/02	07	03/03
Adaptation du code à la nouvelle librairie	03/03	07	10/03
Redaction du pré rapport	07/03	06	13/03
Test et évaluation des résultats	10/03	03	13/03
Validation et choix de paramètres de test	14/03	17	31/03
Préparation de la présentation	20/03	07	27/03
Lecture de la bibliographie (2ème partie)	31/03	14	15/05
Familiarisation avec le code source (ondelettes)	28/04	07	05/05
Construction noyaux avec info spatiale	05/05	10	15/05
Implémentation et test du noyau spatial	12/05	25	06/06
Validation et choix de paramètres de test	19/05	18	06/06
Construction noyaux avec info spatiale et spectrale	15/05	10	25/05
Implémentation et test du noyau combiné	20/05	17	06/06
Validation et choix de paramètres de test	22/05	15	06/06
Redaction du rapport	26/05	15	10/06

Table 7: Énumération des tâche développées au long du stage

Description détaillée des tâches présentes sur le diagramme :

Lecture de la bibliographie (1^{ère} partie) : Lecture de la bibliographie technique concernant les machines à support vecteur et les noyaux utilisés.

Résumé et rédaction de la partie théorique du rapport : Résumé des parties importantes de la bibliographie qui a été incluse dans ce rapport au chapitres 4.

Familiarisation avec le code source (SVM): Lecture et compréhension du code Matlab pour la classification des images hyperspectrales développée dans [5]. Repaire des images utilisées et compréhension de sa codification.

Adaptation du code à la nouvelle librairie : Les codes fournis sont basés sur une toolbox déterminée, mais pendant le stage on va utiliser une autre toolbox développée pour Matlab qu'on peut trouver sur [13]. Procès d'adaptation et compréhension des nouvelles fonctions de cette librairie.

Rédaction du pré-rapport

Test et évaluation des résultats : Classification spectrale de deux images hyperspectrales réelles et obtention des paramètres qui caractérisent l'exactitude de cette classification.

Validation et choix des paramètres de test : Après le test antérieur, réglage des paramètres du classifieur pour améliorer la précision de classification.

Préparation de la présentation

Lecture de la bibliographie (2^{ème} partie) : Lecture de la deuxième partie la bibliographie concernant l'analyse multirésolution des images. Concrètement, la lecture s'est centré en la transformée en ondelettes. La lecture est prévue pendant les 2 mois suivants.

Familiarisation avec le code source (ondelettes) : Après l'étude théorique des ondelettes, lecture et compréhension du code Matlab pour le traitement des images par ondelettes. Des tests ont été effectués pour se familiariser avec la nouvelle toolbox.

Construction des noyaux avec l'information spatiale : Application des aspects théoriques vus dans la littérature pour construire le noyau spatial : d'abord, sous une expression mathématique, puis sous une fonction Matlab.

Implémentation et test des noyaux spatiales : Une fois crée, test du noyau sur des images réelles : une image hyperspectrale et une image panchromatique.

Validation et choix des paramètres de test : Après le test du nouveau noyau, réglage des paramètres du classifieur pour améliorer les performances.

Construction des noyaux combinaison de l'information spatiale et spectrale : Recherche d'un noyau résultat de la combinaison de l'information spatiale et spectrale : d'abord, sous une notation mathématique, puis sous une fonction Matlab.

Implémentation et test du noyaux combiné : Test tu noyau sur les mêmes images réelles que pour le test spatial.

Validation et choix des paramètres de test : Une fois testé, réglage des paramètres du noyau pour éteindre une classification optimale.

Rédaction du rapport

Si l'on compare le plan de travail effectué avec celui qui avait été prévu précédemment, on trouve des différences très importants. D'abord, la différence la plus importante a été un changement sur l'approche de comment extraire l'information spatiale des images. Au début, on avait prévu d'étudier le processus de formation de l'image en modélisant le capteur, les perturbations qui rendent floue l'image, etc. ; alors que finalement on s'est penché vers un analyse multirésolution pour étudier comment les objets de la scène sont représentés sur plusieurs résolutions.

Une autre différence qu'on peut trouver dans le déroulement du stage est la durée totale du même, initialement prévue jusqu'à fin juin, alors qu'elle a été jusqu'à fin mai. Cette réduction du temps dédié au projet a provoqué un décalage et une réduction des temps dédiés à l'amélioration des noyaux et à la recherche des paramètres qui optimisaient la classification.

Finalement, la recherche d'un noyau qui contenait l'information spatiale s'est maintenu. Pourtant, on a dédoublé le temps dédié à la classification spatiale, pour inclure, aussi la classification conjointe avec l'information spectrale et spatiale. Comme on peut observer sur le diagramme de Gantt, après la lecture bibliographique correspondant à la transformée en ondelettes, on a commencé à aborder le noyau spatial, mais avant de finir avec l'implémentation et amélioration de ce noyau, on a commencé à étudier le noyau spatio-spectral pour finir avec une validation des paramètres conjointe entre les noyaux spatial et spatio-spectral.

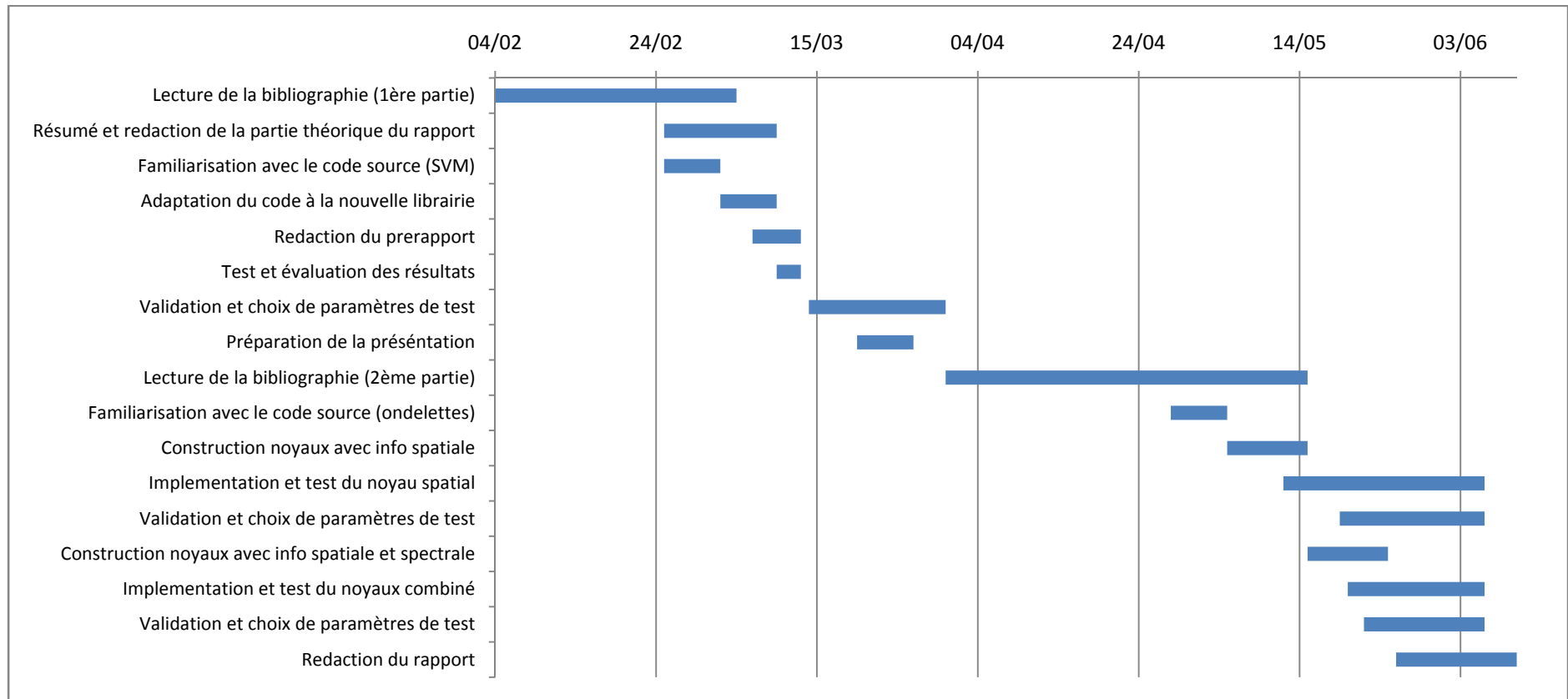


Figure 31 : Diagramme de Gantt du déroulement du stage

CHAPITRE 11 CONCLUSION

La classification d'images hyperspectrales pour caractériser le milieu urbain a été abordée dans ce rapport. L'objectif était de déterminer à quelle classe appartient chaque pixel de l'image avec la précision la plus grande possible. Pour cela, les Machines à Vecteurs Support (SVM), classifieurs basés sur la théorie de l'apprentissage supervisé, ont été étudiées pour la classification automatique des images à haute résolution spectrale. Dans ce rapport, on a étudié trois approches pour la classification des images : l'approche spectrale, l'approche spatiale et l'approche combinée.

Approche spectrale

Dans une première partie de ce stage, après une étude théorique des SVM, on a abordé la classification tenant compte l'information purement spectral. Des résultats expérimentaux ont été présentés pour illustrer les explications théoriques.

Comme on a vu, les SVM se servent des fonctions noyau pour classifier les images. Après avoir étudié plusieurs fonctions noyau, le noyau gaussien a été choisi pour effectuer les tests grâce à ces performances face à d'autres noyaux linéaires et non linéaires. Pour voir comment les paramètres influent sur le résultat de classification on a fait plusieurs essais pour différents valeurs du paramètre du noyau à fin de trouver la valeur optimale qui nous rend la classification la plus précise possible.

Les résultats de classification pour l'approche spectral sont assez précis, mais on trouve dans les images des zones mal classifiées et d'autres où la classification n'est pas homogène (présence du bruit). Cela nous amène à introduire l'information spatiale pour améliorer le résultat de classification

Approche spatiale

L'étude des caractéristiques spatiales des images hyperspectrales est motivée par la haute résolution spatiale des images dont on dispose. Pour extraire l'information spatiale on a fait une analyse multirésolution, basé sur la décomposition en ondelettes de l'image. L'image utilisée lors de l'analyse multirésolution a été projetée sur un des axes principaux pour réduire l'ensemble des variables. Après un étude théorique de la transformée en ondelettes, on a proposé un vecteur de caractéristiques spatiales contenant les coefficients de la décomposition en ondelettes pour chaque résolution et pour chaque pixel de l'image.

Les résultats expérimentaux ont été présentés pour une classification basée sur le noyau gaussien, après avoir optimisé les différents paramètres du noyau comme le niveau de résolution, le nombre de composantes principales sur lesquelles on a projeté l'image ou la variance du noyau gaussien.

Pour l'approche spatiale on a observé une meilleure homogénéité de la classification, mais une précision de classification moins performante que pour l'approche spectrale.

Finalement, on a combiné les deux approches dans un noyau spatio-spectrale et on a obtenu une classification plus précise des deux images. La précision de la classification a été évaluée du point de vue subjectif en regardant les résultats graphiques de classification et objectif, à partir des paramètres de précision définis.

Comme conclusion finale de ce stage, la classification des objets dans une image hyperspectrale a été éteinte en combinant l'information spectrale et spatiale avec un niveau de précision acceptable. Pour une poursuite de ce travail et à fin d'améliorer les performances de classification, un étude d'autres noyaux pourrait être envisagé, ainsi que l'étude d'autres filtres lors de la décomposition d'ondelettes qui tiennent compte d'autres directions sur l'image.

CHAPITRE 12

ÉVALUATION DU COÛT DU PROJET

Pour évaluer le coût de ce projet, on s'est basé sur la Fiche d'Évaluation du Coût d'un Contrat, document pour le calcul approximatif du coût d'un projet dans les laboratoires de l'INPG. Ce document est fourni avec le projet et contient le coût de base de personnel permanent, ainsi que les dépenses spécifiques au contrat.

Le coût de base de personnel tient compte du type de chercheur qui va participer au projet, étant le coût le plus élevé pour un directeur de recherche, et le coût le plus bas pour un chercheur débutant, et il est exprimé en € par an.

Les dépenses spécifiques au contrat comprennent l'achat de matériel, la sous-traitance, les heures de calcul et d'autres dépenses.

Pour la réalisation du stage on a compté sur la participation d'un chercheur confirmé, un chercheur post-doc et un chercheur débutant. La dévouement, ainsi que le coût proportionnel sont détaillés sur la Table 8.

Cadre	Temps dédié	Coût
Chercheur débutant	5 mois	40.652 €
Chercheur post-doc	1 demi-journée / semaine	5.549 €
Chercheur confirmé	3 jours	1.639 €
	Totale	47.840 €

Table 8 : Coût de base du personnel proportionnel au temps dédié

Pour la réalisation de ce projet, on ne compte pas sur des dépenses spécifiques, puisqu'il a été réalisé sur des équipes informatiques déjà installés et appartenant au laboratoire. L'achat de matériel n'a pas été nécessaire, ainsi qu'il n'y a pas eu besoin de sous-traitance.

Alors, le coût approximatif du projet est d'environ 50.000 €.

LÉGENDE DES FIGURES

<i>Figure 1 : Facteur de réflexion de deux pixels différents selon la longueur d'onde</i>	4
<i>Figure 2 : Représentation d'une image hyperspectral sous forme d'hypercube</i>	5
<i>Figure 3: Une image IKONOS.....</i>	13
<i>Figure 4: Ensembles Hyperplan optimal de séparation dans un SVM pour le cas linéairement non-séparables</i>	16
<i>Figure 5 : Schéma de la transformée en ondelettes et de la transformée inverse</i>	23
<i>Figure 6 : Schéma multirésolution avec les coefficients d'ondelettes pour 2 résolutions différentes et le vecteur de caractéristiques</i>	25
<i>Figure 7 : Matrice de covariance pour l'image hyperspectrale correspondant à l'Université de Pavie (Italie)</i>	28
<i>Figure 8 : Contribution de chaque valeur propre pondérée par la somme totale des 103 valeurs propres</i>	29
<i>Figure 9 : Projection sur la 1ère et 3ème composante principale</i>	29
<i>Figure 10 : Structure multirésolution avec coefficients de la décomposition.....</i>	31
<i>Figure 11 : Transformée en ondelettes à l'échelle 2 avec zoom sur l'effet de bords</i>	31
<i>Figure 12 : Transformée en ondelettes à l'échelle 3 avec zoom sur l'effet de bords</i>	32
<i>Figure 13 : Image ROSIS de la zone de l'Université de Pavie, en Italie, composé de 3 bandes</i>	34
<i>Figure 14 : Ensembles d'entraînement et de test pour l'Université de Pavie.....</i>	35
<i>Figure 15 : Résultats de classification sur l'ensemble d'entraînement et sur l'image complète de l'Université de Pavie</i>	36
<i>Figure 16 : Image ROSIS du centre de la ville de Pavie, en Italie, composée de 3 bandes</i>	37
<i>Figure 17 : Ensembles d'entraînement et de test pour le Centre de Pavie.....</i>	38
<i>Figure 18 : Résultats de classification sur l'ensemble d'entraînement et sur l'image complète du Centre de Pavie</i>	38
<i>Figure 19 : Paramètres de précision en fonction des paramètres du noyau</i>	40
<i>Figure 20 : Réalisations des coefficients de précision selon plusieurs niveaux de résolution et σ</i>	41

<i>Figure 21 : Image ROSIS de la zone de l'Université de Pavie, en Italie, composée de 3 bandes de taille 256 x 256.....</i>	<i>42</i>
<i>Figure 22 : Ensembles d'entraînement et de test de taille 256 x 256</i>	<i>42</i>
<i>Figure 23 : Résultats de classification sur l'ensemble d'entraînement et sur l'image complète de l'Université de Pavie de taille 256 x 256</i>	<i>43</i>
<i>Figure 24 : Image PLEIADES de la ville de Toulouse. Résolution spatiale : 0,75 m/pixels</i>	<i>43</i>
<i>Figure 25 : Ensembles d'entraînement et de test de Toulouse.....</i>	<i>44</i>
<i>Figure 26 : Résultats de classification sur l'ensemble d'entraînement et sur l'image complète de Toulouse.....</i>	<i>44</i>
<i>Figure 27 : Paramètres de précision en fonction du valeur de μ pour l'image Université</i>	<i>46</i>
<i>Figure 28 : Paramètres de précision en fonction du valeur de μ pour l'image Toulouse</i>	<i>47</i>
<i>Figure 29 : Classification pour l'ensemble d'entraînement et l'image complète d'Université avec les 3 noyaux</i>	<i>49</i>
<i>Figure 30 : Classification pour l'ensemble d'entraînement et l'image complète de Toulouse avec les 3 noyaux</i>	<i>50</i>
<i>Figure 31 : Diagramme de Gantt du déroulement du stage</i>	<i>55</i>

GLOSSAIRE

AA	Average Accuracy
ACP	Analyse en Composantes Principales
CA	Classification Accuracy
ECG	Electrocardiogramme
EMG	Electromyogramme
IFS	Iteration Function System
OA	Overall Accuracy
RBF	Radial Basis Functions
RGB	Red Green Bleu (Rouge Vert Bleu, RVB)
ROSI	Reflective Optics System Imaging Spectrometer
SVM	Support Vector Machine
UBF	Ultra Basse Fréquence

RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] **J. A. Benediktsson and I. Kanellopoulos**
Classification of multisource and hyperspectral data based on decision fusion. IEEE Transactions on Geoscience and Remote Sensing, vol. 37, pp. 1367-1377, Mai 1999.
- [2] **I. Bloch et al.**
Le Traitement des Images. Polycopié du cours ANIM. Département TSI – Télécom Paris. Paris, 2004.
- [3] **J. Channusot**
Approches vectorielles ou marginales pour le traitement d'images multi-composantes. PhD thèse, Université de Savoie, France, 1998.
- [4] **O. Chapelle, V. N. Vapnik, O. Bousquet, and S. Mukherjee**
Choosing multiple parameters for support vectors machines. Machine Learning, vol. 46, n° 1, 2002.
- [5] **M. Fauvel**
Spectral and spatial methods for the classification of urban remote sensing data. PhD thèse, Institut National Polytechnique de Grenoble, France, 2007.
- [6] **M. Fauvel, J. Chanussot, and J. A. Benediktsson**
A combined support vector machines classification based on decision fusion. IEEE International Geoscience and Remote Sensing Symposium. IGARSS'06. Juillet 2006.
- [7] **M. Fauvel, J. Chanussot, and J. A. Benediktsson**
A joint spatial and spectral SVM's classification of panchromatic images. IEEE International Geoscience and Remote Sensing Symposium. IGARSS'07. Juillet 2007.
- [8] **M. Fauvel, J. Chanussot, and J. A. Benediktsson**
Evaluation of kernels for multiclass classification of hyperspectral remote sensing data. IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP'06. Mai 2006.
- [9] **A. Jalobeanu**
Modèles, estimation bayésienne et algorithmes pour la déconvolution d'images satellitaires et aériennes. PhD thèse, Université de Nice-Sophia Antipolis, France, 2001.
- [10] **S. Mallat**
Une exploration des signaux en ondelettes. Éditions de l'École Polytechnique. Palaiseau, 2000.

[11] **F. Melgani and L. Bruzzone**

Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on Geoscience and Remote Sensing, vol. 42, n° 8, pp. 1778-1790, Août 2004.

[12] **J. Shawe-Taylor and N. Cristianini**

Kernel methods for pattern analysis. Cambridge University Press, 2004.

[13] <http://asi.insa-rouen.fr/enseignants/~arakotom/toolbox/index.htm> . Librairie Matlab utilisée pour la classification avec SVM. Février 2008.

BIBLIOGRAPHIE COMPLÉMENTAIRE

J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson

Classification of hyperspectral data from urban areas based on extended morphological profiles. IEEE Transactions on Geoscience and Remote Sensing, vol. 43, n° 3, pp. 480-491 Mars 2005.

O. Chapelle, P. Haffner, and V. N. Vapnik

Support vector machines of histogram-based image classification. IEEE Transactions on Neural Networks, vol. 10, n° 5, pp. 1055-1064, Septembre 1999.

I. Daubechies

Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics. Philadelphia, 1992. Eight printing.

L. Hong and D. Brzakovic

Bayesian Restoration of Image Sequences Using 3-D Markov Random Fields. IEEE International Conference on Acoustics, Speech, and Signal Processing. ICASSP-89. Mai, 1989.

M. Pesaresi and J. A. Benediktsson

A new approach for the morphological segmentation of high-resolution satellite imagery. IEEE Transactions on Geoscience and Remote Sensing, vol. 39, n° 2, pp. 309-320, Février 2001.

S. Sarkar and G. Healey

Hyperspectral Texture Classification Using Generalized Markov Fields. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'04, Juin 2004.

C. Thomas

Fusion d'Images de Résolutions Spatiales Différentes. PhD thèse, École de Mines de Paris, France, 2006.

PROJET numéro 87

DERNIERE PAGE DE COUVERTURE

ANNEE UNIVERSITAIRE **2007/2008**

PROJET DE TROISIEME ANNEE

RAPPORT FINAL

JUIN 2008

TITRE DU SUJET :

**Classification d'images hyperspectrales pour la caractérisation
du milieu urbain par une approche multirésolution**

RESUME (environ 10 lignes) :

Le stage a comme objectif la classification d'images hyperspectrales pour la caractérisation du milieu urbain en tenant compte de l'information spectrale et spatiale. Pour cette classification, les machines à vecteur support (SVM) ont été choisies pour sa précision de classification. Dans ce rapport, une étude des SVM est présentée ainsi que les fonctions noyau permettant de rendre la classification plus précise. Pour extraire l'information spatiale de l'image, une analyse multirésolution de l'image est proposé basé sur la transformée en ondelettes. Avec cette analyse, on obtient un vecteur de caractéristiques spatiales pour chaque pixel à fournir au classifieur SVM. Des résultats expérimentaux sont présentés pour trois approches de classification : classification spectrale, classification spatiale et classification conjointe spatio-spectrale, ainsi que les paramètres de précision associés à chaque classification. Finalement, un plan du travail effectué sous diagramme de Gantt est fourni.

Proposé par l'Entreprise :

Gipsa-lab - ENSIEG

MOTS-CLES :

**classification, images hyperspectrales, SVM,
ondelettes, multirésolution**