



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par **l'Institut Supérieur de l'Aéronautique et de l'Espace**
Spécialité : Photonique et systèmes optoélectroniques

Présentée et soutenue par **Guillaume ROUSSEL**
le **10 juillet 2012**

**Développement et évaluation de nouvelles méthodes
de classification spatiale-spectrale d'images hyperspectrales**

JURY

M. Nicolas Paparoditis, président
Mme Véronique Achard, co-directrice de thèse
M. Alexandre Alakian
M. Jocelyn Chanussot, rapporteur
M. Jean-Claude Fort, directeur de thèse
M. Grégoire Mercier, rapporteur

École doctorale : **Génie électrique, électronique, télécommunications :
du microsystème au système**

Unité de recherche : **Équipe d'accueil ISAE-ONERA OLIMPES**

Directeur de thèse : **M. Jean-Claude Fort**
Co-directrice de thèse : **Mme Véronique Achard**

Résumé

L'imagerie hyperspectrale, grâce à un nombre élevé de bandes spectrales très fines et contiguës, est capable d'associer à chaque pixel d'une image une signature spectrale caractéristique du comportement réfléchissant du matériau ou du mélange de matériaux présents dans ce pixel. La plupart des algorithmes de classification tirent profit de cette grande profusion d'information spectrale mais exploitent très peu l'information contextuelle existant entre les pixels appartenant à un même voisinage. L'objectif de cette thèse est de réaliser de nouveaux algorithmes utilisant simultanément les informations spectrale et spatiale à des fins de classification et d'étudier la complémentarité de ces deux types d'information dans divers contextes. Dans cette optique nous avons développé trois scénarios de classification sensiblement différents, chacun étant adapté à un type d'application particulier.

Nous avons tout d'abord développé un procédé d'extraction puis de classification vectorielle d'un ensemble de caractéristiques spectrales et spatiales. Les caractéristiques spectrales sont extraites au moyen de méthodes visant à réduire la dimension des images hyperspectrales tout en conservant une majorité de l'information utile. Les caractéristiques spatiales sont quant à elles produites par l'intermédiaire d'outils de caractérisation de la texture (matrices de co-occurrence et spectres de texture) ou de la forme (profils morphologiques). Nous nous sommes ensuite intéressés à la modélisation markovienne et avons entrepris d'adapter un algorithme de classification de type Conditional Random Field à un contexte hyperspectral. Notre troisième et dernière approche s'appuie sur une segmentation préalable de l'image afin de réaliser une classification par zones et non plus par pixels.

L'information spectrale pure permet de regrouper efficacement des pixels présentant des signatures spectrales similaires et suffit généralement dans le cadre de problèmes de classification ne faisant intervenir que des classes sémantiquement très précises, liées à un unique type de matériau. Les classes plus générales (utilisées par exemple pour des applications d'aménagement des sols) se composent en revanche de plusieurs matériaux parfois communs à plusieurs classes et agencés selon des motifs qui se répètent. Caractérisables à la fois spatialement et spectralement, ces classes sont susceptibles d'être plus complètement décrites par une utilisation simultanée de ces deux types d'information. Pour conclure cette

RÉSUMÉ

étude, nous avons effectué une comparaison des trois méthodes d'intégration de l'information spatiale au processus de classification selon les trois critères que sont la précision de classification, la complexité algorithmique et la robustesse.

Abstract

Thanks to a high number of thin and contiguous spectral bands, the hyperspectral imagery can associate to each pixel of an image a spectral signature representing the reflective behaviour of the materials composing the pixel. Most of the classification algorithms use this great amount of spectral information without noticing the contextual information between the pixels that belong to the same neighborhood. This study aims to realize new algorithms using simultaneously the spectral and spatial informations in order to classify hyperspectral images, and to study their complementarity in several contexts. For this purpose, we have developed three different classification scenarios, each one adapted to a particular type of application.

The first scenario consists in a vectorial classification processus. Several spectral and spatial characteristics are extracted and merged in order to form a unique data set, which is classified using a Support Vector Machine method or a Gaussian Mixing Model algorithm. The spectral characteristics are extracted using dimension reduction method, such as PCA or MNF, while the spatial characteristics are extracted using textural characterization tools (co-occurrence matrices and texture spectra) or morphological tools (morphological profiles). For the second scenario, we adapted a Conditional Random Field algorithm to the hyperspectral context. Finally, the last scenario is an area-wise classification algorithm relying on a textural segmentation method as a pre-processing step.

The spectral information is generally sufficient to deal with semantically simple classes, linked to a unique type of material. Complex classes (such as ground amenagment classes) are composed of several materials which potentially belong to more than one class. Those classes can be characterized both spectrally and spatially, which means that they can be more completely described using both spectral and spatial informations. To conclude this study, we compared the three spectral/spatial classification scenarios using three criterions : classification accuracy, algorithmic complexity and strength.

ABSTRACT

Remerciements

Plus de trois années de travail, tiraillées entre deux laboratoires et deux villes - Paris la lumineuse et Toulouse la rose - ça fait un beau paquet de gens à remercier. C'est ainsi tout autant par devoir moral que par désir de ne froisser aucune susceptibilité que je vais tâcher de n'oublier personne en chemin. Je vais tout d'abord remercier mes deux rapporteurs ainsi que le président de mon jury, Jocelyn Chanussot, Grégoire Mercier et Nicolas Paparoditis, d'une part pour l'intérêt qu'ils ont porté à mon travail et d'autre part pour avoir accepté de se déplacer, parfois d'assez loin, pour me voir le soutenir. Un grand merci également à mes encadrants Véronique et Alexandre, ainsi qu'à Jean-Claude, mon directeur de thèse, pour m'avoir supporté (dans les deux sens du terme) tout ce temps, avec une opiniâtreté que la distance n'est pas parvenue à émousser. Je me souviendrai tout particulièrement de l'humour pince-sans-rire accueillant les résultats (trop) souvent décevants de mes nombreuses études de sensibilité, lors de nos réunions hebdomadaires du vendredi ; un humour qui contribua grandement à la conservation de mon équilibre mental. Pierre Desproges disait qu'il faut rire de tout car l'humour désacralise la bêtise, il faut croire que le bougre fonctionne tout aussi bien avec la frustration et le découragement ! Avant de poursuivre la valse des remerciements, précisons que ce qui suit respecte l'ordre chronologique de mes allers et venues entre Toulouse et Paris. De plus, afin de ne pas alourdir le texte avec une surpopulation indigeste de "merci" en tout genre, veuillez considérer que ma gratitude envers les personnes citées par la suite est implicite.

Une thèse se limite-t-elle au seul travail scientifique décrit dans les quelques dizaines de pages de son manuscrit ? Au-delà de l'état de l'art et du développement, au-delà des expérimentations et autres études de sensibilité, n'y a-t-il pas une expérience humaine, à la fois épreuve et accomplissement, qui élargit l'esprit et le porte à maturité ? Alors que je pénétrais pour la première fois sur le site toulousain de l'Onera, en octobre 2008, à la fois jeune, fringuant et terriblement naïf, je n'entrevois pas la moindre bribe du changement que provoquerait en moi la complexe machinerie entre les engrenages de laquelle je venais d'insérer mes petits doigts potelés. Les trois prochaines années ne seraient rien de plus qu'un stage plus long que de coutume, pensais-je sottement. Quelques petits mois, c'est tout ce qu'il fallu pour prendre la mesure de mon erreur.

Travaillaient alors dans mon département d'accueil six autres doctorant. Mentionnons Marie et Michael, déjà au terme du sacerdoce et que je ne côtoyai par conséquent que brièvement, Guillaume le féru de météorologie et Audrey, avec qui je devais plus tard partager les joies et tensions d'une campagne de mesures aéroportées. Enfin attardons-nous sur Steph et Colin, à la fois compères et mentors, qui m'écoutèrent et m'éclairèrent de leurs lumières lorsque l'enthousiasme initial se teinta d'un soupçon de doute. Car une thèse n'est pas un stage. Son cadre est plus flou, sa finalité plus obscure, du moins lorsqu'on la débute, et l'incroyable éventail de possibilités qu'elle implique débouche sur une polymorphie déroutante pour qui a l'habitude d'arpenter des chemins balisés. L'état de l'art, c'est un peu le sketch du train de Coluche à la puissance dix ; un article peut en cacher dix autres qui eux-même en cachent cent autres et ainsi de suite, et l'on se prend bien souvent à oublier l'objectif de cette accumulation de lectures. Qu'est-ce que je fous là ? Ai-je les compétences nécessaire pour aller jusqu'au bout ? Ne ferais-je pas mieux de m'isoler dans une grotte au Népal pour réfléchir au sens de mon existence ? Autant de questions qui appellent les conseils de thésards plus expérimentés pour vous convaincre que cette remise en question fait partie des règles du jeu, que presque tout le monde passe par là ; ne faites pas l'erreur de sous-estimer les vertus d'une petite séance de psychanalyse sur le parking, en face de la machine à café, lorsque la journée a été particulièrement éprouvante.

Mais revenons à nos moutons. J'ai toujours considéré Colin comme un maître en matière de sang-froid, le genre de gaillard qui au fil des années a fini par élever le flegme au rang d'art de vivre, au point de parvenir à boucler son manuscrit de thèse presque sans se départir ni de son calme olympien, ni de sa démarche chaloupée de grand fauve indomptable. A ses côtés j'appris, entre autres indispensables choses, le respect du bon vieux navet qui tâche – le nanard ayant déjà toute ma sympathie – ainsi que les joies de la varappe en grappin sur tour génoise, de la corde à linge motorisée ou encore de l'arnaque au Monaco. A ce propos ne manquez surtout pas, à paraître aux alentours des calendes grecques, le récit illustré de nos aventures en Corse, ou comment nous passâmes à deux doigts d'être jetés dans les flots méditerranéens par un psychotique claqueur de rétroviseur ou bien gavés comme des oies par des Ténardiens rondement flippants.

Stéphanie, l'artiste touche-à-tout de la bande, aussi à l'aise devant une table à dessin que sur les planches d'une scène de théâtre - et je ne parle pas des merveilles de terre cuite qui prennent vie entre ses doigts – cache pour sa part derrière sa frêle silhouette une force de caractère hors du commun, et il en faut pour tirer son épingle du jeu dans ce milieu encore trop masculin, comme pour apprendre au piètre cuisinier que je suis l'art exigeant de la confection de sushis. Ajoutons aux multiples talents de la donzelle une habileté diabolique à coller dans le cerveau d'autrui les ritournelles les plus tenaces – sur un o, à la volette, sur un o, à la volette, sur un oranger – ainsi que la faculté, que je croyais propre à ma seule frangine, de faire vivre par son éloquence les situations et anecdotes en apparence les plus banales.

Citons également Geneviève, notre chère secrétaire, à qui je dois toujours les photos d'un certain chantier de construction situé sur l'autre versant de la rocade ; Xavier, le grand schtroumpf du labo, qui accorda quelques minutes de son temps à relire mon manuscrit alors à ses balbutiements ; le père Rivière, amené à devenir plus tard un personnage majeur de ce récit ; François et Philippe, mes compagnons de vagabondage à Reykjavík ; Olivier du fond du couloir, avec qui la conversation ne fait jamais mine de descendre en dessous du trente-troisième degré, le surréalisme en prime ; Guillaume Anna, toujours près à mettre joyeusement les pieds dans le plat ; et, bien sûr, tous les autres permanents et stagiaires.

Avant de clôturer ce premier chapitre de ma vie de doctorant, je voudrais avoir une pensée toute particulière pour une personne sans qui mon expérience toulousaine n'auraient certainement pas eu le même cachet. Mon arrivée dans cette ville toute de briques vêtue équivalait, pour moi qui sortais tout juste du cocon familial, à un véritable – et tardif, je le concède – plongeon dans l'inconnu. Le souvenir de la première soirée passée dans le vide à la fois terrifiant et libérateur de mon nouvel appartement est d'ailleurs encore vivace dans les coins et recoins de ma mémoire. Après avoir déambulé comme une bête sans but dans les rues de mon quartier sans croiser âme qui vive, recherchant vaguement un endroit où casser la croûte, je me rappelle être rentré chez moi sous la lumière blafarde d'un soleil boudeur, pour m'étendre au milieu d'une chambre ne comportant en fait de meubles qu'un sac de voyage plein à ras bords de fringues diverses. Là, cerné par une abondance de rien, je me surpris à penser qu'abandonner famille et amis pour servir la science à l'autre bout de la France – quand bien même dans la joie – était peut-être une belle connerie. J'aurais sans doute continué à me complaire un petit bout de temps dans cet état d'esprit pour le moins végétatif sans le coup de pied au euh, le coup de pouce de Mariette, mon seul contact en ces terres qui deviendrait au fil des soirées avec ses potes de médecine et autres Gaziers une amie des plus proches. J'ai souvenance d'un soir notamment où nous parcourûmes en long et en large les quais de la Garonne, nous arrêtant toutes les deux minutes pour interroger le premier badaud venu sur l'identité de l'antagoniste principal des tortues ninjas, tout cela pour finir assis au milieu d'inconnus de la dernière heure à prendre des leçons d'accordéon.

C'est beaucoup trop tôt, dix mois plus tard pour être précis, que je fis mes adieux – contraint et forcé, vous pouvez m'en croire – à la douceur du climat méridional, la rapidité des trajets quotidiens et la liberté vespérale, pour retrouver avec un sentiment mitigé la grisaille parisienne et le RER C, une heure et demi de transports depuis ma banlieue profonde pour rallier le treizième étage d'un bâtiment froid et austère où se terrait tout au bout d'un couloir sans fin le nouveau département qui devait m'accueillir. Austère, pourtant, la bande de joyeux lurons qui peuplait ses entrailles ne l'était certes pas. Le tout premier bureau que l'on croisait en pénétrant sur le territoire du DTIM constituait une sorte de salon de

thé improvisé ; on y trouvait tout ce qu'il faut pour confectionner le dit breuvage, de la théière en fonte au combo sonde/multimètre permettant de régler la température de l'eau au millième de degré près. Le visiteur pouvait y rencontrer, outre Robert la plante verte dégingandée, une population bigarrée de thésards hagards venus passer leurs rares moments d'oisiveté autour d'une chopine revigorante de tisane elfique ; Joseph et Aurélien, tout d'abord, propriétaires de l'établissement et développeurs par nécessité d'une technique imparable pour travailler tout en faisant la conversation ; le père Christophe, que l'on entendait venir de fort loin compte tenu de la façon caractéristique dont il faisait claquer ses semelles sur le carrelage ; Laure la non-macheuse de mots au caractère bien trempé et Evangeline la kamikaze du vélocipède, poissarde de la logistique aéroportée et fervente amatrice d'auto-stop. Enfin il y avait Pauline, ma colocataire de bureau, qui supporta sans jamais se plaindre le spectacle agaçant d'humeurs changeantes et de crises violentes et impromptues de gymnastique rythmique et sportive que je lui infligeais quotidiennement, sans compter le panorama en perpétuelle évolution que constituait le patchwork hétéroclite de planches de PHD Comics, photos et autres illustrations en tout genre que je scotchais avec application sur les murs de notre antre. Derrière le masque de sérieux et d'application qu'elle arborait en public, la mistinguette était toutefois bien loin d'être dépourvue des symptômes de folie douce qui font tout le sel d'une personnalité digne d'intérêt – pour un autre fou, du moins ... Ainsi, lorsqu'elle était par trop surchargée de travail ou au contraire terrassée par l'ennui, il lui arrivait de marquer son mécontentement en proférant de petits miaulement suraigus, tout en marmonnant par devers elle quelques terribles imprécations d'une voix de petite fille boudeuse. Je suis également prêt à mettre ma main à couper que, le soir venu, pelotonnée entre les pans d'un large capuchon noir, elle s'en allait distribuer aux passants de la ville lumière des tracts vantant tous les mérites de la vie au grand air à Palaiseau, sa ville, sa patrie, que dis-je, son sang et ses larmes !

Face à notre bureau se trouvait le second lieu de perdition favoris du département, j'ai nommé le secrétariat. Rares étaient ceux qui ne prenaient pas cinq minutes chaque matin pour aller boire un café chez Françoise et échanger quelques propos badins avec les quelques importuns qui avaient eu l'exacte même idée au même exact moment. Placés aux premières loges, Pauline et moi ne perdions pas une goutte du vacarme provoqué par cette porte qui s'ouvrait et se refermait toutes les trente secondes environ, laissant à chaque fois échapper qui un éclat de rires, qui une exclamation indignée de l'aubergiste essayant tant bien que mal de bosser au milieu du désordre ambiant. Parmi les habitués, citons Fabrice J, danseur émérite à l'aimable nature qui s'offrait de temps en temps quelques passes de tango avec la tenancière, quand il ne s'adonnait pas à la lutte gréco-romaine avec son ours de mari ; Martial et son trop plein de gentillesse mêlé parfois d'un soupçon de grognerie, membre le plus assidu du ciné-club local, VIP d'un pub à l'emblème pour le moins inhabituel – était-ce la fouine, le putois, peut-être la gerboise ? – et le seul Parisien mis à part mon encadrant à avoir pu assister à

ma soutenance – merci vieux frère ; Bertrand le hipster de la bande, amateur de fromages italiens et fin connaisseur des coins les plus tendance de la capitale ; Valérie, notre expat' néo-zélandaise férue de grimpette ; Paul le probabiliste intégriste, en compagnie duquel je fis la découverte du Muay Thaï, un sport aussi subtil que relaxant ; Fabrice S, la chemise à moitié ouverte et le sourire en coin, toujours le premier à balancer LA vanne qui fait plaisir, surtout quand Jérôme et/ou Paul se trouvent dans son champs de vision ; et puis Fred et Guy à qui je dois, outre les conseils avisés et autres discussions passionnantes, la découverte de quelques merveilles du sixième art comme du septième, le partage d'expériences aussi marquantes que l'usurpation de l'identité de Fantomas dans une affaire d'enlèvement caractérisé (un certain monsieur S, toujours à la recherche de ses outils de jardinage, faisait également partie du complot) ou bien la couverture photographique d'une répétition théâtrale, et enfin un séjour mémorable sur les berges du lac d'Annecy, où activités physiques et cérébrales telles que kayak, jeux de société, via ferrata et sauts de falaise ne cessaient que pour laisser place à d'orgiasques festins.

Pour parfaire le tableau, mentionnons également les joutes oratoires animant régulièrement la salle Picasso, les pique-nique participatifs et autres concours de pâtisserie, les sorties cinéma aux Gobelins ou à la Cinémathèque et les fins de soirée à la Belette – paix à son âme. En définitive, c'est bel et bien l'alchimie subtile de toutes ces personnalités hétéroclites (que celles que j'ai omis me pardonnent) qui contrebalança avantageusement le déplaisir induit initialement par mes retrouvailles avec la capitale. Rien que pour cela, merci à vous tous !

Mes travaux étaient plus proches de leur conclusion que de leur commencement lorsque j'entrepris de faire mon come-back à Toulouse. Mariette, que des affaires pressantes appelaient à Paris, me permit généreusement de squatter son appartement pendant son absence ; situé au treizième étage – étrange, il me vient comme une sensation de déjà-vu – d'un des immeubles les plus hauts de la ville, le bougre bénéficiait d'un panorama tout à fait fascinant, allant jusqu'à laisser apparaître, au loin et par temps clair, le relief accidenté du massif Pyrénéen. A son retour, six mois plus tard, nos impératif professionnels respectifs ne nous permirent pas de reprendre notre train de vie vespéral d'antan. Je parvins toutefois à la contaminer d'une addiction particulièrement virulente, que j'avais pour ma part contractée quelques mois plus tôt, concernant les aventures télévisuelles et néanmoins rocambolesques d'un vieux fou voyageant dans le temps et l'espace en cabine téléphonique. Pas rancunière pour un sou, elle n'accepta pas moins de travailler à mes côtés durant les derniers instants de la rédaction du présent manuscrit. Tandis que je me malmenais pour terminer à temps, elle s'infligea en effet un programme de révisions aux antipodes de la sinécure, afin de s'assurer une place de choix en tant qu'interne à Paris.

Mais ne nous emballons pas, avant de relater la fin de mes aventures scien-

tifiques, il me reste encore une foule de gens formidables à évoquer, à commencer par Stéphanie et Colin, mes amis thésards de la première heure, tous deux docteurs de plein droit désormais, qui réussirent l'exploit de pimenter mon futur voyage à l'autre bout du monde avant même mon départ, chapeau bas les artistes ! Notre trio, déjà joyeusement bigarré, s'étoffait de quelques personnages hauts en couleurs avec qui nous partageâmes bon nombre de moments inoubliables. Entamons les hostilités avec Nicolas, déjà présent à l'Onera Toulouse lors de mon premier passage mais avec qui je ne me liai d'amitié qu'au retour de mes dix-huit mois d'exil. Véritable gardien des secrets du DOTA, cet hyperactif au regard pétillant d'une inépuisable malice, formé à la prestigieuse école de veille cancanesque de Gouvieux-lès-petits-pois, jetterait probablement tout et n'importe quoi dans une torchère pour s'approprier la plus petite anecdote, pourvu qu'elle soit croustillante. Virtuose du sous-entendu, l'animal serait également capable de faire comprendre dix choses différentes à dix personnes à la fois avec une seule et unique phrase. En sa compagnie, je goûtai aux joies du tourisme sauvage et du bain de foule permanent à Prague et expérimentai le mystère-arnaque de la route magique (celle qui monte, mais qui en fait descend), pour le simple amour de l'art. Pierre-Yves et Hélène, baroudeurs de l'extrême et artisans touche-à-tout de la bande, bénéficieront quant à eux de mon respect éternel pour avoir transformé par leurs propres moyens une bâtisse à moitié délabrée en maison de rêve, le tout en l'espace de deux petites années. Témoins de la croissance de ce gigantesque être de bois, pierre et autres matières diverses, nous nous restaurâmes plus d'une fois au fond de ses entrailles, appréciant avec enthousiasme les truites, saucisses, gâteaux au chocolat polymorphes et autre poussier-miel d'Espigoule que nous concoctaient les maîtres de maison. Et parce qu'il ne faut se priver de rien, je vais en profiter pour souhaiter une belle et intense existence à leur fille Emmanuelle, dont la naissance précéda ma soutenance d'une toute petite semaine. Comme le disait l'ami Poulidor : "Je ferai mieux la prochaine fois !" Enfin Colin, moi et Damien, un troisième larron partageant nos goûts cinématographiques douteux, frappâmes un grand coup en fondant le CSFI ou Comité de Soutien aux Films Inutiles, dans le cadre duquel se succédèrent des soirées à thème toutes plus passionnantes les unes que les autres, se focalisant sur des artistes aussi emblématiques que Steven Seagal, Nicolas Cage ou Bill Murray – bonjour le grand écart – et se terminant invariablement par un petit jeu sans conséquence : rire ou mourir, les amateurs de comédies musicales à base de chats-robots me comprendront...

Mon retour au DOTA coïncida également avec l'arrivée d'une kyrielle de nouvelles têtes, stagiaires et autres post-doctorants, qui avec les quatre thésards arrivés fin 2010 et Myriam, seule recrue de l'année précédente, contribuèrent à faire chuter drastiquement la moyenne d'âge du département. Commençons par immortaliser par l'encre et le papier les deux grands frères, Max le Rahan de Saint-Aubin, surfeur à ses heures et compagnon de galère des fins de journées, lorsque tout le monde est parti vaquer à d'autres occupations tandis que nous, pauvres chevaux de trait harassés par la tâche, continuons de turbiner, et Fix le

géologue des ténèbres, grand prophète du demotivational et spécialiste reconnu d'utilité publique du "c'est toi" en milieu hostile. Puis viennent les petits frères et les petites soeurs, anciens stagiaires-chenilles fraîchement métamorphosés en doctorants-papillons, parmi lesquels Thomas le collectionneur de canettes et président de la Fédération Onerienne de Coinche, et Romain, fantôme évanescent du bureau du fond du couloir qui ne quittait l'entêtante mélopée de son laboratoire sous-terrain que pour de rares et courtes pauses café. De l'autre côté du dit couloir, dans un bureau réaménagé avec goût par les soins attentionnés de Fix, Thomas et moi-même se terrait Karine, redoutable esclavagiste spécialisée dans la marchandise Singapourienne et fervente admiratrice de la race bovidé, avec qui je partage, outre une même origine lozérienne, quelques souvenirs dantesques de randonnées, à pied comme à vélocyclette. La dernière, et non la moindre, n'est autre que Marianne, dit "la délurée", qui eut l'insigne honneur - ou l'injuste poisse, tout dépend du point de vue - d'hériter de la malédiction de Pauline, à savoir endurer quotidiennement ma présence inopportune. La réputation de son irrationnelle antipathie à l'égard des nouveaux venus n'étant plus à faire, elle commença par m'épier suspicieusement de derrière son écran d'ordinateur en me balançant périodiquement à la figure des pense-bête jaunes fluo roulés en boule, sans doute afin d'évaluer mes velléités de riposte en cas d'agression papetière caractérisée. Ce n'est que peu à peu, brique après brique pour ainsi dire, que je parvins à abattre le mur de sa défiance pour découvrir derrière une source inépuisable de bonne humeur, de dynamisme ... et de curiosité, car la donzelle n'a rien à envier à Nicolas sur le plan de la soif de confidences ; je ne peux d'ailleurs m'empêcher de sourire à la simple pensée de nos séances de psychanalyse vespérale. Sérieusement, Marianne, t'es vraiment un scandale !

Enfin, je ne puis tirer ma révérence sans évoquer nos confrères les plus jeunes, je veux bien sûr parler de nos amis stagiaires, en alternance ou non, qui constituèrent cette année un cru des plus précieux, aussi bien sur le plan qualitatif que quantitatif. Ainsi n'oublions-nous pas Florence la fée du bois, Caroline aux ongles arc-en-ciel, Adrien et Clément "grognon" (et sa mère) les deux mâles alphas, ainsi que Béatrice, Lilian, Elsa et David, qui nous aura tout de même bien fait rire ; et bien sûr notre cher Clément "dino", l'amoureux des tyrannosaures - surtout ceux qui pilotent des F14 - et le co-créateur de l'imagerie Roberspectrale, qui compte maintenant les moutons à l'INRA. Prends garde aux ovidés, jeune padawan, car un jour ils pourraient bien se retourner contre toi, prendre les armes pour te dire sans une once de pitié dans le regard : All your base ... are belong to US!!!

Parce que c'est dans l'ordre des choses et qu'ainsi le veut l'éternel cycle laborantin, toute cette bleusaille scientifique finit par se disperser aux quatre vents à la fin de l'été, laissant dans leur sillage une poignées de bureaux désespérément vides. Des bureaux qui restèrent à l'état de jachère pendant de nombreuses semaines avant d'accueillir pour un énième tour de manège une nouvelle fournée de jeunes gens plein d'espoir, Manu l'homme au chapeau, Ramzi, puis un peu

REMERCIEMENTS

plus tard Céline et Nicolas. Et parce qu'il faut toujours garder le meilleur pour la fin, je voudrais avoir une pensée émue pour mes deux grandes amies de fin de thèse, qui me soutinrent à la fois moralement et pâtisseriesment durant l'ultime soirée précédant ma soutenance. La première, c'est Emmanuelle la Berruyère, amoureuse de son propre fief au point que je m'étonne de ne pas la voir arborer sur la poitrine, en permanence et avec fierté, le blason d'azur à trois moutons passants de sa cité. La seconde est un souffle rafraîchissant né de la péninsule, Rosa la Catalane à la broune chevelure, infatigable fêtarde qui réussit l'exploit de mettre le turbulent Fix à genou au terme de son premier séjour parmi nous. Que d'excellents moments passés en ta compagnie, chère coéquipière de belote, du défi défilé en pyjama à la chasse à l'enjoliveur le long du canal du midi, en passant par le concours de paella dans ce terrain vague de Valencia ou bien les soirées terminées à minuit sous les Tilleuls. Tout cela n'est pas prêt de tomber dans les limbes de l'oubli.

Pour finir, j'aimerais remercier en vrac et dans le désordre tous mes amis de Paris et d'ailleurs qui ne jouèrent pas de rôle direct dans le déroulement de cette thèse mais me soutinrent à leur façon : la bande des "Gueules", Amine, Mohamed et Thomas, nemesis des restaurateurs chinois, Razaël et Marcin, mes deux compagnons d'aventure en pays Cathare qui n'oublieront certainement pas de sitôt les dangers que peut receler un chemin de traverse sous-estimé, Camillette la spécialiste bien involontaire des canassons vidéoludiques, Xavier et Rolando, les obsessionnel du coup Smash, chasseurs de meuf, voleurs de transformation et gâcheurs de glace, et puis JB, Bertolus, Benny LittleStep, Nath, Gwen, Hiraika, Omar, Conje, Béa, Ralph Baer, Chris Avellone et Shigeru Miyamoto, Jérôme Garcin et toute l'équipe du Masque, Willy Ronis, le Cyrano et le De Danu, l'assiette libanaise des Avions, Sergio du ferry Toulon-Bastia et tous les autres, même les pires, parce que si tout le monde était beau, gentil et parfaits, la vie serait bien monotone. Alors merci aux opportunistes et aux grandes gueules metteurs de pieds dans le plat, aux coincés du derche et aux menteurs invétérés, aux misogynes, aux vendus, aux imbéciles m'as-tu-vu, merci d'assumer tous ces petits travers qui nous font tant rire une fois que la pilule est passée. Continuez de nous faire rêver !

Et, bien évidemment, toute ma reconnaissance à mes parents Alain et Catherine, d'une part pour m'avoir donné la chance de trimballer ma carcasse dans ce petit monde à la fois fou et magnifique - tout le monde ne leur en saura pas gré mais moi, en tout cas, j'apprécie le geste - et d'autre part pour m'avoir élevé dans le respect de la bonne chère, l'amour du cinéma d'art et essai, de la randonnée et de la photographie, et enfin et surtout de m'avoir donné la liberté de choisir la façon dont je voulais rater ma propre vie, comme dirait l'autre. Merci à Amandine ma frangine et à Victor son dévoué tueur de zombis, à Jeanne, ma

petite grand-mère aux pieds plus petits encore, à Hélène et Dominique et Samuel et Aurélie, au petit Zouk et au turbulent Sushi, toute la fine équipe qui s'affiche joyeusement sur les plages bretonnes une fois venu le printemps. La famille, c'est peut-être la seule chose qu'on ne peut pas choisir, alors quand on l'aime et qu'elle nous le rend bien, que peut-on faire d'autre que fermer son clapet et remercier la providence ?

Pas facile de trouver une conclusion à ce qui pourrait bien représenter un vingt-cinquième de ma vie - si j'ai de la chance. Faut-il terminer sur cette note familiale, plus classique tu meurs, ou bien peut-on s'autoriser le plaisir coupable d'une dernière envolée lyrique, au risque de passer pour un illuminé mystique ou pire, un snobinard suffisant ? Sans risque, pas de gloire, alors prenons-le et ainsi soit-il !

Une personne dont mon coeur s'est épris m'a confié un jours des mots qui m'ont ému, touché, retourné, de ceux dont n'importe qui peut comprendre la portée mais que seuls les plus courageux sont capables de suivre. Ces mots du poète Cavafis, les voici :

*Lorsque tu mettras le cap sur Ithaque,
fais de sorte que ton voyage soit long,
plein d'aventures et d'expériences.
Les Lestrygons et les Cyclopes,
et la colère de Poséidon ne crains,
ils ne se trouveront point sur ton chemin
si ta pensée reste élevée, si une émotion de qualité
envahit ton esprit et ton corps. Lestrygons Cyclopes,
et la fureur de Poséidon tu n'auras à affronter
que si tu les portes en toi,
si c'est ton âme qui les dresse devant toi.*

*Fais de sorte que ton parcours soit long.
Que nombreux soient les matins
où - avec quel délice et quelle joie ! -
tu découvriras des ports inconnus,
des ports nouveaux pour toi, et tu iras
t'arrêter devant les échoppes Phéniciennes
pour acquérir les belles marchandises
nacres, coraux, ambres, ébènes
et des parfums voluptueux,
surtout beaucoup de parfums voluptueux ;
et tu iras d'une ville Egyptienne à l'autre
pour apprendre, et encore apprendre, de la bouche des savants.*

REMERCIEMENTS

*La pensée d'Ithaque ne doit pas te quitter.
Elle sera toujours ta destination.
Mais n'écourte pas la durée du voyage.
Il vaut mieux que cela prenne des longues années
et que déjà vieux tu atteignes l'île,
riche de tout ce que tu as acquis sur ton parcours
et sans te dire
qu'Ithaque t'amènera des richesses nouvelles.*

*Ithaque t'a offert le beau voyage.
Sans elle, tu n'aurais pas pris la route.
Elle n'a plus rien à te donner.*

*Et si tu la trouvais pauvre, Ithaque ne t'a pas trompé.
Sage à présent et plein d'expérience,
tu as certainement compris
ce que pour toi Ithaque signifie.*

Heureux qui comme Ulysse a fait un beau voyage. Après plusieurs années passées sur les berges d'Ogygie, l'île de Calypso, l'occasion m'est enfin donnée de reprendre le mien. Il me mènera j'espère dans bien des contrées, emmêlera à mon sort des destinées plus nombreuses encore. Chaque fois que je prendrai la mer, par nécessité, par caprice ou par amour, ce sera toujours pour rallier un port dont je ne saurai rien, et une seule chose m'apparaîtra aussi claire que l'eau fraîchement née de la roche : Ithaque, patiemment, devra ronger son frein ...

Liste des abréviations

AC	Area Classification
ACP	Analyse en Composantes Principales
ACPS	Analyse en Composantes Principales Spatiale
CCD	Charged Couple Device
CEM	Classification Expectation Maximisation
CH	Coefficients de haralick
CHA	Classification Hiérarchique Ascendante
CO	Colmar
CRF	Conditional Random Field
CS	Composantes Spectrales
EM	Expectation Maximisation
GMM	Gaussian Mixture Model
IP	Indian Pines
IRLS	Iteratively Reweighted Least Square
JSEG	J-Segmentation
LBP	Loopy Belief Propagation
MAP	Maximum A Posteriori
MCRD	Minimum Change Rate Deviation
MLR	Multinomial Logistic Regression
MNF	Minimum Noise Fraction
MR	MozaRosis
MR2	MozaRosis 2
PM	Profils Morphologiques
RC	Rapport de Croissance
ST	Spectres de Texture
SVM	Séparateur à Vastes Marges

LISTE DES ABRÉVIATIONS

- VC** Variance Cumulée
- VM** Variance Maximale
- VT** Vérité Terrain

Table des matières

Résumé	iii
Abstract	v
Remerciements	vii
Liste des abréviations	xvii
Table des matières	xix
Table des figures	xxiii
Liste des tableaux	xxv
Introduction	1
1 Contexte de l'étude	5
1 Introduction	5
2 L'imagerie hyperspectrale	5
2.1 Principe	5
2.2 Modes de représentation des données	7
2.3 Un aspect spatial	12
3 Classification	14
3.1 Principe	14
3.2 Etat de l'art des outils de classification	15
4 Etat de l'art des méthodes de classification spectrale/spatiale d'images hyperspectrales	19
5 Objectifs de la thèse et jeux de données	21
6 Conclusion	24
2 Classification vectorielle de caractéristiques spectrales et spatiales	27
1 Introduction	27
2 Méthodes de réduction de dimension pour l'extraction de caractéristiques spectrales	27
2.1 Analyse en Composantes Principales	28
2.2 Minimum Noise Fraction	29
2.3 Analyse discriminante descriptive	30
2.4 Minimum Change Rate Deviation (MCRD)	31
2.5 Sélection du nombre de caractéristiques à conserver	32

3	Extraction d'indices spatiaux	34
3.1	Matrices de co-occurrence	34
3.2	Spectres de texture	37
3.3	Profils morphologiques	40
4	Processus de classification	41
4.1	Classification supervisée par SVM	43
4.2	Classification non supervisée par modèle de mélange gaussien	46
5	Résultats	49
5.1	Mesures de précision des résultats	50
5.2	Etude de sensibilité	50
6	Conclusion	66
3	Conditional Random Field (CRF) adapté aux données hyper-	
	spectrales	69
1	Introduction	69
2	Modèles markoviens	69
3	Conditionnal Random Field (CRF)	71
3.1	Calcul des paramètres du modèle	71
3.2	Modes de représentation alternatifs des cliques d'ordre deux	74
3.3	Classification de l'image	75
4	Résultats	77
4.1	Comparaison à l'existant	77
4.2	Extension à l'analyse d'autres jeux de données	83
5	Conclusion	88
4	Classification par zones	89
1	Principe	89
2	Segmentation JSEG	89
2.1	Quantification	90
2.2	Calcul du critère d'hétérogénéité	90
2.3	Algorithme de croissance de régions	92
3	Calcul d'attributs spectraux/texturaux de zone et classification .	95
4	Résultats	97
5	Conclusion	104
5	Comparatif des approches de classification spectrale/spatiale	105
1	Précision des résultats	105
2	Complexité algorithmique	107
3	Robustesse	110
4	Conclusion	111
	Conclusion et perspectives	113

Bibliographie

117

TABLE DES MATIÈRES

Table des figures

1.1	Système d'acquisition Push-Broom	6
1.2	Comparaison de signatures spectrales	7
1.3	Liste des instruments hyperspectraux aéroportés	8
1.4	Données hyperspectrales : représentation image	8
1.5	Données hyperspectrales : espace vecteur	9
1.6	Données hyperspectrales : matrice de covariance	10
1.7	Séparation de classes	11
1.8	Divers aspects de l'information spatiale	12
1.9	Influence de la résolution sur l'information spatiale.	13
1.10	Apport de l'information spatiale	13
1.11	Exemple de classification d'une image.	14
1.12	Classification des outils de classification	15
1.13	Étapes de l'algorithme K-means	16
1.14	Classification ascendante hiérarchique	17
1.15	Principe de la classification supervisée	18
1.16	Image IP	22
1.17	Image MR	23
1.18	Image MR2	24
1.19	Image CO	24
2.1	MCRD	31
2.2	Sélection de caractéristiques : seuillage	32
2.3	Sélection de caractéristique : Scree-test	33
2.4	Création de matrices de co-occurrence	35
2.5	Exemples de coefficients de Haralick	37
2.6	Création d'une unité de texture	38
2.7	Exemple de spectre de texture pour un voisinage de taille 31×31	39
2.8	Indices spatiaux obtenus à partir des spectres de texture	39
2.9	Éléments structurants	40
2.10	Erosion et dilatation classiques	41
2.11	Ouverture et fermeture classiques et par reconstruction	42
2.12	Profil morphologique étendu	42
2.13	Schéma du processus de classification	43

TABLE DES FIGURES

2.14	Exemple d'hyperplan	44
2.15	Cas non linéairement séparable	45
2.16	Variabilité inter-classe	54
2.17	Décallages de pixels	57
2.18	Distribution de différences entre pixels voisins	58
2.19	Impact de la taille du voisinage	61
2.20	Résultats de classification obtenus avec le classifieur SVM2	64
2.21	Résultats de classification obtenus avec le classifieur GMM	65
3.1	Types de voisinage	70
3.2	Représentation alternative des cliques d'ordre 2	75
3.3	Graphe LBP	76
3.4	Impact du nombre d'itérations de l'algorithme IRLS (IP)	80
3.5	Evolution des paramètres de IRLS (IP)	81
3.6	Evolution des gradients de paramètres (IP)	82
3.7	Impact du nombre d'itérations de l'algorithme IRLS (CO, MR et MR2)	84
3.8	Résultats de l'algorithme CRF	87
4.1	Aperçu général du processus de classification par zones	90
4.2	Calcul du critère d'hétérogénéité	91
4.3	Comparaison du filtrage de Sobel avec le calcul d'une image d'hétérogénéité	92
4.4	Calcul du seuil τ_V	94
4.5	Identification des graines	94
4.6	Croissance de régions	95
4.7	Création des attributs spectraux/texturaux de zone	96
4.8	Taille de voisinage et nombre de composantes	101
4.9	Apport de l'information de texture (JSEG)	102
4.10	Illustration des diverses étapes du processus de classification par zones.	103
5.1	Cartes de classification associées aux statistiques présentées par le tableau 5.1	106
5.2	Etude de la complexité	109

Liste des tableaux

2.1	Système de calcul des précisions moyennes	52
2.2	Résultats moyens de classification vectorielle (IP)	53
2.3	Résultats de classification spectrale	55
2.4	Influence du critère de sélection du nombre de composantes à conserver	55
2.5	Influence de l'algorithme de réduction de dimension	56
2.6	Influence des paramètres associés à la création des coefficients de Haralick	59
2.7	Influence du niveau de granulométrie	59
2.8	Influence des paramètres associés à la création des spectres de texture	60
2.9	Comparaison des divers types de classification vectorielle (MR, MR2 et CO)	62
2.10	Complexité associée à la création des caractéristiques texturales .	66
2.11	Résumé	67
3.1	Pourcentage de bonne classification des résultats obtenus sur l'image IP	78
3.2	Pourcentage de bonne classification des résultats obtenus sur l'image MR	85
3.3	Pourcentage de bonne classification des résultats obtenus sur l'image MR2	86
3.4	Pourcentage de bonne classification des résultats obtenus sur l'image CO	86
4.1	Meilleurs résultats de l'algorithme AC	99
4.2	Résultats de l'étude de sensibilité	100
5.1	Etude de la précision de classification	106
5.2	Etude de la complexité	107
5.3	Récapitulatif des avantages et inconvénients de chaque algorithme	112

LISTE DES TABLEAUX

Introduction

Avec la multiplication du nombre de capteurs aéroportés, satellitaires ou autres, de plus en plus performants, notamment en termes de résolution spatiale et/ou spectrale, le traitement d'image est devenu depuis quelques dizaines d'années un domaine de recherche extrêmement porteur, menant à de nombreuses applications. Parmi ces dernières, la classification thématique de données constitue une problématique d'importance. Elle vise à transformer de façon automatique une information physique brute en information thématique directement interprétable par des spécialistes de divers horizons tels que la médecine, l'agriculture, la géologie, etc.

L'imagerie hyperspectrale est un sous-groupe de l'imagerie classique caractérisée par une résolution spectrale très fine, une image pouvant compter jusqu'à plusieurs centaines de bandes associées à des intervalles de longueur d'onde très étroits. Chaque pixel est ainsi associé à une signature spectrale permettant d'identifier précisément le ou les matériaux qui le constituent. Dans le cadre d'une application visant à regrouper les pixels correspondant à un matériau similaire, les algorithmes de classification purement spectrale sont généralement très performants. A l'inverse, lorsque les classes sont plus macroscopiques, un même matériau peut se retrouver simultanément dans plusieurs d'entre elles, ce qui complexifie considérablement la tâche de l'information spectrale. La solution consiste alors à l'associer au contexte spatial. Interprétable à différentes échelles, ce second type d'information permet de classifier le même jeu de données de plusieurs façons. Prenons l'exemple d'une image composée de vergers et de champs de vignes, deux classes partageant un matériau en commun, le sol nu, en plus des matériaux végétaux qui leur sont propres. Là où une classification purement spectrale, même supervisée, placerait les pixels de sol nu dans une même classe, la prise en compte de l'information texturale permettrait deux classes plus macroscopiques : "vergers" et "champs de vignes".

Les travaux visant à utiliser simultanément les informations spectrale et contextuelle sont nombreux. Dans sa thèse, Rellier [41] a développé deux processus de classification spatiale/spectrale en adaptant un modèle de champs de Markov gaussien à l'imagerie hyperspectrale. Wang et al. [55] ont étudié pour leur part une série d'attributs spatiaux à base de spectres de texture permettant de caractériser des textures à grande échelle. Zhong et al. [58] ont quant à eux introduit

un modèle spatial/spectral à partir d'un concept de *Conditionnal Random Field* (CRF) nécessitant toutefois la détermination d'un nombre très élevé de paramètres. Plus récemment enfin, Tarabalka [48, 51] a entrepris de régulariser des résultats de classification purement spectraux au moyen de cartes de segmentation réalisées au préalable. L'ensemble de ces travaux adopte un point de vue spatial/spectral de l'information thématique à mettre en évidence, de sorte que ces deux types d'information sont utilisés conjointement à des fins de classification. Malgré tout, les approches développées à ce jour sont très majoritairement supervisées et de nombreuses façons de mettre en évidence l'information spatiale restent à étudier dans le cadre d'une utilisation complémentaire à l'information spectrale. C'est ce qui justifie cette étude durant laquelle nous avons développé de nouvelles méthodes de classification spatiale/spectrale d'images hyperspectrales, supervisées ou non.

Le sujet de cette thèse met en relation un certain nombre de concepts, domaines et applications scientifiques qu'il nous faut avant tout définir. Nous commençons ainsi par présenter ce qu'implique l'imagerie hyperspectrale en termes de physique, de capteurs et de représentation sur l'outil informatique, avant de décrire les différentes formes que peut revêtir l'information spatiale au sein d'une image, ainsi que la plus-value qu'elle peut apporter à l'information spectrale dans le cadre d'une application de classification thématique. Suite à un état de l'art des outils de classification spectraux d'une part, et spectraux/spatiaux d'autre part, nous procédons à un inventaire des jeux de données dont nous disposons, chacun d'eux présentant des particularités exploitables par une méthode de classification spectrale/spatiale bien précise.

Trois méthodes, correspondant chacune à une façon différente d'extraire l'information spatiale, ont été mises en oeuvre puis analysées. La première est une approche de classification vectorielle fonctionnant sur la base d'attributs spectraux et texturaux rassemblés sous la forme d'un unique jeu de données traité soit de façon supervisée avec un algorithme SVM, soit de façon non supervisée avec un modèle de mélange gaussien. Les attributs spectraux sont obtenus par réduction de dimension de l'image hyperspectrale originale tandis que les attributs texturaux sont réalisés par l'intermédiaire d'une étude statistique (coefficients de Haralick, spectres de texture) ou morphologique (profils morphologiques) du voisinage de chaque pixel. Cette méthode est simple et relativement rapide mais n'inclut pas véritablement l'information spatiale au sein du modèle de représentation des données. Avec la seconde approche, nous avons ainsi entrepris de bâtir un modèle de représentation spatial/spectral en nous inspirant de l'algorithme proposé par Zhong et al. [58] que nous avons mentionné précédemment. La version que nous présentons vise d'une part à diminuer de façon drastique le nombre de paramètres à estimer et d'autre part à améliorer le mode de représentation de l'information spatiale au sein du modèle. Enfin, la dernière approche est un pro-

cessus de classification par zones déterminées par l'intermédiaire d'une segmentation préalable de l'image, l'algorithme utilisé présentant la particularité de tenir compte de l'information de texture pour éviter une éventuelle sur-segmentation. Les zones ainsi identifiées sont ensuite associées à un ensemble d'attributs spectraux et/ou texturaux puis fusionnées au moyen d'une classification hiérarchique ascendante jusqu'à atteindre le nombre de classes désiré.

La dernière partie de ce manuscrit compare ces trois méthodes sur la base de trois critères : précision des résultats, complexité et robustesse vis-à-vis des paramètres.

Chapitre 1

Contexte de l'étude

1 Introduction

Ces travaux de thèse ont pour objectif d'exploiter simultanément les dimensions spatiale et spectrale d'une image, l'idée étant de tirer parti de la haute résolution spectrale et spatiale des données hyperspectrales les plus récentes afin d'améliorer la précision des algorithmes de classification. Dans ce chapitre nous commençons par présenter brièvement les caractéristiques propres à l'imagerie hyperspectrale, ainsi que ce que sous-entend la notion de contexte spatial en imagerie. Nous nous focalisons ensuite sur le concept de classification avant de réaliser un état de l'art des méthodes de classification alliant informations spatiale et spectrale dans un contexte hyperspectral. Enfin, nous présentons les quatre jeux de données utilisés dans le cadre de cette étude en insistant notamment sur la raison pour laquelle leur usage est pertinent vis-à-vis des méthodes que nous avons développées.

2 L'imagerie hyperspectrale

2.1 Principe

La télédétection vise à capter, enregistrer puis traiter un signal électromagnétique suite à son interaction avec une cible - la Terre, dans notre cas. En imagerie passive, le signal en question est le rayonnement émis par le soleil. Une image dite *monospectrale* est une matrice 2D dont chaque élément (appelé pixel) correspond à un échantillon spatial de la scène sur lequel est intégré le signal reçu. Nous considérerons plusieurs grandeurs radiométriques. La luminance est définie par :

$$L = \frac{d^2\Phi}{d\omega d\Sigma} \quad (1.1)$$

et correspond à la mesure du flux $d^2\Phi$ (exprimée en $W.m^{-2}.sr^{-1}$) correspondant à un angle solide $d\omega$ percutant une surface $d\Sigma$ perpendiculaire à l'axe du

rayon électromagnétique. La fraction de luminance contenue dans un certain intervalle de longueurs d'onde $d\lambda$ est appelé luminance spectrale et s'exprime en $W.m^{-2}.sr^{-1}.\mu m^{-1}$. Enfin dans l'hypothèse lambertienne où la diffusion est uniforme dans toutes les directions, la réflectance :

$$\rho = \frac{F_r}{F_i} \quad (1.2)$$

correspond au rapport entre la densité de flux (luminance) réfléchi F_r et la densité de flux incidente F_i . Cette dernière ne dépend que de la longueur d'onde, de la nature du matériau ainsi que de la configuration angulaire de visée et d'éclairage dans le cas d'une surface anisotrope. L'intensité du signal électromagnétique et les effets atmosphériques ne rentrent pas en ligne de compte, ce qui justifie notre choix d'utiliser des données en réflectance plutôt que des données en luminance.

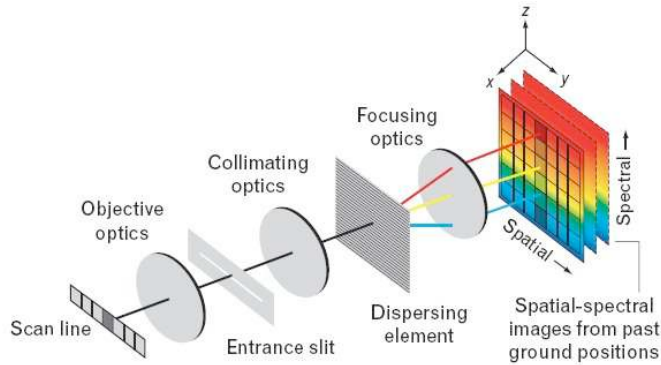


FIG. 1.1: Système d'acquisition Push-Broom : l'image est acquise ligne par ligne au moyen d'une matrice de détecteurs CCD (Charge-Coupled Device) que l'on fait glisser tel un balais au-dessus de la zone d'étude. Cette matrice est constituée de N lignes, chacune associée à un intervalle de longueurs d'onde. Lors de chaque acquisition, le rayonnement électromagnétique est divisé par un système dispersif de sorte que chaque ligne de CCD ne capte que la fraction de rayonnement de l'intervalle de longueurs d'onde auquel elle est associée.

En parallèle d'une tendance visant à obtenir des images de mieux en mieux résolues spatialement, l'imagerie *hyperspectrale* est issue de la volonté de modéliser aussi précisément que possible le comportement réfléchissant des matériaux et objets constituant une scène, de sorte que chaque matériau soit associé à une signature spectrale (appelée spectre) qui lui est propre. Un capteur hyperspectral échantillonne le signal électromagnétique reçu en une série de bandes spectrales étroites et contiguës (cf. figure 1.1); un pixel n'est donc pas associé à une valeur scalaire mais à un vecteur de dimension M égale au nombre de bandes de l'image (de l'ordre de quelques centaines). La figure 1.2 témoigne de la variété de comportements spectraux permettant de différencier divers matériaux.

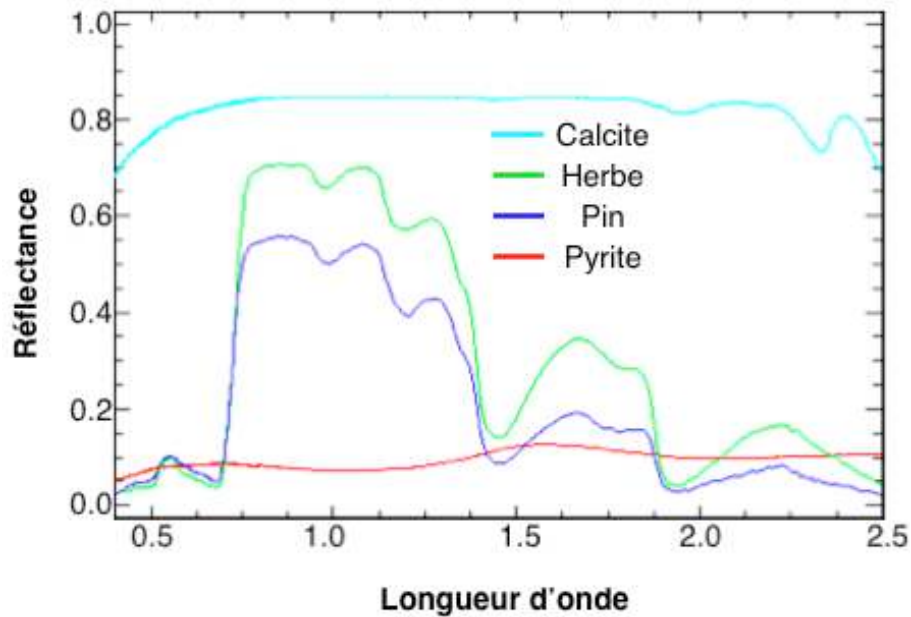


FIG. 1.2: Comparaison de signatures spectrales (en réflectance) correspondant à différents matériaux : en cyan le calcite, en rouge la pyrite, en vert l'herbe et en bleu le pin. Ces matériaux se distinguent les uns des autres par leur façon d'absorber le rayonnement électromagnétique en fonction de la longueur d'onde

Il existe aujourd'hui plusieurs dizaines de capteurs hyperspectraux couvrant des intervalles de longueurs d'onde allant de 0,3 à 15 μm . La majorité d'entre eux sont placés sur support aéroporté (voir figure 1.3).

2.2 Modes de représentation des données

Nous considérerons principalement trois façons de représenter les données hyperspectrales [29] : la représentation image, la représentation spectrale et l'espace vecteur.

La représentation image est sans conteste la plus intuitive puisqu'elle permet l'affichage des données sous la forme d'une image correspondant à la vue que l'observateur aurait s'il se mettait à la place du capteur (cf. figure 1.4). C'est également la seule représentation capable de mettre en valeur les relations contextuelles existant entre pixels d'un même voisinage. Cependant, si elle est exhaustive d'un point de vue spatial, elle ne rend pas compte de la richesse de l'information spectrale contenue dans une image hyperspectrale.

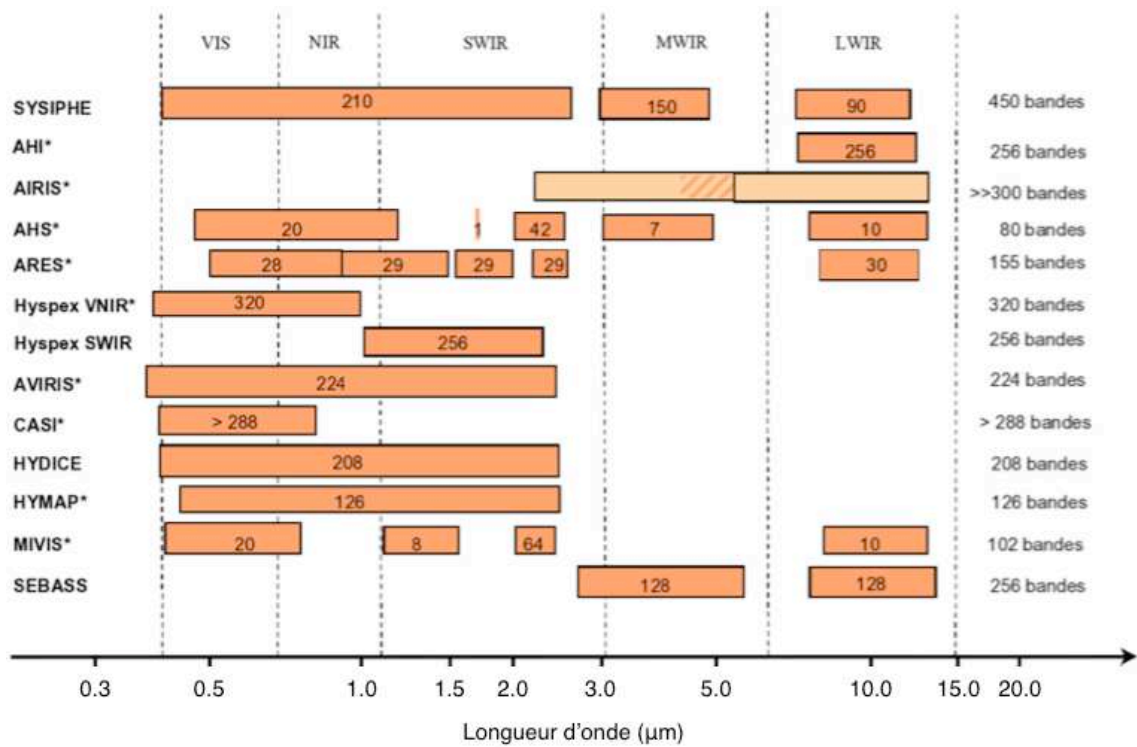


FIG. 1.3: Liste des instruments hyperspectraux aéroportés



FIG. 1.4: La représentation image, la plus intuitive. L'exemple ci-dessus est une composition colorée faite à partir de trois bandes d'une image hyperspectrale (proche infra-rouge : 830 nm, rouge : 690 nm et vert : 550 nm).

Cette richesse est plus aisément perceptible et analysable par le biais de la représentation spectrale. Cette dernière correspond à un graphique décrivant le spectre d'un ou plusieurs pixels (cf. figure 1.2) en fonction de la longueur d'onde du rayonnement. Remarquons toutefois que c'est l'information contextuelle qui est cette fois complètement absente de ce type de représentation.

Enfin, l'analyse spectrale fine de l'ensemble des pixels d'une image hyperspectrale est plus facile à réaliser avec le dernier type de représentation, l'espace

vecteur, représenté sur la figure 1.5. Le spectre associé à chaque pixel est représenté par un vecteur dans un espace à M dimensions, M étant le nombre de bandes spectrales de l'image. Les composantes de ce vecteur sont égales aux valeurs de luminance ou de réflectance perçues par le capteur sur chacune des bandes. Nous assimilons ainsi l'image hyperspectrale à une matrice $M \times N$ - où N est le nombre de pixels - dont il est aisé de faire ressortir une information statistique utilisable pour l'analyse de données. Il faut toutefois remarquer que la notion de voisinage est de nouveau absente de ce type de représentation. La taille d'une telle matrice peut également être une source de problèmes.

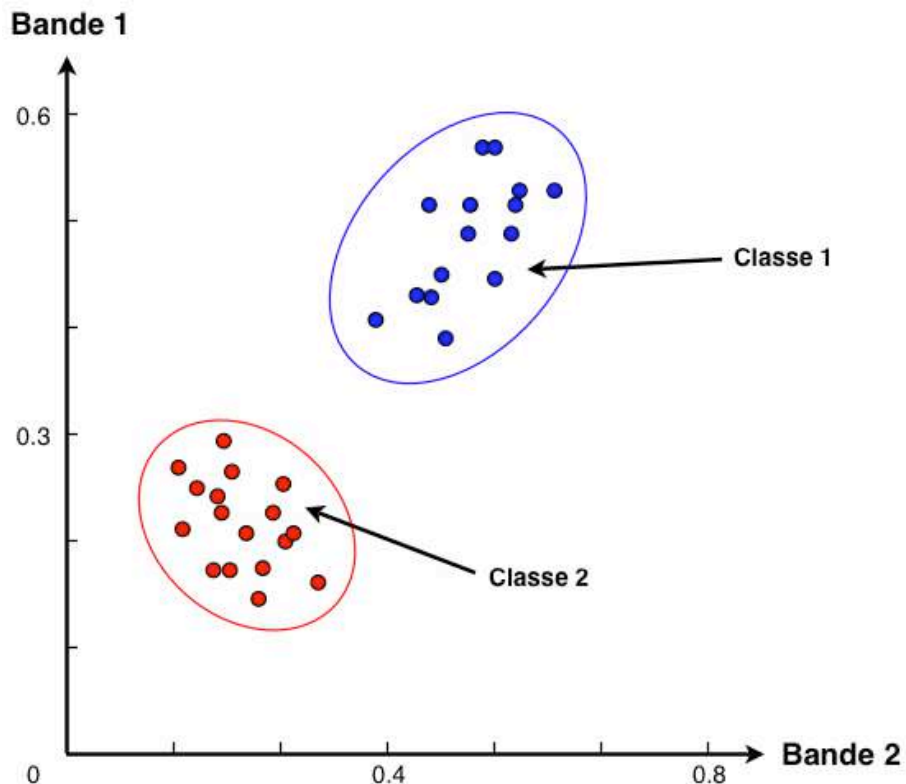


FIG. 1.5: L'espace vecteur permet de regrouper les pixels par groupes présentant des statistiques spectrales similaires.

2.2.1 Réduction de dimension

Dans un contexte hyperspectral, la grande profusion de données rend difficile l'extraction d'une information claire et pertinente à partir des seules données brutes. En effet, leur caractère très fortement multivarié les rend sensibles au phénomène de Hughes : plus la dimension des données étudiées est grande et plus le nombre de paramètres nécessaire à l'élaboration d'un modèle permettant de les représenter augmente. Ces paramètres sont d'autant plus délicats à estimer qu'ils sont nombreux, ce qui entraîne une dégradation de la qualité du traitement. De

plus, une image de grande dimension implique une multiplication non négligeable du temps de calcul des algorithmes d'analyse.

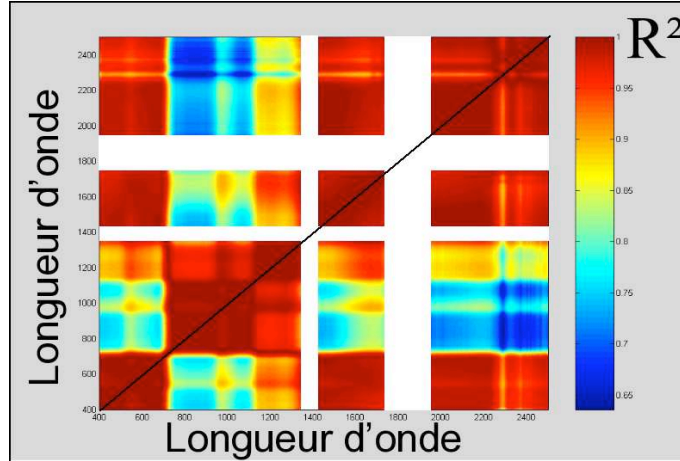


FIG. 1.6: Matrice de covariance d'une image hyperspectrale : les similitudes inter-bandes par blocs sont parfaitement visibles.

Pour contrebalancer ce problème, il existe des procédés tirant parti de la précision de l'échantillonnage spectral et des redondances inter-bandes des données hyperspectrales pour obtenir un nouveau jeu réduit de variables, ces dernières étant définies à partir d'un critère adapté à l'application désirée (la variance pour la classification, le kurtosis pour la détection d'anomalie, etc...). En effet, les bandes spectrales adjacentes présentent parfois de fortes similitudes, entraînant une forte corrélation inter-bande (cf. figure 1.6). Ce comportement est dû au fait que les propriétés de réflectance spectrale des matériaux restent souvent stables sur des intervalles de longueurs d'onde plus ou moins grands.

Cette réduction de dimension de l'espace vecteur peut être fait de deux manières, soit par des méthodes de sélection de bandes [2, 42], soit par projection des données sur une espace de dimension inférieure, ce dernier procédé débouchant sur l'obtention de nouvelles primitives résultant d'une combinaison des bandes spectrales originales (voir figure 1.7). Dans les deux cas, l'objectif est de minimiser l'écart en termes d'information utile entre l'image d'origine et l'image réduite. Nous utiliserons cependant exclusivement les méthodes de projection qui présentent l'avantage de tenir compte de la totalité des bandes spectrales originales. Ces méthodes tendent de plus à rendre les distributions gaussiennes, plus faciles à modéliser.

Il existe de nombreux algorithmes de projection, dont certains seront détaillés dans le chapitre 2 de ce manuscrit.

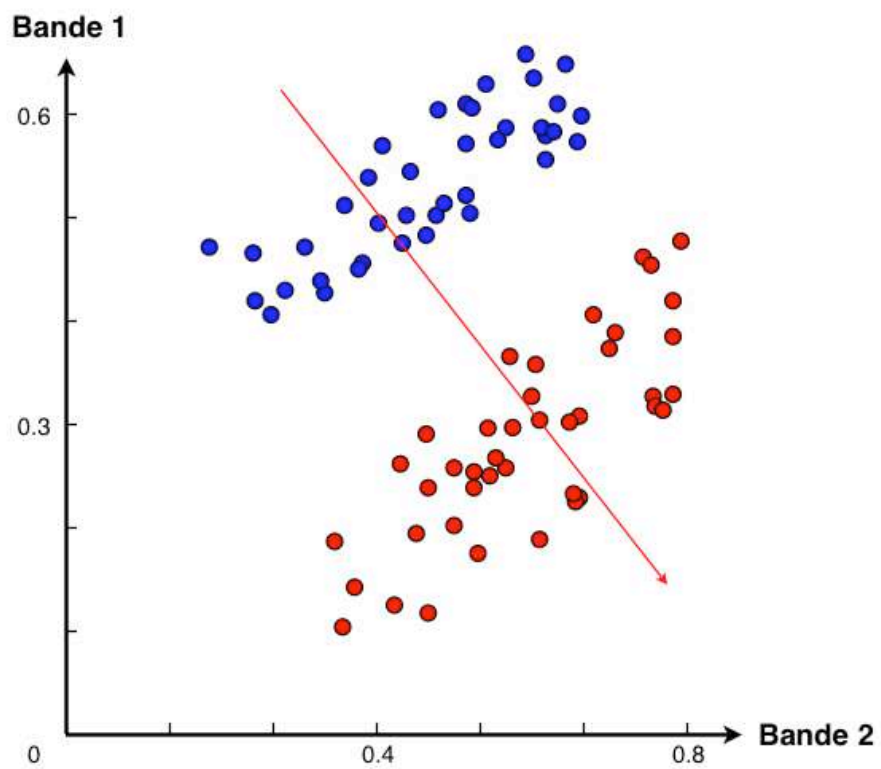


FIG. 1.7: Exemple d'axe de projection (en rouge) permettant de séparer de manière optimale deux classes de pixels sur une image à deux bande.

2.3 Un aspect spatial

Dans une image, l'information spatiale peut revêtir plusieurs aspects : forme, orientation ou texture ; des aspects dont l'intérêt dépend de l'application considérée (cf. figure 1.8). Toutefois, quel que soit le point de vue adopté, l'information spatiale reste une information de type contextuel. Autrement dit les pixels ne sont pas considérés comme des entités ponctuelles totalement indépendantes de leur environnement. Ils font partie d'une population et présentent à ce titre un lien aussi bien statistique que sémantique avec leur voisinage immédiat.

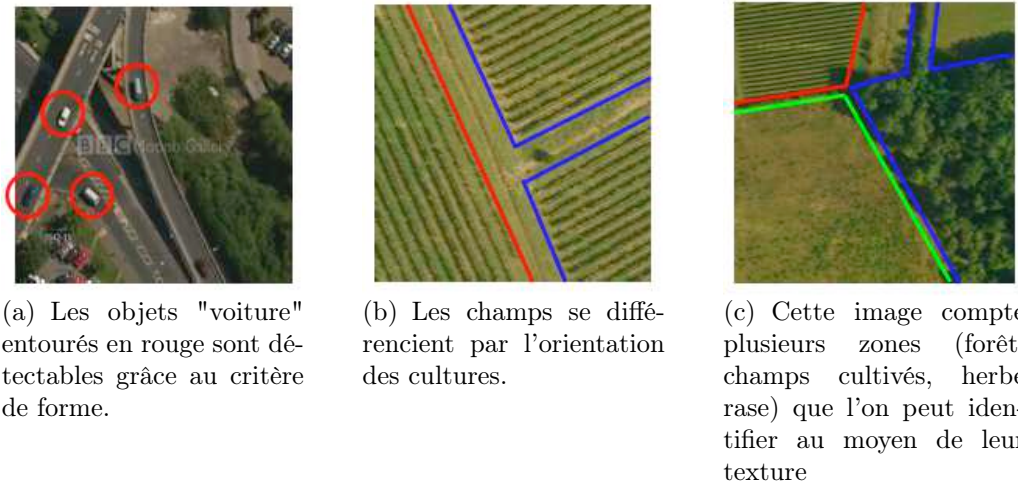


FIG. 1.8: Divers aspects de l'information spatiale

L'aspect spatial d'une image est directement lié à sa résolution spatiale. Cette dernière correspond aux dimensions réelles de la zone couverte par un pixel de l'image. Elle conditionne le plus petit niveau de détails proposé par une image mais détermine également le type d'information que l'on va pouvoir retirer de cette image, que ce soit à l'oeil nu ou au moyen de traitements informatiques. En effet selon l'échelle utilisée, l'information spatiale relative à une même cible peut être très différente (cf. figure 1.9).

La figure 1.10 met en valeur l'intérêt de l'information spatiale. L'image 1.10(a) est constituée de deux zones, en haut et en bas, composées de plusieurs occurrences des deux mêmes spectres. Cependant, si ces deux zones sont identiques d'un point de vue spectral, elles se distinguent par l'agencement spatial des pixels qui la composent. Le résultat 1.10(b) a été obtenu à partir d'une classification SVM (Séparateur à Vastes Marges) appliquée sur un ensemble d'indices spectraux tandis que pour le résultat 1.10(c), le classifieur a été nourri à la fois par des indices spectraux et spatiaux. Là où le premier classifieur se contente de discriminer admirablement bien les deux types de spectre, le second parvient à distinguer très clairement les deux zones texturées, ce qui nous montre que l'information spatiale permet de considérer des classes d'un autre niveau sémantique.

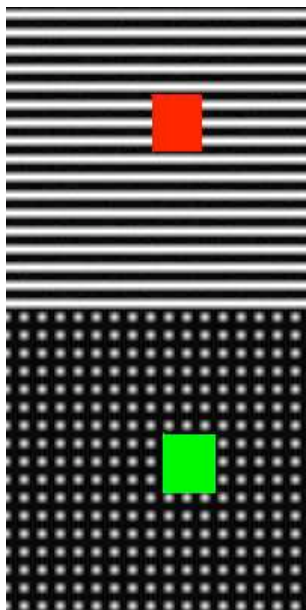


(a) Cette image Landsat, de faible résolution, ne présente qu'une information de texture (tapis urbain et forêt).

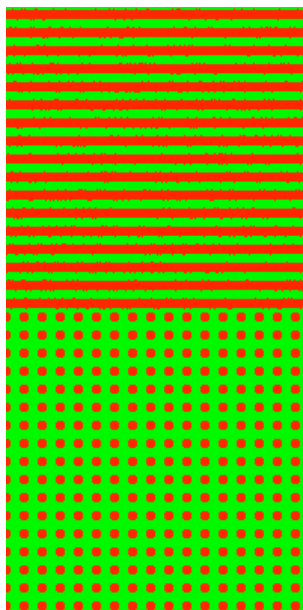


(b) Cette image haute résolution présente une information de texture et de forme (arbres et bâtiments).

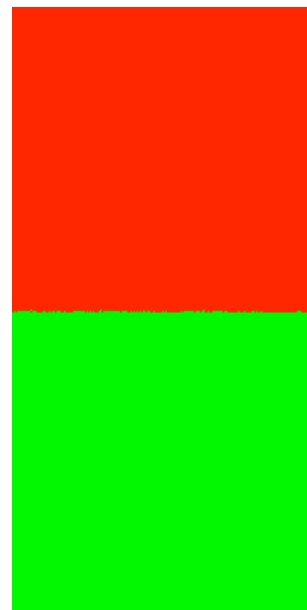
FIG. 1.9: Influence de la résolution sur l'information spatiale.



(a) Première composante ACP de l'image originale (avec les échantillons d'apprentissage pour chaque zone en rouge et vert)



(b) Résultat de Classification SVM sur un ensemble d'indices spectraux



(c) Résultat de Classification SVM sur un ensemble d'indices spectraux et spatiaux

FIG. 1.10: Apport de l'information spatiale

3 Classification

L'objectif de ce travail est d'utiliser conjointement les deux types d'information évoqués dans la section précédente à des fins de classification. Dans cette section, nous introduisons brièvement le concept de classification ainsi que les outils les plus couramment utilisés à ce jour.

3.1 Principe

La classification est un procédé algorithmique visant à diviser l'ensemble des éléments $\{p_i, i \in \{1, \dots, N\}\}$ d'une population P en un nombre C de partitions, selon un critère de similarité dépendant du contexte et du résultat désiré. En imagerie où P correspond à une image et p_i à un pixel, l'objectif est de produire des cartes thématiques mettant en valeur de manière simple et schématique différentes régions élémentaires d'une scène (cf. figure 1.11). Une partition - ou classe - représente une famille de pixels ayant des propriétés communes et qui identifie ainsi une ou plusieurs zones présentant des caractéristiques physiques (type de sol, de végétation, taux d'humidité) ou contextuelles propres. Cela rend la notion de similarité entre pixel double : spectrale et/ou spatiale. La qualité des résultats obtenus dépendant grandement de la robustesse des classes, ces dernières doivent satisfaire aux trois exigences suivantes :

- Exhaustivité : tous les pixels de l'image doivent pouvoir être associés à une classe.
- Séparabilité : les classes doivent être suffisamment différenciables pour qu'un pixel ne puisse être associé qu'à une seule classe.
- Pertinence : l'ensemble des classes doit refléter l'information que l'utilisateur veut retirer de l'image.

Ces exigences doivent être vérifiées simultanément, les deux premières étant liées aux données tandis que la dernière dépend de l'utilisateur.

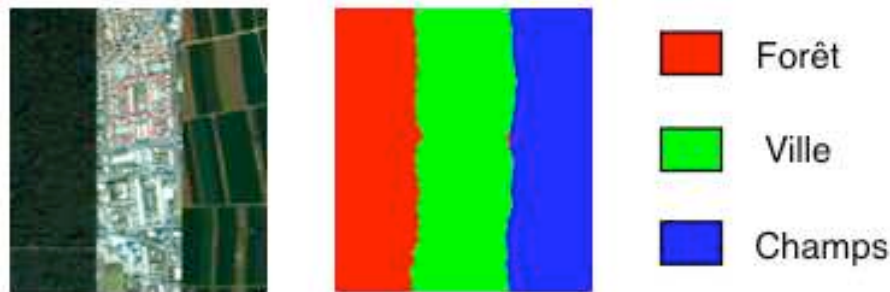


FIG. 1.11: Exemple de classification d'une image.

3.2 Etat de l'art des outils de classification

Les algorithmes de classification peuvent être repertoriés de plusieurs manières (voir figure 1.12). Un premier critère est le niveau de supervision, c'est-à-dire l'importance de l'intervention de l'utilisateur dans le processus de classification. Il est également possible de séparer les algorithmes paramétriques - associés à un modèle probabiliste - des algorithmes non paramétriques. Nous utiliserons le premier critère dans l'état de l'art qui suit.

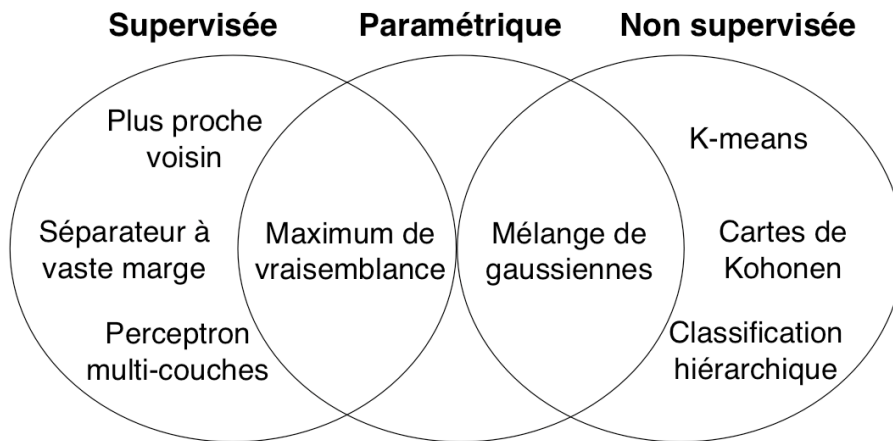


FIG. 1.12: Classification des outils de classification

3.2.1 Classification non supervisée

Dans le cas d'une classification non supervisée, l'image est partitionnée sans utiliser de connaissance a priori (nombre de classes, ensembles d'échantillons déjà étiquetés). De tels algorithmes de classification analysent la répartition des données au sein de leur espace de représentation de sorte à déterminer des groupements naturels correspondant à des zones de forte densité. Ces outils sont très prisés, principalement parce que l'acquisition de vérités terrain précises (un processus long et délicat) n'est pas systématiquement possible.

Il existe principalement deux grandes familles d'approches : les classifications par partitionnement et les classifications hiérarchiques. La première considère que l'espace de représentation est divisible en un nombre fini de régions disjointes où toutes les observations appartiennent à la même classe. L'algorithme K-means [32] est le représentant le plus répandu de cette famille. Son objectif est de partitionner un ensemble de n observations (x_1, x_2, \dots, x_n) en k groupes (G_1, G_2, \dots, G_k) tout en minimisant la somme des distances à la moyenne :

$$\arg \min \sum_{i=1}^k \sum_{x_j \in G_i} \|x_j - \mu_i\|^2, \mu_i = \frac{1}{|G_i|} \sum_{x_j \in G_i} x_j \quad (1.3)$$

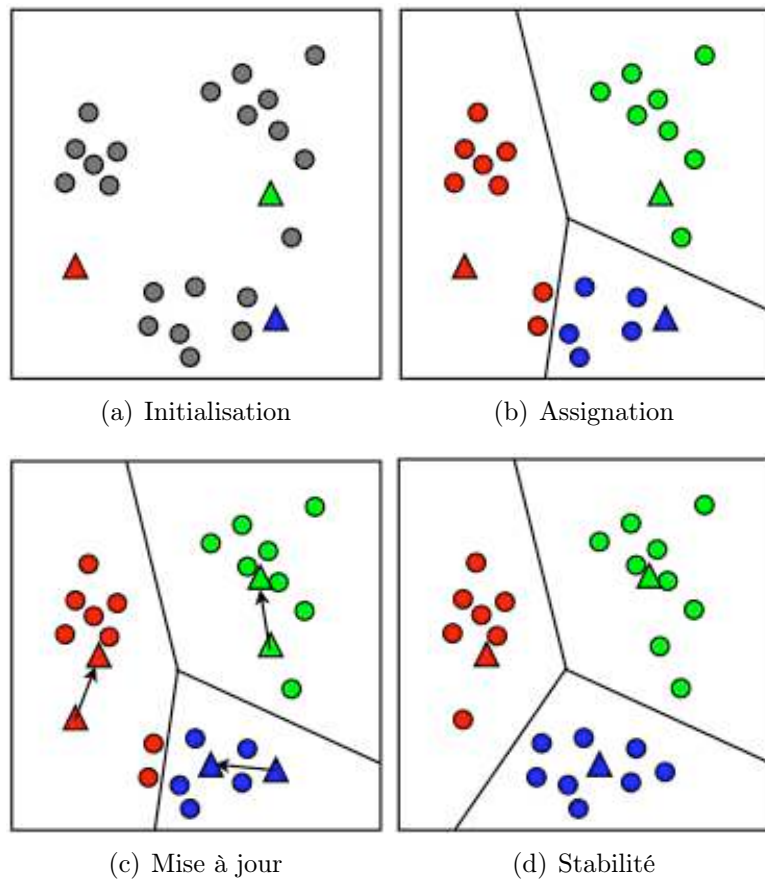


FIG. 1.13: Étapes de l'algorithme K-means : les centroïdes de classes, représentés par les triangles, sont choisis aléatoirement parmi les pixels de l'image (les ronds). Ces derniers sont associés au plus proche centroïde qui est alors mis à jour. Ces deux étapes sont répétées jusqu'à stabilité.

La solution optimale n'étant pas atteignable en temps polynomial, c'est généralement une heuristique qui est employée. L'heuristique standard consiste, partant d'un ensemble de centroïdes (c_1, c_2, \dots, c_k) choisis aléatoirement parmi les pixels de l'image et représentant chacun un groupe (G_1, G_2, \dots, G_k) , à alterner de manière itérative entre une étape d'assignation de label et de mise à jour des centroïdes. Durant l'étape d'assignation, tous les pixels de l'image à classifier sont associés au groupe G_i dont le centroïde est le plus proche au sens de la distance euclidienne. Puis les centroïdes de chaque groupe sont recalculées durant l'étape de mise à jour. Ces deux étapes sont répétées jusqu'à stabilité (voir figure 1.13).

Cet algorithme est à la fois rapide et simple mais outre le fait de devoir fixer au préalable le nombre de groupes, ses résultats présentent également une forte dépendance au choix des centroïdes initiaux, ce qui implique généralement de réitérer plusieurs fois le processus complet afin de minimiser les chances de converger vers un optimum local. Il existe plusieurs extensions de cet algorithme,

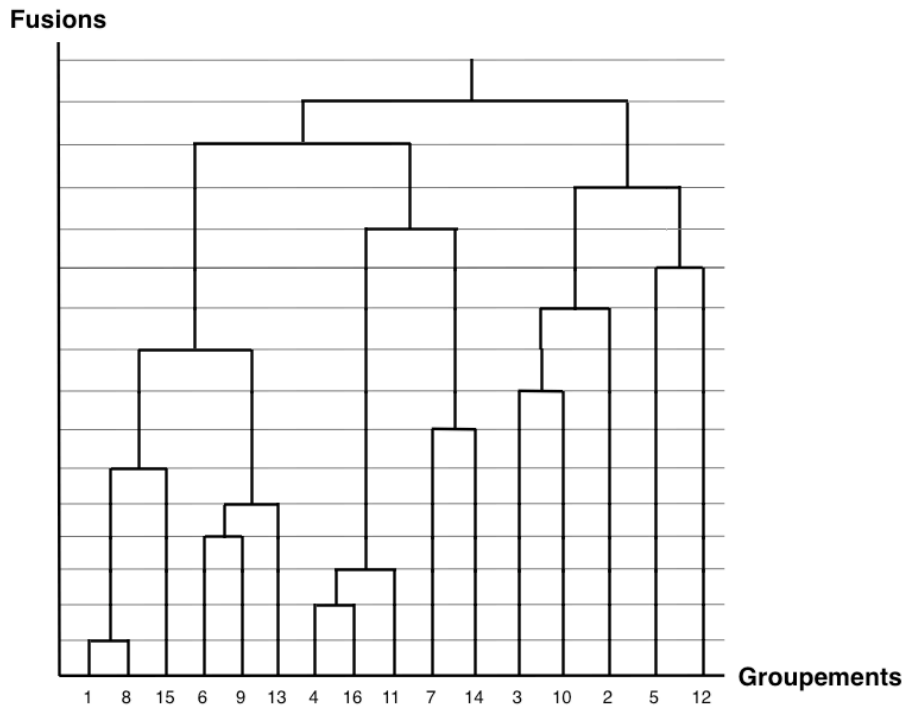


FIG. 1.14: Dendrogramme représentant la succession de fusion d'une classification ascendante hiérarchique

dont les cartes de Kohonen [25] ou cartes auto-organisatrices qui introduisent un système de topologie a priori au niveau des centroïdes de classe. Ces cartes sont généralement représentées par des réseaux de neurones double couches que l'on fait évoluer par apprentissage en leur soumettant des observations qui seront associés à l'un ou l'autre des centroïdes de classe. Les relations de voisinage entre ces derniers permettent d'en faire évoluer plusieurs simultanément (en partant du principe que s'il y a une zone d'intérêt dans une direction, il y en a sans doute d'autres dans la même direction qui sont représentés par des centroïdes différents) et donc de réduire la durée du processus d'apprentissage ; la classification par carte de Kohonen est ainsi particulièrement rapide. Cependant, la topologie optimale des centroïdes est difficile à estimer à l'avance et tout comme l'algorithme K-means, ce type de méthode peut être mis en difficulté en cas de chevauchements de classes. Afin de réduire l'impact de ces derniers, il existe une seconde extension (paramétrique, cette fois) du K-means, les modèles de mélange gaussiens [33], qui laisse plus de libertés quant à la forme des distributions de classes. Les distributions n'y sont plus forcément définies sphériques mais peuvent être ellipsoïdales et de taille variée. Ces modèles seront décrits plus avant dans la partie 4.2.

Parallèlement à ces méthodes dites de *clustering*, les algorithmes de classification ascendante ou descendante hiérarchique génèrent, respectivement par fusion

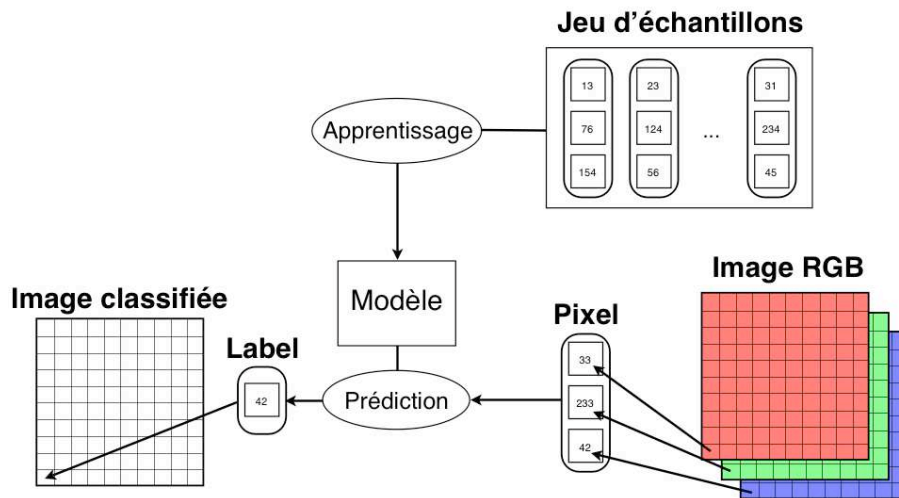


FIG. 1.15: Principe de la classification supervisée : le modèle est tout d'abord créé par apprentissage des données d'entraînement présentes dans le jeu d'échantillons. Les pixels de l'image sont ensuite injectés un à un dans le modèle (lois de probabilité, réseau de neurones, jeu d'hyperplans) qui leur associe un label.

ou par division, une suite de partitions des observations allant de la partition monoclasse, où la même classe englobe la totalité des observations, à la partition où chaque observation correspond à une classe (voir figure 1.14). L'objectif de ces outils est de proposer à l'analyste de choisir manuellement ou automatiquement la partition de classe la plus réaliste. Il en existe plusieurs variantes qui se distinguent essentiellement par leur critère de fusion ou de division des classes (distance euclidienne, angle spectral, etc).

3.2.2 Classification supervisée

L'approche supervisée consiste, à partir de plusieurs échantillons de l'image cible (que l'on appelle également données d'apprentissage) représentatifs de chacune des classes, à élaborer un modèle (probabiliste, géométrique ou structurel) à partir duquel le reste des éléments de l'image cible seront étiquetés (voir figure 1.15). La pertinence du résultat de classification est en grande partie déterminée par la représentativité des données d'apprentissage, qu'il n'est pas forcément évident de rassembler de manière précise et qui dépendent le plus souvent de l'information thématique recherchée.

Parmi les méthodes non paramétriques le procédé le plus simple est celui des k plus proches voisins [10]. Ce dernier consiste à assigner au pixel étudié la classe majoritaire parmi les k échantillons qui lui sont le plus proche dans le nuage de points considéré. Ce procédé est simple mais une base d'apprentissage trop

grande a tendance à augmenter très fortement les temps de calcul. Plus récent, l'algorithme des Séparateurs à Vastes Marges [7] (SVM) utilise des hyperplans pour séparer les classes au sein de l'espace de représentation des données. Sa complexité est également dépendante de la taille de la base d'apprentissage, mais son avantage réside en ce qu'il donne de bons résultats même lorsque le nombre d'échantillons utilisés pour l'apprentissage est limité. Le seul défaut réside dans la difficulté de choisir puis paramétrer un noyau adapté à tel ou tel jeu de données. Nous évoquerons cet algorithme plus en détails dans le chapitre 2. Les réseaux de neurones multi-couches (*Multi-Layer Perceptron*) [26] constituent une autre approche de classification supervisée non paramétrique. Ces réseaux sont constitués d'un minimum de trois couches de neurones. Afin de paramétrer les poids associés à ces derniers, l'algorithme d'apprentissage le plus répandu est cette fois-ci la rétropropagation du gradient. Ce dernier propage l'erreur quadratique obtenu sur les échantillons d'apprentissage et détermine la contribution des poids à l'erreur générale, puis il les corrige de manière à se rapprocher le plus possible du résultat souhaité.

Enfin, les outils de classification supervisées paramétriques déterminent les frontières de décision entre classes grâce à un modèle probabiliste défini par apprentissage. Partant de ce modèle, la phase de prédiction permet alors d'associer une classe au reste des pixels de l'image. L'exemple typique est la classification par maximum de vraisemblance [45] qui associe une loi de probabilité conditionnelle présumée gaussienne à chacune des classes considérées, auxquelles vont correspondre une zone de probabilité maximale au sein de l'espace vecteur.

4 Etat de l'art des méthodes de classification spectrale/spatiale d'images hyperspectrales

Dans la présente section, nous passons en revue diverses méthodes issues de la littérature utilisant simultanément les informations spectrale et spatiale d'une image hyperspectrale à des fins de classification.

Une première approche consiste à calculer diverses caractéristiques spectrales et spatiales pour chaque pixel de l'image puis à procéder à une classification pixel à pixel de ces dernières. Les deux types d'informations peuvent être classifiées indépendamment, résultant en une carte de classification spatiale et une carte de classification spectrale qui sont fusionnées selon diverses règles. C'est ainsi que procèdent Benediktsson et al. [4], utilisant un réseau de neurones pour classifier un ensemble de profils morphologiques préalablement calculés sur les premières composantes principales de l'image et une approche type maximum de vraisemblance pour classifier des bandes spectrales issues d'une réduction de dimension par *Nonparametric Weighted Feature Extraction* (NWFE). Les carac-

téristiques peuvent également être concaténées les unes aux autres pour former un unique jeu de données, auquel on applique une classification vectorielle. Lepisto et al. adoptent cette approche avec des caractéristiques spatiales de type bancs de filtres [30] ou coefficients de Haralick [31] et un classifieur de type k -plus proches voisins. Les bancs de filtres sont des caractéristiques spatiales obtenues par filtrage directionnel de l'image au moyen de divers outils de traitement du signal tels que les filtres de Gabor [14] ou les transformées en ondelettes [47]. Les coefficients de Haralick sont quant à eux obtenus en étudiant les transitions de niveaux de gris pour un décalage de pixels donné. Enfin, Palmason et al. [36] et Fauvel et al. [15] font de même avec des profils morphologiques, un troisième type de caractéristique spatiale obtenue par filtrage morphologique de l'image. Les classifieurs utilisés sont alors respectivement un réseau de neurones et une machine à vecteurs supports. Mis à part les profils morphologiques, ces caractéristiques spatiales sont encore peu employées dans le cadre de l'imagerie hyperspectrale, et encore moins comparées les unes aux autres. De plus, elles sont généralement utilisées dans un contexte supervisé nécessitant une information a priori sur les images à classifier dont nous ne disposons pas toujours.

Une seconde approche consiste à adapter la théorie markovienne à l'imagerie hyperspectrale. La modélisation par champ de Markov constitue une approche probabiliste classique permettant de prendre en compte les interactions spatiales entre les pixels d'une image. Elle assimile l'image à un champ de variables aléatoires (chacune associée à un pixel, ou site, du champ) muni d'un système de voisinage. Ce champ est dit markovien si la loi de la variable en un site quelconque conditionnellement au reste des sites du champ est égale à la loi de la variable en ce même site conditionnellement à l'ensemble de ses voisins. Les interactions sont ainsi considérées à la fois localement et globalement par transitivité. Pour classifier des données hyperspectrales, une première possibilité consiste à réaliser une pré-classification vectorielle afin de travailler dans un second temps avec un modèle markovien scalaire. Ce fut notamment le choix de Jackson et al. [22] (maximum de vraisemblance), Bovolo et al. [6], Tarabalka et al. [52] et Moser et al. [34] (SVM). Une autre manière d'approcher le problème consiste à remplacer les valeurs scalaires originellement utilisées dans le modèle par des vecteurs. Rellier [41] introduit ainsi un modèle markovien gaussien multivarié adapté aux données hyperspectrales, les interactions entre bandes adjacentes n'étant pas modélisées directement mais découlant des paramètres du modèle. Panjwani et al. [37] proposent un modèle similaire dans le cadre d'une segmentation de données RGB, tandis que Kato et al. [24] y ajoutent des composantes texturales réalisées à partir de filtres de Gabor. Zhong et al. [58] ont pour leur part utilisé un modèle de *Conditional Random Field*, intéressant pour plusieurs raisons que nous listerons dans le chapitre 3. Un tel modèle implique cependant l'utilisation d'un nombre élevé de paramètres (quelques milliers) qu'il est possible de diminuer par réduction de dimension préalable.

Enfin, une dernière approche consiste à tirer parti de l'information de forme

contenue dans les images par l'intermédiaire d'une carte de segmentation réalisée préalablement. La segmentation est un procédé visant à partitionner les données en régions afin d'en simplifier l'interprétation. Une région est un ensemble de pixels voisins similaires, cette notion de similarité pouvant être quantifiée à l'aide de diverses caractéristiques telles que la signature spectrale ou l'aspect textural du voisinage. Tarabalka et al. se sont beaucoup intéressés à la question dans leurs travaux. Dans [48], [51] et [50] le résultat d'une classification par SVM est régularisé au moyen d'une carte de segmentation et d'un vote majoritaire. Cette carte de segmentation est respectivement obtenue au moyen d'algorithmes de clustering (ISODATA, mélange de gaussiennes), d'une segmentation par lignes de partage des eaux et d'une segmentation par forêt couvrante de poids minimal (*Minimum Spanning Forest*). Dans ce dernier cas, une approche par classifieurs multiples [49] a également été réalisée afin d'augmenter la pertinence de la segmentation. Cependant, les approches de segmentation utilisées ne tiennent pas compte de l'information de texture des zones, ce qui peut aboutir sur un résultat sur-segmenté. De plus, la carte de segmentation reste un outil d'appoint permettant de régulariser une image classifiée pixel à pixel. Nous proposons ainsi un nouveau processus de classification par zones basé sur le calcul d'une carte de segmentation tenant compte de l'information de texture.

5 Objectifs de la thèse et jeux de données

Ce travail a pour but d'évaluer l'apport de l'information spatiale à l'imagerie hyperspectrale dans divers contextes. Durant notre étude, nous avons ainsi considéré plusieurs cas de figure. Nous nous intéressons tout d'abord au cas où les classes présentent une information spatiale à très courte portée, de l'ordre du pixel. Dans un tel contexte, nous avons choisi de privilégier une approche markovienne car son modèle se base sur des relations contextuelles entre pixels très proches les uns des autres. Dans le second cas de figure, les classes présentent une information spatiale à plus longue portée et sont constituées de motifs texturaux bien visibles. Pour traiter de telles données, nous avons cette fois opté pour une classification vectorielle de diverses caractéristiques spectrales et spatiales, ces dernières reflétant l'information spatiale sur de plus grands voisinages (de l'ordre de la dizaine de pixels). En ce qui concerne le dernier cas de figure, nous considérons en plus l'hypothèse que les pixels appartenant à la même classe sont répartis dans l'image en groupes connexes. Nous introduisons ainsi une troisième approche basée sur la segmentation préalable de l'image afin d'identifier ces groupes et de les utiliser dans le cadre d'une classification par zones.

Notre but est d'étudier l'aptitude de ces trois approches à s'adapter à différentes manifestations de l'information spatiale dans les images. Pour ce faire, nous les avons appliquées à quatre jeux de données choisis de sorte que chacun relève d'un des cas de figure définis dans le paragraphe précédent. Un jeu de don-

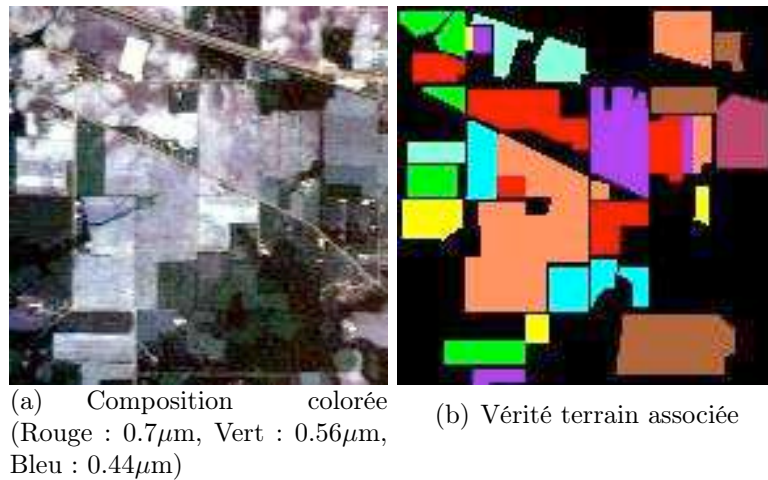


FIG. 1.16: Image IP

nées utilisable pour la validation et l'intercomparaison de plusieurs méthodes de classification nécessite d'être associée à une vérité terrain très complète, adaptée aux objectifs de la classification et suffisamment complexe pour que la validation soit pertinente. Dans le domaine de l'imagerie hyperspectrale, de tels jeux de données sont rares.

Notre premier jeu de données, très utilisé dans la littérature [16, 57, 58], est l'image *Indian Pine*, que nous désignerons par IP. C'est un extrait de 145×145 pixels d'une image AVIRIS réalisée sur le site *Indian Pine*, de 20 mètres de résolution spatiale et constituée de 220 bandes spectrales couvrant un intervalle de longueurs d'onde allant de $0.37\mu\text{m}$ à $2.5\mu\text{m}$ (voir figure 1.16(a)). Certaines bandes spectrales, fortement bruitées (les bandes 1, 219 et 220, ainsi que les bandes [104–109] et [150–163] correspondant aux zones d'absorption de la vapeur d'eau) et très faiblement corrélées aux autres, ont été écartées afin de ne pas perturber les traitements. Au final, nous utilisons donc un total de 197 bandes spectrales. L'extrait est associé à une vérité terrain composée de seize classes, dont sept sont ignorées en raison de leur faible nombre d'échantillons. Les neuf classes restantes sont réparties sur un ensemble de 9345 échantillons (voir figure 1.16(b)). Ce sont pour la plupart des classes agricoles de type blé, soja ou herbe rase présentant très peu d'information spatiale. L'image est ainsi majoritairement composée de parcelles agricoles et forestières séparées par des structures linéaires représentant des routes ou des chemins. La taille des parcelles étant en moyenne assez faible (rarement plus d'une trentaine de pixels de large), la considération de vastes voisinages ne constitue a priori pas l'approche la plus efficace. En effet, une fenêtre comportant assez d'échantillons pour définir correctement une texture couvrirait très certainement une zone constituée de plusieurs classes différentes. L'image semble par conséquent plus adaptée à l'usage d'une approche markovienne.

Les jeux de données MR et MR2 ont pour leur part été synthétisés à partir

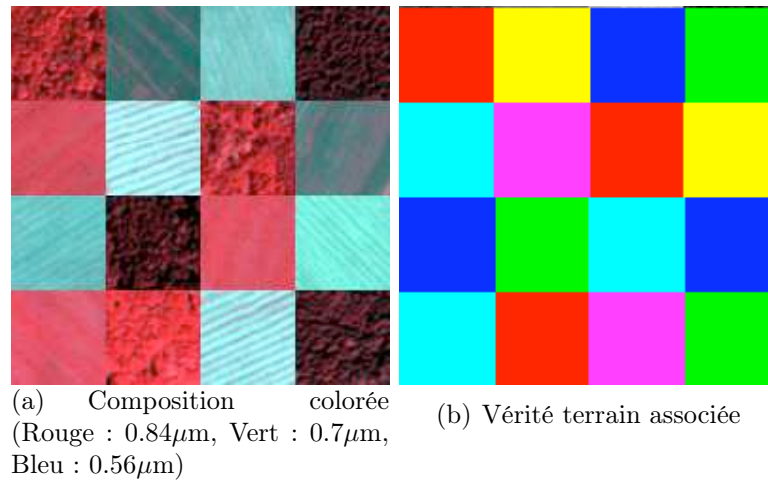


FIG. 1.17: Image MR

d'une image ROSIS de la région d'Oberpfaffenhofen, en Allemagne. Cette image - fournie par le DLR - est caractérisée par une résolution spatiale de 1.6 mètres et une couverture spectrale de 103 bandes (dont 91 sont conservées, les bandes [1 – 12] étant fortement bruitées) allant du visible au proche infrarouge (430nm à 837nm). De nombreuses zones ont été extraites de l'image originale puis juxtaposées les unes aux autres de sorte à former respectivement une image de 200×200 et 300×240 pixels (voir figure 1.17(a) et 1.18(a)). La seule différence entre MR et MR2 réside dans la forme des zones. Dans le cas de MR, celles-ci sont toutes des carrés de 50×50 pixels, tandis que la seconde image est constituée de zones de forme et de taille variées, afin de représenter avec plus de réalisme les frontières entre classes. L'ensemble des pixels d'une zone appartient à l'une des six classes considérées (2 types de forêt et 4 types de culture). La structure de l'image encourage aussi bien l'utilisation d'une approche de segmentation que d'une approche de classification vectorielle d'indices spectraux/spatiaux. En effet, la grande taille des zones permet d'étudier de vastes voisinages sans forcément risquer de mélanger plusieurs textures différentes.

Nous avons enfin synthétisé un quatrième jeu de données, désigné par CO, à partir d'une image HyMap représentant la ville de Colmar et ses environs. L'image a une résolution spatiale de 3 mètres et est constituée de 128 bandes spectrales (dont 126 sont conservées) couvrant le même intervalle de longueur d'onde que l'image précédente. Trois zones de 60×200 pixels ont été prélevées sur cette image et fusionnées de sorte à former une image de 180×200 pixels (voir figure 1.19(a)). Nous considérons ici trois types de sol, une classe forestière, une classe urbaine et une classe rurale, chacune étant associée à l'une des zones extraites de l'image originale (voir figure 1.19(b)). Ce sont des classes sémantiquement plus complexes dans le sens où elles ne sont pas uniquement liées à un matériau précis mais à un type d'aménagement du territoire (forêt, ville ou campagne). Ce type de classe

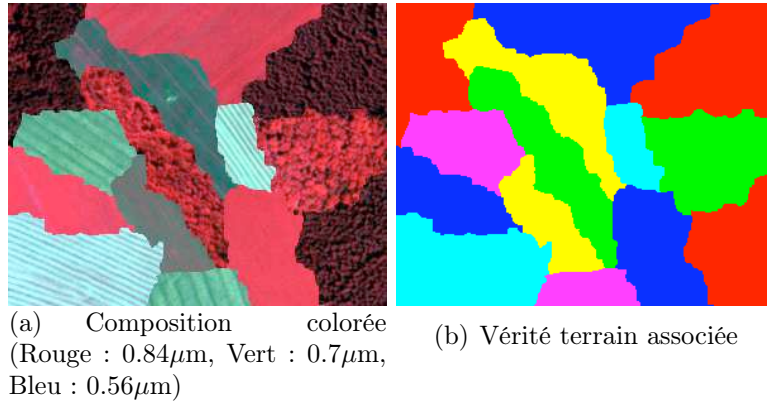


FIG. 1.18: Image MR2

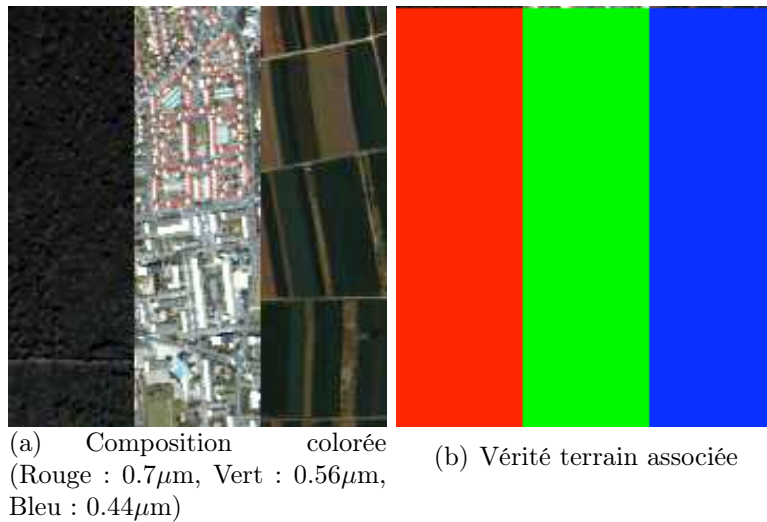


FIG. 1.19: Image CO

s'étend généralement sur des ensembles de pixels assez grands. De ce fait, les deux dernières approches sont là encore préférables à l'approche markovienne.

6 Conclusion

Dans ce chapitre, nous avons commencé par présenter succinctement le contexte de l'étude : la complémentarité des informations spectrale et spatiale pour la classification d'images hyperspectrales. Le passage en revue des méthodes de la littérature s'intéressant également à cette complémentarité nous a permis de lister plusieurs approches de classification spectrale/spatiale, chacune disposant de forces et de faiblesses relatives à la nature de l'information spatiale reflétée par les classes étudiées. A partir des quatre jeux de données à notre disposition, nous avons ainsi défini trois scénarios de classification que nous étudierons et compare-

rons dans la suite de ce manuscrit. Le premier, détaillé dans le chapitre 2, consiste à classifier de manière vectorielle une série de caractéristiques spectrales et spatiales calculées au préalable. Le second scénario, décrit dans le chapitre 3, est une classification par Maximum A Posteriori basée sur un modèle de type *Conditional Random Field* amélioré. Enfin le dernier scénario s'inspire des travaux de Deng et al. [12] sur la segmentation d'images multispectrales pour proposer une méthode de classification par zones et non plus pixel à pixel.

Chapitre 2

Classification vectorielle de caractéristiques spectrales et spatiales

1 Introduction

Dans ce chapitre, nous nous intéressons à l'extraction de caractéristiques spectrales et spatiales pour la classification vectorielle d'images hyperspectrales. Nous commençons par évoquer les divers procédés de réduction de dimension que nous avons considérés pour l'extraction de caractéristiques spectrales. Ces dernières constituent la base à partir de laquelle sont créées les caractéristiques spatiales tels que les spectres de texture ou les coefficients de Haralick que nous détaillons dans un second temps.

Enfin, le processus de classification vectorielle permettant l'utilisation simultanée de ces deux types de caractéristique est décrit dans une troisième partie. Les résultats obtenus sur les quatre jeux de données mentionnés dans le chapitre précédent, dans un contexte supervisé d'une part et non supervisé d'autre part, sont présentés en fin de chapitre.

2 Méthodes de réduction de dimension pour l'extraction de caractéristiques spectrales

Nous avons vu dans la section 2.2.1 du chapitre 1 qu'il était souhaitable de réduire la dimension de l'espace de travail afin de s'affranchir des effets du phénomène de Hughes. Dans un contexte hyperspectral où les bandes adjacentes comportent souvent de fortes similarités, l'objectif est d'extraire un nombre limité de variables reflétant la totalité de l'information utile tout en supprimant l'information redondante. Nous considérons ainsi ces variables comme autant d'indices

mettant en valeur l'information spectrale. Dans cette section, nous évoquons les divers algorithmes de réduction de dimension que nous avons utilisés.

2.1 Analyse en Composantes Principales

L'analyse en composantes principales (ACP), également connue sous le nom de transformée de Karhunen-Loève, vise à réduire la dimension d'une image multidimensionnelle en identifiant les combinaisons linéaires (appelées composantes principales) des bandes originales maximisant la variance (cf section 2.2.1 du chapitre 1). Soit \mathbf{X} une matrice de taille $M \times N$, N représentant le nombre de pixels de l'image et M le nombre de bandes. Chaque ligne de \mathbf{X} représente une bande de l'image que l'on a préalablement centrée. Considérons alors la matrice de covariance de ces données :

$$\Sigma = \frac{1}{N} \mathbf{X} \mathbf{X}^T \quad (2.1)$$

Cette matrice peut être décomposée comme suit :

$$\Sigma = \mathbf{U} \Lambda \mathbf{U}^T \quad (2.2)$$

où Λ est la matrice diagonale des valeurs propres $\{\lambda_1, \dots, \lambda_M\}$ de Σ rangées par ordre décroissant et \mathbf{U} une matrice $M \times M$ orthogonale contenant les vecteurs propres de Σ . Les composantes principales sont finalement obtenues en projetant l'image originale sur la base formée par ces vecteurs propres :

$$\mathbf{Y} = \mathbf{U}^T \mathbf{X} \quad (2.3)$$

Ces derniers forment une base orthogonale de l'espace de départ, sont décorrélés les uns des autres et ordonnés selon la proportion de variance du nuage de point qu'ils caractérisent.

L'ACP étant une analyse purement spectrale, le contexte spatial des pixels n'intervient pas. Pour prendre en compte le voisinage de chaque pixel dans le calcul de la covariance, nous proposons une variante de l'algorithme où les composantes principales ne sont plus calculées à partir de la matrice de covariance de l'image mais sur une matrice de "pseudo-covariance" :

$$\tilde{\Sigma} = \frac{1}{N} \mathbf{X} \tilde{\mathbf{X}}^T \quad (2.4)$$

où $\tilde{\mathbf{X}}$ correspond à l'image originale convoluée par un filtre moyenneur. L'objectif est de privilégier les pixels issus de zones homogènes dans le calcul de la variance afin de discriminer au mieux ces dernières sur les composantes principales. Ce procédé permet de mettre en valeur les classes non texturées.

Dans le cas d'images volumineuses, l'ACP peut être coûteuse en temps de calcul car la matrice de covariance est calculée sur la totalité de l'image. Elle

garantit cependant une perte d'énergie du signal minimale, ce qui représente une des mesures non supervisées de l'information les plus efficaces. De plus la réduction de dimension de l'image permet en contrepartie d'améliorer la complexité de l'étape de classification finale.

2.2 Minimum Noise Fraction

Très semblable à l'ACP, ce procédé de réduction de dimension identifie les combinaisons linéaires des bandes originales de l'image qui maximisent non plus la variance mais le rapport signal sur bruit. Les données \mathbf{X} sont tout d'abord décomposées de la sorte :

$$\mathbf{X} = \mathbf{S} + \mathbf{N} \quad (2.5)$$

où \mathbf{S} et \mathbf{N} sont respectivement les composantes (non corrélées) de signal et de bruit de \mathbf{X} . Cette hypothèse de non corrélation nous permet d'écrire :

$$\Sigma = \Sigma_{\mathbf{S}} + \Sigma_{\mathbf{N}} \quad (2.6)$$

tel que $\Sigma_{\mathbf{S}}$ et $\Sigma_{\mathbf{N}}$ sont respectivement les matrices de covariance du signal et du bruit. La proportion de bruit NF_i d'une bande $i \in \{1, \dots, P\}$ est ainsi définie par :

$$NF_i = \frac{Var\{\mathbf{N}_i\}}{Var\{\mathbf{X}_i\}} \quad (2.7)$$

Si l'on considère que le rapport signal sur bruit d'une bande i est :

$$SNR_i = \frac{Var\{\mathbf{S}_i\}}{Var\{\mathbf{N}_i\}} \quad (2.8)$$

alors

$$SNR_i = \frac{1}{NF_i} - 1 \quad (2.9)$$

et maximiser le rapport signal sur bruit revient à maximiser $\frac{1}{NF_i}$. On recherche ainsi un vecteur directeur u de l'espace tel que :

$$\frac{u^T \Sigma u}{u^T \Sigma_{\mathbf{N}} u} \quad (2.10)$$

soit maximal, ce qui revient à résoudre l'équation suivante :

$$\Sigma \Sigma_{\mathbf{N}}^{-1} u = \lambda u \quad (2.11)$$

Les vecteurs $\{u_i\}_{i \in \{1, \dots, P\}}$ solutions de l'équation sont les vecteurs propres de la matrice $\Sigma \Sigma_{\mathbf{N}}^{-1}$, la valeur propre λ_i associée correspondant au rapport signal sur bruit dans la direction indiquée par u_i . Par convention, ces valeurs propres sont arrangées par ordre décroissant de sorte à ce que les composantes MNF soient également arrangées par ordre décroissant de rapport signal sur bruit.

Cette méthode est un pendant intéressant de l'ACP car le rapport signal sur bruit est souvent considéré comme une mesure plus pertinente de la qualité de l'information présente dans une bande que la variance. Le fait de devoir disposer d'une connaissance a priori sur le bruit de l'image, ce qui est assez délicat si l'on ne connaît pas les caractéristiques du capteur et les conditions d'acquisition des données, est toutefois problématique. La matrice de covariance du bruit est souvent calculée sur la zone de l'image considérée comme la plus homogène et où les variations du signal sont uniquement dues au bruit mais en pratique, il existe rarement dans une image une zone parfaitement homogène permettant de faire une bonne estimation du bruit.

2.3 Analyse discriminante descriptive

L'analyse discriminante descriptive est une méthode statistique supervisée qui, à partir d'un jeu de N observations décrites par M variables et réparties en K groupes, vise à déterminer l'espace de représentation des données qui séparera au mieux les K groupes. Cette méthode est essentiellement géométrique et ne repose sur aucune hypothèse probabiliste.

Soit $n_i, i \in \{1, \dots, K\}$ le nombre d'échantillons appartenant au groupe i . μ_i correspond au centroïde des échantillons du groupe i et W_i est la matrice de covariance de ce dernier. L'objectif de l'algorithme est de déterminer une série d'axes non corrélés deux à deux qui discriminent au mieux les individus de groupes différents tout en rapprochant les individus de même groupe. Pour cela, nous considérons deux matrices de dispersion. La matrice de dispersion intra-classe :

$$W = \frac{1}{N} \sum_{i=1}^K n_i W_i \quad (2.12)$$

correspond à la moyenne pondérée des matrices de covariance de chaque classe et représente la dispersion moyenne des individus d'une classe autour de leur centroïde. La matrice de dispersion inter-classe :

$$B = \frac{1}{N} \sum_{i=1}^K n_i (\mu_i - \mu)^T (\mu_i - \mu) \quad (2.13)$$

où $\mu = \frac{1}{N} \sum_{i=1}^K n_i \mu_i$ est le centroïde de l'ensemble des échantillons, représente la dispersion moyenne entre centroïdes de classes différentes.

On recherche ainsi u , un vecteur directeur de l'espace tel que $u^T B u$ soit maximal et $u^T W u$ soit minimal. Cela consiste à maximiser $\frac{u^T B u}{u^T V u}$, où $V = B + W$ est la matrice de dispersion totale, ce qui revient à résoudre l'équation suivante :

$$V^{-1} B u = \lambda u \quad (2.14)$$

Au final, les axes recherchés sont ainsi les vecteurs propres correspondant aux valeurs propres non-nulles de la matrice $V^{-1}B$. Dans le cas où $N > M > K$, il existe $K - 1$ axes discriminants.

Cette méthode a l'avantage d'être rapide et de tenir compte des classes de l'image dans sa recherche d'axes discriminants. Cependant, elle nécessite de ce fait la connaissance d'un ensemble d'échantillons pour chaque classe et peut être prise en défaut si les classes ne sont pas assez uniformément réparties dans l'espace initial, notamment lorsque deux classes ont une moyenne très proche.

2.4 Minimum Change Rate Deviation (MCRD)

Considérons f , une fonction non linéaire de réordonnement du spectre de chaque pixel par valeur de réflectance décroissante (voir figure 2.1) telle que $\mathbf{D} = f(\mathbf{X})$, où X est le jeu de données multivariées de départ. Grâce à cette réorganisation des spectres, il est plus facile de les approcher avec des méthodes d'optimisation simple (type moindre carré, par exemple). La méthode MCRD [13] est une méthode itérative et paramétrique qui, tout en conservant les propriétés de décorrélation des projetés et d'orthogonalité des projecteurs de l'ACP, recherche les combinaisons linéaires de bandes permettant de reconstruire le plus fidèlement le jeu de données \mathbf{D} . Ces combinaisons sont représentées par les vecteurs $(\mathbf{u}_i)_{i \in [1, \dots, M]}$ que l'on détermine par optimisation de l'équation quadratique suivante :

$$I_{\mathbf{X}}(\mathbf{u}_i) = (\mathbf{D}_i - \mathbf{X}\mathbf{u}_i)^T(\mathbf{D}_i - \mathbf{X}\mathbf{u}_i) \quad (2.15)$$

où \mathbf{D}_i est la i -ème bande de \mathbf{D} . A chaque itération k de la méthode, on cherche ainsi le vecteur \mathbf{u}_k orthogonal aux vecteurs $(\mathbf{u}_j)_{j \in [1, \dots, k-1]}$ et tel que $\mathbf{X}\mathbf{u}_k$ et $(\mathbf{X}\mathbf{u}_j)_{j \in [1, \dots, k-1]}$ soient non corrélés. On choisit finalement de retenir K vecteurs de projection répartis de façon homogène sur l'intervalle $[1, \dots, m]$ de sorte à approcher le plus précisément possible les spectres de \mathbf{D} .

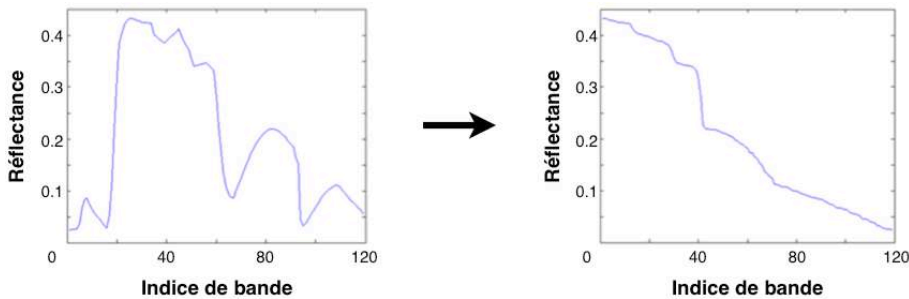


FIG. 2.1: Réordonnement décroissant des valeurs de réflectance du spectre

2.5 Sélection du nombre de caractéristiques à conserver

Les algorithmes de réduction de dimension cités précédemment produisent pour la plupart un nombre de composantes égal au nombre de bandes de l'image originale. Afin de travailler avec un nombre de variables raisonnable, il est nécessaire de ne retenir qu'un petit nombre de ces composantes. Pour l'analyse discriminante, le nombre d'axes discriminants correspond au nombre de classes de l'image moins un. Pour la méthode MCRD, il dépend de la précision avec laquelle on désire approximer le nouveau cube hyperspectral. Concernant l'ACP et la MNF, les composantes sont ordonnées en fonction de la part de variance expliquée par les composantes. Reste à définir le nombre minimal de composantes qu'il faut conserver tout en gardant un niveau informatif quasiment aussi élevé que les données originales. Pour calculer ce nombre, nous proposons plusieurs méthodes : le seuillage, le Scree-test de Cattell et le rapport de croissance.

2.5.1 Seuillage

Cette méthode utilise le fait que la variance expliquée par chaque composante est proportionnelle à la valeur propre associée à son vecteur de projection.

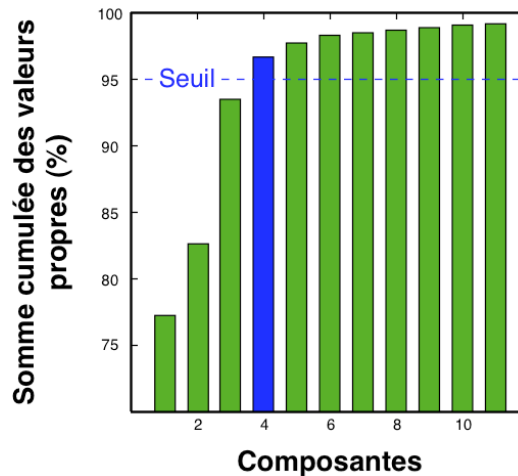


FIG. 2.2: Seuil de pourcentage des valeurs propres cumulées : pour un seuil de 95%, les quatre premières composantes sont conservées

Elle consiste à fixer un seuil d'information expliqué et de choisir le plus petit nombre de bandes pour lequel ce seuil est dépassé. Généralement, on conserve les composantes permettant d'atteindre un certain pourcentage de la somme cumulée des valeurs propres (voir figure 2.2), ou bien celles dont la valeur propre est associée est supérieure à un certain seuil.

2.5.2 Scree-test de Cattell

Cette méthode se base sur une approche purement visuelle de la décroissance des valeurs propres. Lorsqu'un coude se forme, on choisit de ne conserver que les bandes se trouvant en amont (voir figure 2.3).

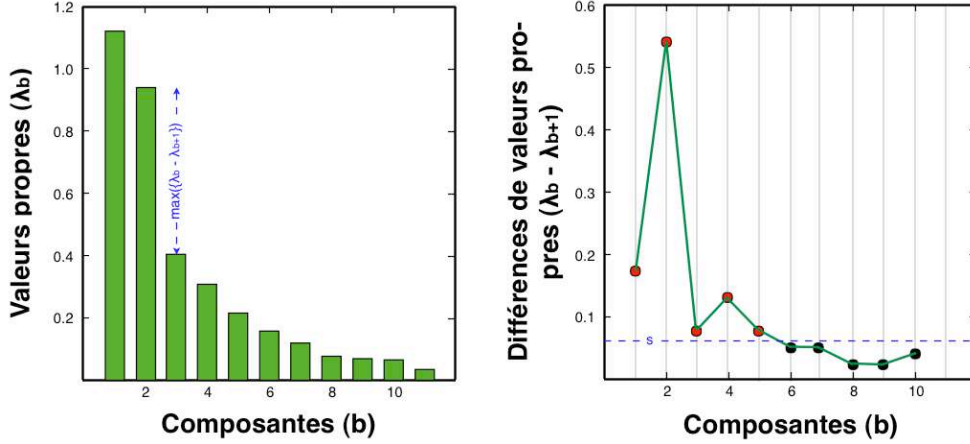


FIG. 2.3: Méthode du coude : exemple avec $\alpha = 10\%$. Les cinq premières composantes sont conservées

Pour détecter automatiquement ce coude, considérons un seuil α généralement fixé entre 10% et 20% et :

$$s = \alpha \max(\lambda_1 - \lambda_2, \dots, \lambda_{k-1} - \lambda_k, \dots, \lambda_{M-1} - \lambda_M) \quad (2.16)$$

La plus petite composante l pour laquelle $\lambda_{l-1} - \lambda_l < s$ constituera alors la frontière à partir de laquelle on ne conserve plus les composantes.

2.5.3 Rapport de croissance

Le principal inconvénient des deux méthodes précédentes est la nécessité pour l'utilisateur de fixer un paramètre a priori. Ce paramètre (la valeur de seuil de la première méthode, notamment) influe non seulement sur le choix du nombre de bandes conservées mais également indirectement sur les résultats du traitement (classification, détection d'anomalies, etc...) effectué ensuite. Ahn et al. [1] proposent un estimateur entièrement automatique appelé rapport de croissance et basé sur l'information restante si on enlève les composantes les unes après les autres. Testé sur des données macro-économiques et financières, cet estimateur s'est révélé être un moyen puissant de détecter les composantes faiblement informatives. Considérons λ_k la valeur propre associée à la k -ième composante de l'image réduite et :

$$\lambda_k^* = \frac{\sum_{i=k}^M \lambda_i}{\sum_{i=k+1}^M \lambda_i} \quad (2.17)$$

Soit RC (rapport de croissance) un critère d'entropie correspondant au rapport des logarithmes de deux λ^* consécutifs :

$$RC(k) = \frac{\log \lambda_k^*}{\log \lambda_{k+1}^*} \quad (2.18)$$

Le nombre de bandes à conserver est la valeur k pour laquelle RC est maximal.

3 Extraction d'indices spatiaux

Un indice (ou caractéristique) spatial correspond à une mesure du comportement spatial du voisinage d'un pixel. Cette mesure peut être issue de traitements statistiques, morphologiques ou purement radiométriques et, calculée sur l'ensemble des pixels d'une image, elle permet de créer de nouveaux attributs caractérisant non plus le comportement spectral d'une scène mais son comportement spatial local. Nous ferons la différence entre deux types d'indices spatiaux, les indices morphologiques d'une part, produits à partir d'opérateurs de morphologie mathématique, et les indices texturaux d'autre part. Bien que largement étudiée, il est assez difficile de trouver une définition précise et communément acceptée de la notion de texture. Sklansky [43] suggère qu'une région d'image est caractérisée par une texture constante lorsqu'un ensemble de statistiques ou de propriétés locales de la fonction image sont constantes, varient faiblement ou bien de manière approximativement périodique dans l'espace. Dunn et al. [14] partent du principe qu'une texture peut être définie comme une collection d'objets primitifs similaires - c'est-à-dire qu'ils ne sont pas nécessairement identiques - agencés selon un motif répétitif. Wang et al. [55] avancent de leur côté que la texture est un terme permettant de caractériser les variations de tons d'une image. Beaucoup d'autres définitions ont toutefois été proposées [35, 39, 40].

Dans cette partie nous présentons successivement des indices texturaux extraits à partir de matrices de co-occurrence ou de spectres de texture, puis nous finirons par la présentation des profils morphologiques.

3.1 Matrices de co-occurrence

Dans le contexte de l'analyse texturale, les matrices de cooccurrence ont été introduites par Haralick [17] et représentent un moyen simple et très utilisé d'étudier l'information de texture dans une image. Ces matrices décrivent l'ensemble des variations radiométriques associées au voisinage d'un pixel, c'est-à-dire qu'elles recensent la proportion de pixels passant d'une valeur radiométrique à une autre pour un déplacement de a colonnes et b lignes au sein de ce voisinage.

Considérons le pixel (x, y) d'une image mono-canal I prenant ses valeurs dans l'ensemble $G = \{0, \dots, N_g - 1\}$ où N_g est le nombre de niveaux de gris. Soit $V_{x,y}^r$

le voisinage de rayon r associé au pixel (x, y) :

$$V_{x,y}^r = \begin{bmatrix} I(x-r, y-r) & \cdots & I(x, y-r) & \cdots & I(x+r, y-r) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ I(x-r, y) & \cdots & I(x, y) & \cdots & I(x+r, y) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ I(x-r, y+r) & \cdots & I(x, y+r) & \cdots & I(x+r, y+r) \end{bmatrix} \quad (2.19)$$

La matrice de co-occurrence $\mathbf{C}_{a,b}$ associée à ce voisinage est une matrice de taille $N_g \times N_g$ représentant une carte d'identité locale des variations spectrales pour un décalage spatial (a, b) donné. Afin de limiter la taille des matrices de co-occurrence manipulées, l'image est tout d'abord quantifiée de sorte à obtenir un N_g assez faible tel qu'il permet tout de même de refléter efficacement le comportement textural local de l'image. La figure 2.4 donne deux exemples de matrices de co-occurrence associées à un même voisinage. Par exemple, l'élément $\mathbf{C}_{1,0}(1, 2)$ de la matrice de co-occurrence 2.4(b) est égal au nombre de couples $(V_{x,y}^r(i_v, j_v), V_{x,y}^r(i_v + 1, j_v))$ tels que $V_{x,y}^r(i_v, j_v) = 1$ et $V_{x,y}^r(i_v + 1, j_v) = 2$, c'est-à-dire 2.

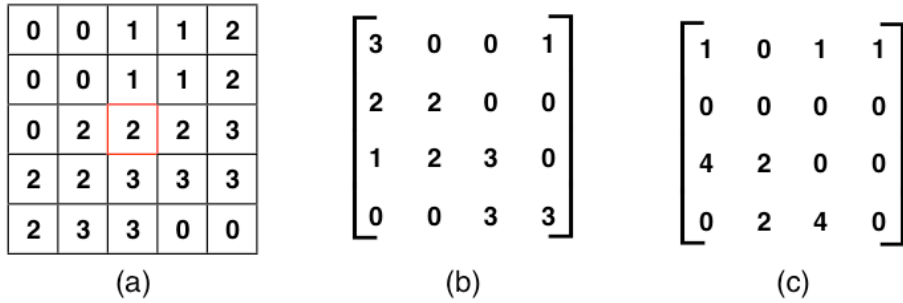


FIG. 2.4: (a) Voisinage 5×5 à 4 niveaux de gris et matrices de cooccurrence correspondantes pour un déplacement (b) $(1, 0)$ et (c) $(0, 2)$

Quelle que soit sa taille, une matrice de cooccurrence représente une masse de données conséquente qu'il n'est pas aisé de manipuler. C'est pourquoi Haralick a élaboré un certain nombre d'indices calculés à partir d'une matrice de co-occurrence. Quatorze indices sont proposés dans [17], chacun synthétisant une caractéristique texturale spécifique, telles que l'homogénéité ou le contraste, tandis que d'autres se focalisent sur la complexité et la nature des transitions entre niveaux de gris. Parmi ces indices, Cossu [9] considère l'énergie (ou second moment angulaire), le contraste, la variance, la corrélation, l'entropie et le moment de différence inverse (ou homogénéité) comme les plus pertinents, les autres étant soit moins riches en information, soit fortement corrélés à ces six coefficients. Soient p_x le vecteur des sommes de lignes de $C_{a,b}$ et p_y le vecteur des sommes de

colonnes de $C_{a,b}$, μ_x , μ_y , σ_x et σ_y étant respectivement les moyennes et écarts-type de p_x et p_y . Ces coefficients sont définis par :

- Energie = $\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} C_{a,b}(i, j)^2$
- Contraste = $\sum_{n=0}^{N_g-1} n^2 \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} C_{a,b}(i, j), |i - j| = n$
- Variance = $\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - \mu)^2 C_{a,b}(i, j)$
- Entropie = $-\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} C_{a,b}(i, j) \log C_{a,b}(i, j)$
- Moment de différence inverse = $\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{1}{1+(i-j)^2} C_{a,b}(i, j)$
- Corrélacion = $\frac{1}{\sigma_x \sigma_y} (\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{ij C_{a,b}(i, j)\} - \mu_x \mu_y)$

La figure 2.5 présente des exemples de ces coefficients obtenus à partir d'une image en niveaux de gris. L'énergie, également appelée second moment angulaire, mesure l'uniformité d'une texture. Si un voisinage est homogène spatialement, quelques valeurs de $C_{a,b}$ seront fortes tandis que toutes les autres seront très faibles, auquel cas l'énergie sera proche de son maximum. Le contraste, comme son nom l'indique, mesure le niveau de contraste d'un voisinage, c'est-à-dire la différence entre les valeurs les plus élevées et les valeurs les plus faibles. Un voisinage faiblement contrasté se caractérise par une matrice de co-occurrence dont les fortes valeurs se rassemblent près de la diagonale tandis que pour un voisinage fortement contrasté, les fortes valeurs ont plutôt tendance à s'en éloigner.

La variance mesure quant à elle l'hétérogénéité d'un voisinage ; elle est d'autant plus forte que les valeurs de ce dernier s'éloignent de leur moyenne. Variance et contraste ne sont pas complètement dépendants l'un de l'autre car si une variance faible implique un contraste également faible, l'inverse n'est pas toujours vrai. L'entropie constitue une mesure du désordre associé à un voisinage. Lorsque ce dernier n'est pas uniforme d'un point de vue textural, la plupart des éléments de $C_{a,b}$ ont une valeur faible, ce qui résulte sur une forte valeur d'entropie. La corrélation est une mesure des dépendances linéaires de niveaux de gris d'un voisinage. Cette mesure est forte lorsqu'il existe une relation linéaire entre les niveaux de gris des paires de pixels. Elle n'est pas dépendante des répétitions de paires de pixels et est donc décorrélée de l'énergie et de l'entropie. Enfin, le moment de différence inverse mesure l'homogénéité d'un voisinage tout en privilégiant les faibles différences de niveaux de gris dans les paires de pixels. Cette dernière mesure est inversement proportionnelle à l'énergie et au contraste à la fois.

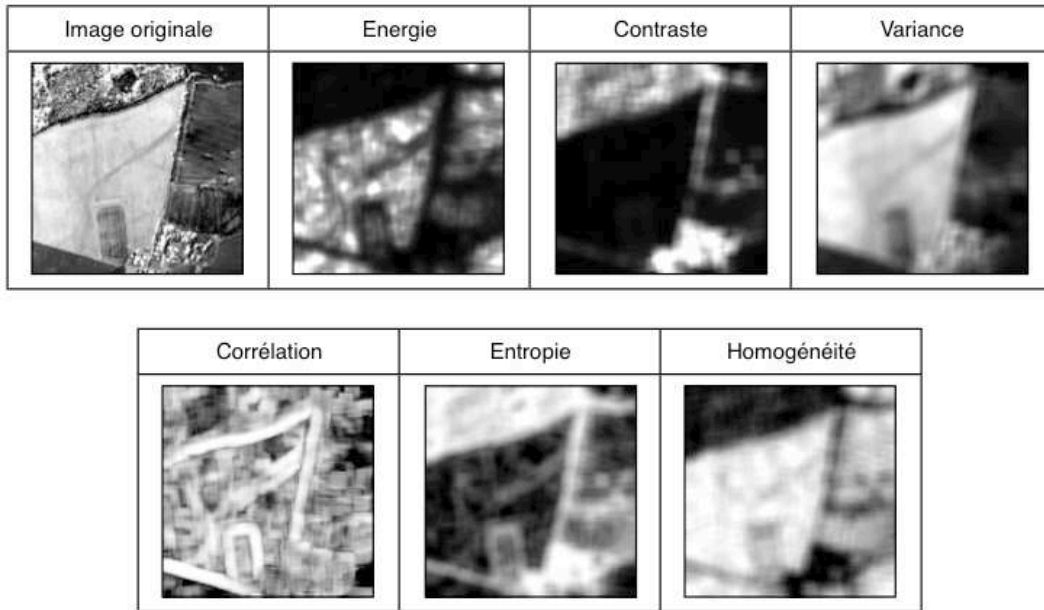


FIG. 2.5: Exemples de coefficients de Haralick

3.2 Spectres de texture

L'approche de modélisation de la texture par spectres de textures [19] repose sur le calcul d'une image d'unités de texture, une grandeur qui caractérise le plus localement possible la notion de texture autour d'un pixel en considérant les huit plus proches voisins. Le comportement textural local relatif à chaque pixel est identifié par le calcul d'un histogramme des unités de texture sur un certain voisinage.

Considérons le voisinage immédiat d'un pixel, c'est-à-dire la fenêtre 3×3 dont il est le pixel central (cf. figure 2.6(a)). Ce voisinage est défini par $L = \{L_0, L_1, \dots, L_8\}$, où L_0 est la valeur du pixel central et L_i les valeurs de ses voisins. L'unité de texture en elle-même est définie par un ensemble $TU = \{E_1, E_2, \dots, E_8\}$ tel que :

$$E_i = \begin{cases} 0 & \text{si } L_i < L_0 - \psi \\ 1 & \text{si } L_0 - \psi \leq L_i \leq L_0 + \psi \\ 2 & \text{si } L_i > L_0 + \psi \end{cases} \quad \forall i \in \{1, 2, \dots, 8\} \quad (2.20)$$

ψ étant un seuil fixé par l'utilisateur.

Une valeur scalaire N_{TU} est ensuite associée à l'unité de texture (cf. 2.6(b)) :

$$N_{TU} = \sum_{i=1}^8 3^{i-1} E_i, \quad N_{TU} \in \{0, 1, \dots, 6560\}$$

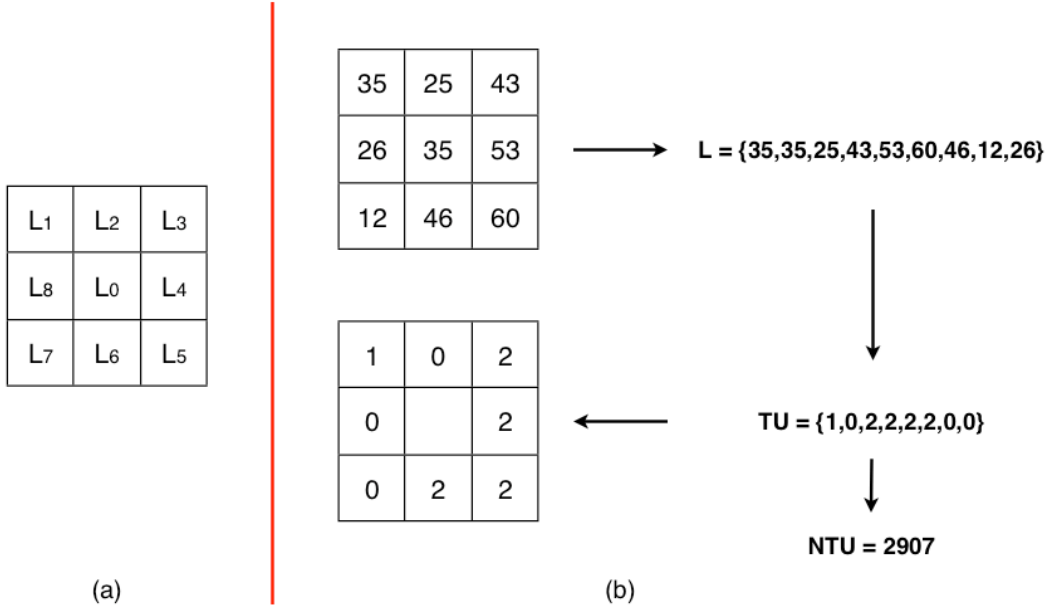


FIG. 2.6: A partir du voisinage immédiat d'un pixel et en utilisant l'ordre des aiguilles d'une montre, il existe 8 manières d'ordonner les éléments de façon à créer un nombre en base 3, en partant de chacun des pixels numérotés de L_1 à L_8 de la figure (a). A partir de cet ensemble de valeurs scalaires, la figure (b) présente comment une unité de texture en base 3 est produite (TU) puis convertie en base 10 (NTU).

L'histogramme des unités de texture d'une fenêtre $V_{x,y}^r$ est appelé spectre de texture (cf figure 2.7).

Il permet de considérer la texture à un niveau plus global, en plus du niveau local induit par l'unité de texture. Wang et al. [55] propose de dériver trois indices texturaux à partir de ces spectres de texture : BWS, GS et DD.

$$BWS = \left[1 - \frac{\sum_{i=0}^{3279} |S(i) - S(3281 + i)|}{\sum_{i=0}^{6560} S(i)} \right] \times 100$$

où $S(i)$ est le nombre d'occurrences de l'unité de texture i dans le spectre de texture. Cet indice appelé *Black-White Symmetry* mesure le degré de symétrie du spectre de texture.

$$GS = \left[1 - \frac{1}{4} \sum_{j=1}^4 \frac{\sum_{i=0}^{6560} |S_j(i) - S_{j+4}(i)|}{2 \times \sum_{i=0}^{6560} S_j(i)} \right] \times 100$$

où $S_j(i)$ est le nombre d'occurrences de l'unité de texture i dans le spectre de texture, selon la façon d'ordonner j (cf. figure 2.6a). Cette indice appelé *Geo-*

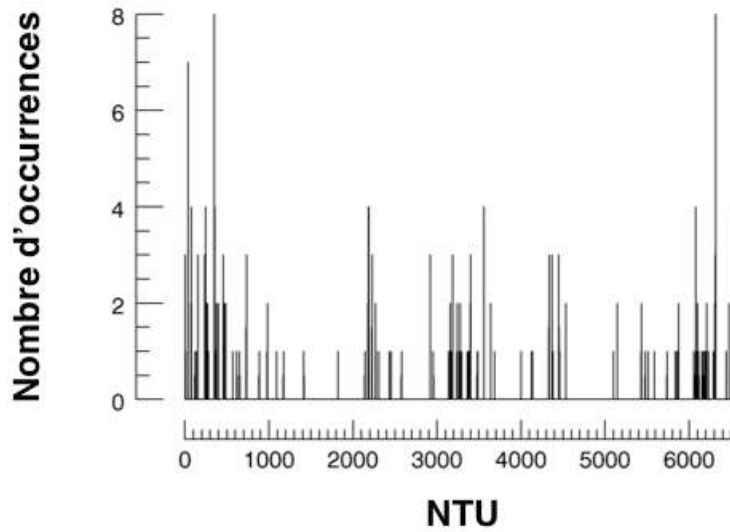


FIG. 2.7: Exemple de spectre de texture pour un voisinage de taille 31×31

metric Symmetry mesure le degré de symétrie d'une zone de l'image autour des quatre axes principaux (vertical, horizontal, diagonal et anti-diagonal).

$$DD = \left[1 - \frac{1}{6} \sum_{m=1}^3 \sum_{n=m+1}^4 \frac{\sum_{i=0}^{6560} |S_m(i) - S_n(i)|}{2 \times \sum_{i=0}^{6560} S_m(i)} \right] \times 100$$

Cet indice appelé *Degree of Direction* mesure la proportion de structures linéaires de la zone de l'image où le spectre de texture est calculé. Une valeur faible indiquera que le spectre de texture est sensible à l'orientation de l'image. La figure 2.8 présente un exemple de chacun de ces indices.

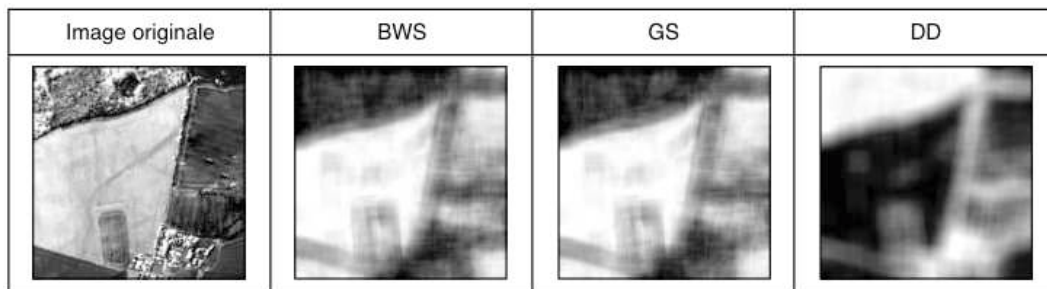


FIG. 2.8: Les trois indices spatiaux obtenus à partir des spectres de texture. (a) BWS est d'autant plus fort que le spectre de texture est symétrique. (b) GS mesure le degré de symétrie d'un certain voisinage tandis que (c) DD reflète son caractère linéaire.

3.3 Profils morphologiques

Cette approche, proposée initialement par Pesaresi et al. [38], repose sur l’emploi d’opérateurs de morphologie mathématique [44]. On utilise le concept de granulométrie, impliquant l’application sur l’image d’opérateurs morphologiques associés à des voisinages de taille croissante, afin de construire des attributs mettant en valeur des structures spatiales de plus en plus grandes à mesure que le voisinage grossit. Fauvel et al. [15] ont notamment utilisé ces opérateurs dans un contexte de classification supervisée d’images hyperspectrales.

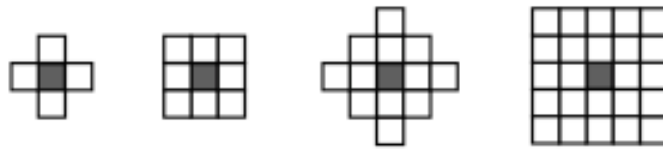


FIG. 2.9: Différents exemples d’élément structurant. Le pixel gris représente à chaque fois le centre de l’élément

L’érosion et la dilatation constituent les deux opérateurs fondamentaux en morphologie mathématique, à partir desquels sont dérivés la plupart des autres opérateurs. Ils sont appliqués à l’image au moyen d’un élément structurant, une fenêtre pouvant prendre diverses formes et tailles (voir figure 2.9), que l’on déplace sur toute l’image de sorte à ce que son centre passe par la totalité des pixels de cette dernière. Pour chaque pixel, on compare les valeurs prises par les voisins inclus dans le domaine de l’élément structurant, puis on lui affecte soit la valeur minimale dans le cas d’une érosion, soit la valeur maximale dans le cas d’une dilatation. Ces deux transformations, illustrées par la figure 2.10, tendent à homogénéiser l’image mais là où l’érosion l’assombrit en étalant les bords des objets les plus sombres, la dilatation l’éclaircit en étalant plutôt les bords des objets les plus clairs.

L’ouverture morphologique est une érosion suivie d’une dilatation et la fermeture morphologique une dilatation suivie d’une érosion. Ces deux opérations suppriment respectivement les objets clairs ou sombres dont la taille est inférieure à celle de l’élément structurant. Ces opérateurs ayant tendance à modifier les contours des objets, on leur préfère les ouvertures et fermetures par reconstruction. Concrètement, une ouverture par reconstruction est une érosion suivie de dilatations successives répétées jusqu’à idempotence, la fermeture par reconstruction étant l’opération duale. Ces opérateurs sont dits connexes et permettent de restituer complètement les objets respectivement clairs ou sombres s’ils n’ont pas été éliminés par le filtrage (voir figure 2.11).

Pour caractériser la forme et la taille de tous les objets d’une image, il est

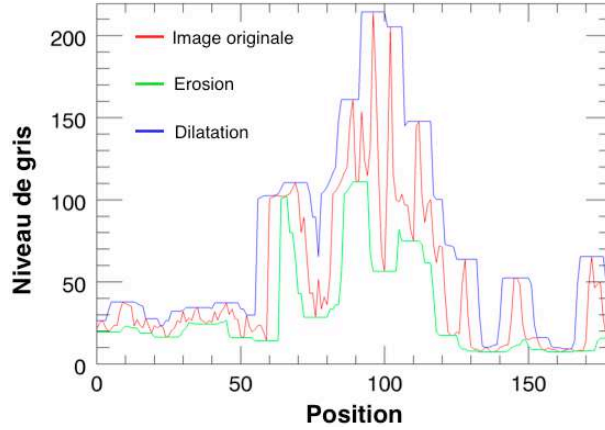


FIG. 2.10: Exemple d'érosion (en vert) et de dilatation (en bleu) d'un signal 1D représentant une ligne de pixels d'une image (en rouge).

nécessaire d'utiliser en parallèle des éléments structurants de différentes tailles. C'est ce que l'on appelle le concept de granulométrie : l'image est filtrée par des opérateurs morphologiques de taille croissante qui simplifient progressivement l'image. Soit \mathbf{X} une image monodimensionnelle, $\gamma_R^i(\mathbf{X})$ l'ouverture par reconstruction de \mathbf{X} par un élément structurant de taille $2i + 1$ et $\phi_R^i(\mathbf{X})$ la fermeture par reconstruction correspondante. Un profil morphologique est la concaténation bande à bande des images de l'ensemble suivant :

$$MP = \{\gamma_R^K(\mathbf{X}), \gamma_R^{K-1}(\mathbf{X}), \dots, \gamma_R^1(\mathbf{X}), \mathbf{X}, \phi_R^1(\mathbf{X}), \dots, \phi_R^{K-1}(\mathbf{X}), \phi_R^K(\mathbf{X})\} \quad (2.21)$$

A partir d'une image monodimensionnelle, on obtient donc une image de dimension $2K + 1$ contenant un indice spectral et $2K$ indices spatiaux. Dans un contexte hyperspectral, nous considérons un profil morphologique étendu [3] composé de la concaténation de l'ensemble des profils morphologiques associés aux indices spectraux évoqués dans la section précédentes :

$$EMP(\mathbf{X}) = \{MP_{IS^1}, MP_{IS^2}, \dots, MP_{IS^p}\} \quad (2.22)$$

où p est le nombre d'indices spectraux retenus. On obtient cette fois une image de dimension $p(2K + 1)$ composée de p indices spectraux et $2Kp$ indices spatiaux (voir figure 2.12). Le nombre d'indices pouvant être assez élevé, il est alors recommandé de procéder à une réduction de dimension du profil $EMP(\mathbf{X})$.

4 Processus de classification

Le processus de classification est divisé en plusieurs étapes (voir figure 2.13).

Tout d'abord, la dimension de l'image est réduite au moyen d'une des méthodes évoquées dans la section 2 afin d'extraire P caractéristiques spectrales. Les

CHAPITRE 2. CLASSIFICATION VECTORIELLE DE CARACTÉRISTIQUES SPECTRALES ET SPATIALES

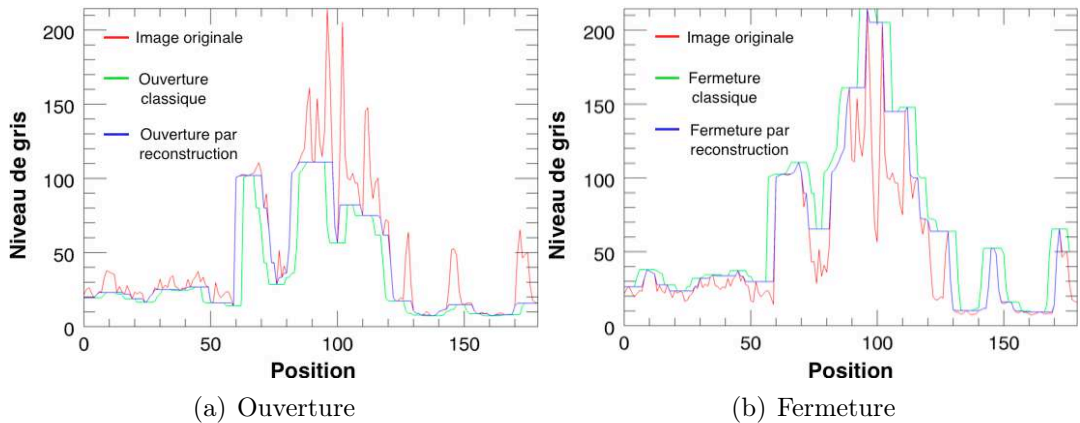


FIG. 2.11: Comparaison de l'ouverture et de la fermeture morphologique classiques (en vert) avec les opérateurs équivalents de reconstruction (en bleu) sur un signal 1D représentant une ligne de pixels d'une image (en rouge). L'ouverture et la fermeture standards permettent de supprimer les éléments (respectivement clairs ou obscurs) trop petits mais modifient également les contours des objets, ce qui n'est pas le cas des opérateurs de reconstruction.

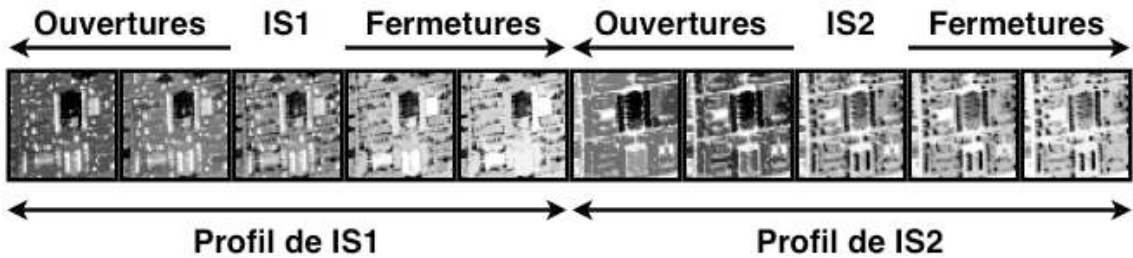


FIG. 2.12: Exemple de profil morphologique étendue pour une image à deux indices spectraux

méthodes d'extraction détaillées dans la section 3 sont ensuite appliquées sur les caractéristiques spectrales de sorte à obtenir Q caractéristiques spatiales. Toutes ces caractéristiques sont ensuite concaténées pour former une image \mathbf{I}_{SS} comportant $P+Q$ bandes. L'image \mathbf{I}_{SS} est finalement classifiée au moyen d'une approche vectorielle [8]. Ce type d'approche, à l'opposé de l'approche marginale appliquant un même traitement sur chaque variable d'une image multi-composantes avant de procéder à une fusion des résultats, traite la totalité des composantes de manière simultanée.

Pour réaliser l'étape d'affectation des classes, nous avons utilisé une méthode supervisée qui fait référence, les Séparateurs à Vastes Marges (SVM), ainsi qu'une méthode non supervisée, les modèles de mélanges gaussiens.

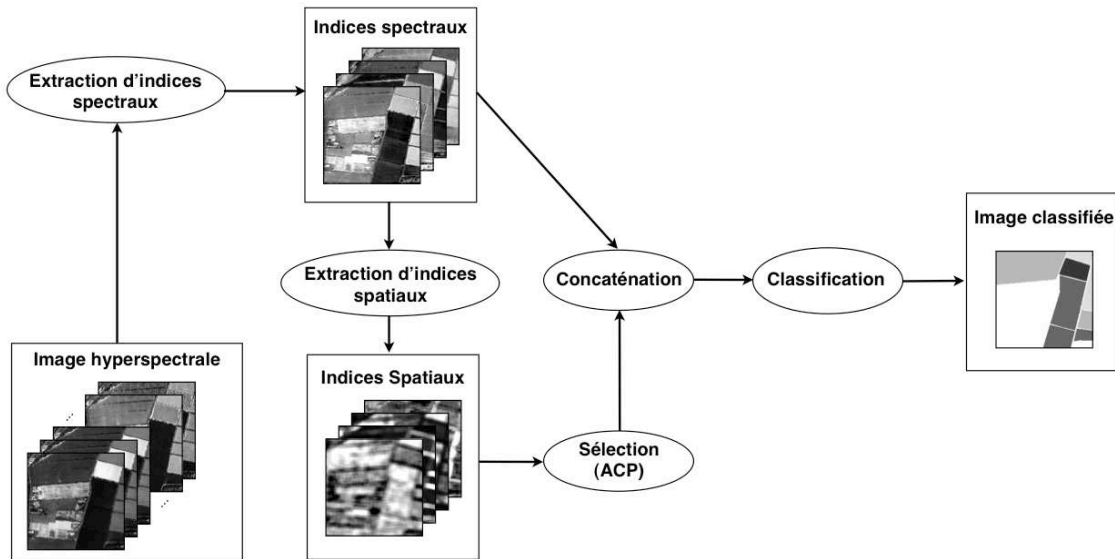


FIG. 2.13: Schéma du processus de classification

4.1 Classification supervisée par SVM

Les séparateurs à vastes marges (SVM) ont à l'origine été introduits dans le domaine de la théorie de l'apprentissage statistique pour la résolution de problèmes tels que la régression ou la classification. Ces outils ont pour objectif de chercher géométriquement la surface de séparation optimale entre les classes en déterminant le sous-ensemble d'échantillons d'apprentissage le plus représentatif des frontières entre classes (appelés vecteurs supports). Cette méthode est intéressante dans le cas où le nombre d'échantillons d'apprentissage est faible, d'une part car la détermination des vecteurs supports peut être faite à partir de peu d'échantillons, et d'autre part car la complexité de la méthode dépend en grande partie du nombre d'échantillons et non de la dimension de l'image à traiter. Les SVM peuvent ainsi classifier des données multivariées de grande dimension sans souffrir du phénomène de Hughes tout en séparant des classes très proches si les échantillons d'apprentissage sont suffisamment représentatifs.

Considérons un problème de classification bi-classes où l'on dispose de T échantillons d'apprentissage $(y_i, \mathbf{x}_i), i = \{1, \dots, T\}$ tels que $y_i \in \{-1, 1\}$ est l'étiquette de la classe d'appartenance de l'échantillon et $\mathbf{x}_i \in \mathbb{R}^M$ la valeur du pixel associé. La frontière entre les deux classes étudiée sera représentée par un hyperplan de séparation optimale validant deux conditions. D'une part, tous les échantillons appartenant à la même classe doivent être situés du même côté de l'hyperplan. D'autre part, la distance des vecteurs les plus proches à l'hyperplan (les vecteurs supports) dans chacune des deux classes doit être maximale (voir figure 2.14). C'est la contrainte de maximisation de la marge. L'équation de

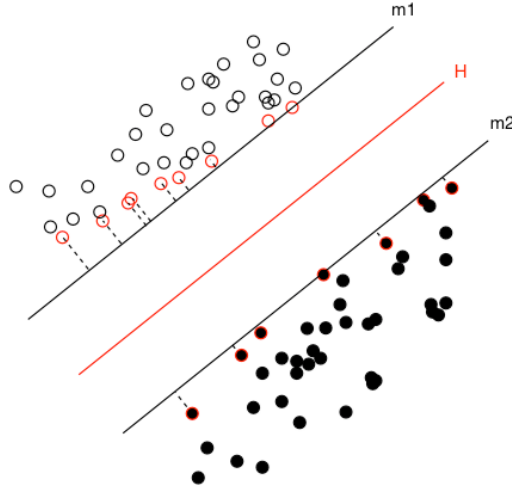


FIG. 2.14: Hyperplan 2D (en rouge) d'un SVM. Les points cerclés de rouge sont les vecteurs supports qui définissent la marge maximale (définie par $m1$ et $m2$) entre les échantillons et l'hyperplan H

l'hyperplan de séparation est ainsi définie par :

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \quad (2.23)$$

où \mathbf{w} est un vecteur normal à l'hyperplan, \mathbf{x} est un point de ce dernier et b est tel que $\frac{|b|}{\|\mathbf{w}\|}$ est la distance orthogonale entre l'origine et l'hyperplan. La première condition évoquée plus haut peut être modélisée par la contrainte suivante :

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1, \forall i \quad (2.24)$$

tandis que la seconde revient à trouver \mathbf{w} et b minimisant $\frac{1}{2}\mathbf{w}^2$. Cette procédure d'optimisation peut être résolue efficacement par maximisation du Lagrangien :

$$L(\lambda) = \sum_{i=1}^T \lambda_i - \frac{1}{2} \sum_{i=1}^T \sum_{j=1}^T \lambda_i \lambda_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \quad (2.25)$$

où les $(\lambda_i)_{i \in \{1, \dots, T\}}$ sont des multiplicateurs de Lagrange, tel que $\sum_{i=1}^T \lambda_i y_i = 0$, $\mathbf{w} = \sum_{i=1}^T \lambda_i y_i \mathbf{x}_i$ et $\lambda_i \geq 0, \forall i$. Un échantillon de la base d'apprentissage est identifié en tant que vecteur support si le multiplicateur de Lagrange associé est assez proche de 1. Au final, le classifieur est représenté par la fonction signe :

$$f(\mathbf{x}, \mathbf{w}, b) = \text{sgn}(\mathbf{w} \cdot \mathbf{x} + b) \quad (2.26)$$

La méthode décrite précédemment permet de gérer des classes linéairement séparables mais voit son efficacité décroître lorsque ces dernières ne le sont pas.

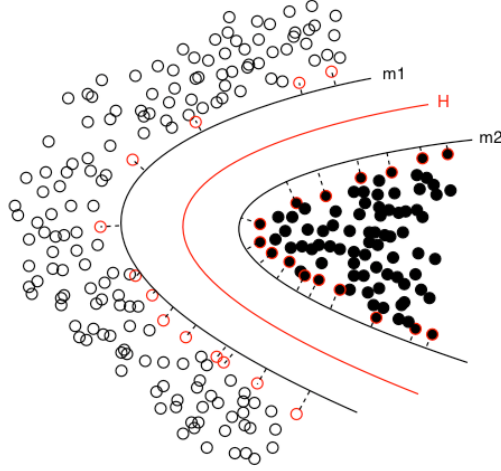


FIG. 2.15: Exemple d'hyperplan capable de gérer les cas non linéairement séparables.

Dans de tels cas, il existe une généralisation des SVM permettant de modéliser des surfaces de décision non linéaires en projetant les données dans un espace H de dimension supérieure à l'espace initial, tel que les données soient linéairement séparables dans H (cf figure 2.15). Soit \mathbf{x} un vecteur de l'espace initial. Son projeté sur H s'écrit $\Phi(\mathbf{x})$ où $\Phi : \mathbb{R}^M \rightarrow H$ est une fonction de projection. Le produit scalaire $\mathbf{x}_i \cdot \mathbf{x}_j$ de l'équation 2.27 devient alors $\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$ que l'on peut calculer avec une fonction noyau $K(\mathbf{x}_i, \mathbf{x}_j)$ ayant les propriétés d'un produit scalaire sans connaître la forme explicite de la fonction Φ . Le Lagrangien à optimiser devient donc :

$$L(\lambda) = \sum_{i=1}^T \lambda_i - \frac{1}{2} \sum_{i=1}^T \sum_{j=1}^T \lambda_i \lambda_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \quad (2.27)$$

Tout ce que l'on a décrit jusqu'à présent n'est valable que pour des cas bi-classes. Pour les cas multiclassés, deux stratégies existent. Soit K le nombre de classes. La première stratégie, appelée *Une contre toutes*, consiste à déterminer pour chaque classe un hyperplan la séparant de toutes les autres, ce qui implique l'utilisation de K classifieurs SVM. Au final, on choisit pour chaque pixel la classe associée au classifieur donnant la valeur de confiance (distance vis-à-vis de l'hyperplan) la plus élevée. Pour la seconde, *Une contre une*, on applique la méthode SVM à chaque paire de classes et lors de la classification, on associe à un pixel l'étiquette qui a été obtenue le plus souvent. C'est cette dernière stratégie que nous retiendrons. En effet, si elle nécessite l'utilisation de $K(K-1)/2$ classifieurs au lieu de K , le nombre d'échantillons considérés à chaque fois est beaucoup plus faible, ce qui réduit sensiblement la charge de calcul. Enfin, remarquons que l'algorithme SVM peut éventuellement tenir compte des échantillons aberrants en intégrant au modèle des coefficients de pénalisation que l'on peut estimer par

validation croisée.

4.2 Classification non supervisée par modèle de mélange gaussien

Un modèle de mélange gaussien est un modèle statistique modélisant une distribution de variables aléatoires sous la forme d'une somme de gaussiennes de formes et de tailles variées. Les paramètres à optimiser (selon un critère de maximum de vraisemblance) sont les caractéristiques de ces gaussiennes, telles que leur variance, leur moyenne et leur amplitude. Cette optimisation est le plus souvent faite au moyen de l'algorithme EM (*Expectation-Maximisation*). Dans le cadre de cette étude, nous utilisons le logiciel Mixmod [33] dont nous expliquons le fonctionnement théorique dans la suite de cette section.

4.2.1 Description du modèle et algorithme EM

Soit $\mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ l'ensemble des N vecteurs pixels de taille M d'une image multi-bandes. Chaque pixel \mathbf{x}_i peut être considéré comme une variable aléatoire de densité de probabilité :

$$f(\mathbf{x}_i|\Lambda) = \sum_{k=1}^K p_k h(\mathbf{x}_i|\mu_k, \Sigma_k) \quad (2.28)$$

où $h(\cdot|\lambda_k)$ est la distribution gaussienne associée à la classe k paramétrée par $\lambda_k = (\mu_k, \Sigma_k)$ tel que μ_k est la moyenne de la distribution et Σ_k sa matrice de covariance. p_k représente le poids de la distribution h associée dans le mélange et est tel que $0 < p_k < 1, \forall k$ et $p_1 + p_2 + \dots + p_K = 1$. $\Lambda = (p_1, \dots, p_K, \lambda_1, \dots, \lambda_K)$ représente le vecteur paramètre à estimer. Chaque pixel \mathbf{x}_i est associé à un label $\mathbf{z}_i = \{z_{i1}, z_{i2}, \dots, z_{iK}\}$ tel que $z_{ik} = 1$ si \mathbf{x}_i est associé à la k -ième distribution (la classe k). Dans notre cas, les labels \mathbf{z}_i sont inconnus, ce qui revient à résoudre un problème d'estimation de densité où Λ doit être optimisé de sorte à maximiser la log-vraisemblance :

$$L(\Lambda|\mathbf{x}_1, \dots, \mathbf{x}_N) = \sum_{i=1}^N \ln \left(\sum_{k=1}^K p_k h(\mathbf{x}_i|\lambda_k) \right) \quad (2.29)$$

Comme nous l'avons mentionné précédemment, cette optimisation peut être réalisée au moyen de l'algorithme EM. A partir d'un ensemble initial de paramètres Λ^0 choisi arbitrairement, on alterne itérativement entre une étape de calcul de l'espérance et une étape de maximisation. Lors de la première étape, les probabilités conditionnelles relatives à chaque classe sont calculées de la façon suivante :

$$t_{ik}^q = t_k^q(\mathbf{x}_i|\Lambda^{q-1}) = \frac{p_k^{q-1} h(\mathbf{x}_i|\lambda_k^{q-1})}{\sum_{l=1}^K p_l^{q-1} h(\mathbf{x}_i|\lambda_l^{q-1})} \quad (2.30)$$

où q est l'itération courante. Puis l'ensemble de paramètres de l'itération suivante Λ^q est calculé par maximisation de :

$$F(\Lambda|\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{t}^q) = \sum_{i=1}^N \sum_{k=1}^K t_{ik}^q \ln(p_k^{q-1} h(\mathbf{x}_i|\lambda_k^{q-1})) \quad (2.31)$$

avec $\mathbf{t}^q = (t_{ik}^q)_{i \in \{1, \dots, N\}, k \in \{1, \dots, K\}}$. p_k est ensuite mis à jour :

$$p_k^q = \frac{\sum_{i=1}^N t_{ik}^q}{N} \quad (2.32)$$

La mise à jour des λ_k dépend quant à elle de la distribution que l'on a choisi pour $h(\cdot|\lambda_k)$. Dans notre cas, une distribution gaussienne de paramètres $\lambda_k = (\mu_k, \Sigma_k)$ aura pour expression :

$$h(\mathbf{x}_i|\mu_k, \Sigma_k) = (2\pi)^{-d/2} |\Sigma_k|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{x}_i - \mu_k)^T \Sigma_k^{-1} (\mathbf{x}_i - \mu_k) \right\} \quad (2.33)$$

La classe associée à une telle distribution est ellipsoïdale, centrée sur la moyenne μ_k et ses caractéristiques géométriques sont reflétées par sa matrice de covariance Σ_k . Afin de simplifier le modèle, nous considérons que Σ_k est diagonale, ce qui nous permet de diminuer de façon drastique le nombre de paramètres du modèle. A chaque maximisation, la moyenne μ_k et la matrice de covariance Σ_k sont ainsi mises à jour de la façon suivante :

$$\mu_k^q = \frac{\sum_{i=1}^N t_{ik}^q \mathbf{x}_i}{\sum_{i=1}^N t_{ik}^q} \quad (2.34)$$

et :

$$\Sigma_k^q = \frac{\text{diag}(W_k^q)}{\sum_{i=1}^N t_{ik}^q} \quad (2.35)$$

telle que :

$$W_k^q = \sum_{i=1}^N t_{ik}^q (\mathbf{x}_i - \mu_k^q)(\mathbf{x}_i - \mu_k^q)^T \quad (2.36)$$

Soit $\mathbf{z}_i = (z_{ik})_{k \in \{1, \dots, K\}}$ le label associé au pixel \mathbf{x}_i , tel que $z_{ik} = 1$ si \mathbf{x}_i appartient à la classe K et $z_{ik} = 0$ sinon. Une fois l'itération finale f terminée, les \mathbf{z}_i sont calculés par l'intermédiaire d'un algorithme type Maximum A Posteriori :

$$z_{ik} = \begin{cases} 1 & \text{si } k = \underset{l=1, \dots, K}{\text{argmax}}(t_{il}^q) \\ 0 & \text{sinon} \end{cases} \quad (2.37)$$

L'algorithme CEM, une variante de l'algorithme EM spécifiquement dédiée à la classification, ajoute une étape d'étiquetage entre le calcul des probabilités conditionnelles (E) et la maximisation de la log-vraisemblance (M). Lors de

cette étape (C), une classification intermédiaire est réalisée en associant chaque pixel \mathbf{x}_i à la classe maximisant la probabilité conditionnelle. Une partition $P^q = (P_1^q, \dots, P_K^q)$ est ainsi créée. Dans l'étape suivante, la log-vraisemblance à maximiser n'est plus celle de l'équation 2.31 mais la suivante :

$$CF(\Lambda|\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{t}^q) = \sum_{i=1}^N \sum_{k=1}^K z_{ik}^q \ln(p_k^{q-1} h(\mathbf{x}_i | \lambda_k^{q-1})) \quad (2.38)$$

De même, la mise à jour des paramètres est modifiée de sorte que :

$$p_k^q = \frac{|P_k^q|}{N} \quad (2.39)$$

$$\mu_k^q = \frac{\sum_{i=1}^N z_{ik}^q \mathbf{x}_i}{|P_k^q|} \quad (2.40)$$

et

$$\Sigma_k^q = \frac{\text{diag}(W_k^q)}{|P_k^q|} \quad (2.41)$$

telle que :

$$W_k^q = \sum_{i=1}^N z_{ik}^q (\mathbf{x}_i - \mu_k^q)(\mathbf{x}_i - \mu_k^q)^T \quad (2.42)$$

Contrairement à l'algorithme EM, la variante CEM converge en un nombre fini d'itérations. Elle ne converge cependant pas vers un maximum local de la vraisemblance de Λ et peut ainsi résulter sur des paramètres suboptimaux, surtout lorsque les différents éléments du mélange se superposent ou que la proportion de chacun dans le mélange est trop disparate.

4.2.2 Sélection du nombre de classes

Dans cette étude, nous considérons des images impliquant des classes plus ou moins complexes (nombre de signatures spectrales différentes, présence ou non d'une information de texture). Toutefois, les applications de classification non supervisée telles que Mixmod ne tiennent pas compte de ces différents niveaux de complexité pour définir le nombre de classes K présentes dans une image et ont tendance à le surestimer. Nous avons ainsi décidé que ce nombre de classes serait la seule information connue *a priori* du processus de classification. Partant de cette base, nous avons considéré trois approches différentes pour aboutir à une carte de classification comportant le bon nombre de classes. L'approche la plus directe consiste à imposer à Mixmod un nombre de classes unique. La seconde, plus souple, permet au logiciel de choisir un nombre de classes optimal $K' \geq K$ (selon un certain critère) parmi un ensemble de valeurs prédéfinies, avant de

fusionner itérativement ces classes les unes aux autres par l'intermédiaire d'une Classification Hiérarchique Descendante (CHD), jusqu'à ce que le bon nombre de classes soit atteint.

Le logiciel Mixmod propose plusieurs critères de sélection automatique du nombre de classes. Nous avons pour notre part choisi d'utiliser le critère BIC (*Bayesian Information Criterion*) qui selon [33] est le plus efficace dans le cadre d'une estimation de densité. Son objectif est de sélectionner le couple (g, K) , g étant le modèle de distribution h utilisé, maximisant la vraisemblance intégrée :

$$\mathbf{f}(\mathbf{x}|g, K) = \int_{\Theta_{g,K}} \left(\prod_{i=1}^N f(\mathbf{x}_i|g, K, \Lambda) \right) \pi(\Lambda|g, K) d\Lambda \quad (2.43)$$

où $\Theta_{g,K}$ est l'espace de paramètres du couple (g, K) et $\pi(\Lambda|g, K)$ un a priori sur Λ associé à ce couple. Cela mène à minimiser le critère BIC suivant :

$$BIC_{g,K} = -2 \log \mathbf{f}(\mathbf{x}|g, K, \hat{\Lambda}) + \nu_{g,K} \ln N \quad (2.44)$$

tel que $\nu_{g,K}$ est le nombre de paramètres libres du modèle g à K classes et $\hat{\Lambda} = \arg \max \mathbf{f}(\mathbf{x}|g, K, \Lambda)$. Dans le cas où l'on considère que Σ_k est une matrice diagonale, on a $\nu_{g,K} = 2MK$.

Afin de réduire le nombre de classes, la CHD a besoin d'un critère de fusion permettant de mesurer la similarité de deux classes C_a et C_b . Le plus classique est la distance de Ward qui correspond à la distance euclidienne entre les centroïdes des deux classes pondérée par leur cardinal :

$$D_W(C_a, C_b) = \sqrt{\frac{|C_a||C_b|}{|C_a| + |C_b|}} D_E(C_a, C_b) \quad (2.45)$$

où $D_E(C_a, C_b)$ est la distance euclidienne entre les centroïdes des classes C_a et C_b . Ce critère permet de sélectionner les deux classes dont le regroupement fera le moins diminuer l'inertie interclasse. Nous avons également considéré un second critère, basé sur le calcul de la matrice de covariance $\Sigma(C_a, C_b)$ de l'ensemble des échantillons formé par le regroupement de C_a et C_b . A chaque itération sont fusionnées les deux classes C_a et C_b telles que la trace de $\Sigma(C_a, C_b)$ soit minimale.

5 Résultats

Nous présentons dans cette partie les résultats de classification obtenus par les algorithmes décrits dans la section 4, appliqués sur des jeux de données composés de divers mélanges d'indices spectraux et spatiaux (décrits dans les sections 2 et 3). Précisons qu'afin de ne pas privilégier un type d'information par rapport à l'autre, le nombre Q d'attributs spatiaux est réduit par ACP de sorte qu'il soit égal au nombre P d'attributs spectraux.

5.1 Mesures de précision des résultats

Afin d'estimer la qualité des résultats, nous utilisons une mesure obtenue par comparaison avec une carte de référence réalisée à partir de la vérité terrain : la précision générale. Cette mesure est extraite d'une matrice de confusion Γ , qui est un tableau de taille $K \times K$ tel que $\Gamma(i, j)$ où $i, j \in \{1, \dots, K\}$ correspond au nombre de pixels de la carte de classification qui ont été associés à la classe i alors qu'ils appartenaient à la classe j sur la carte de référence. Les valeurs diagonales de Γ correspondent donc aux pixels bien classés tandis que les autres valeurs correspondent aux confusions inter-classes.

La précision générale (P_G) représente la proportion de pixels bien classés, sans distinguer les classes :

$$P_G = \frac{\sum_{k=1}^K \Gamma(k, k)}{\sum_{k=1}^K \sum_{k'=1}^K \Gamma(k, k')} \quad (2.46)$$

Dans le cadre du modèle de mélange gaussien (non supervisé), seuls les pixels associés à un label de la vérité de terrain des images sont utilisés dans le processus de classification (c'est-à-dire 9345 pixels pour *Indian Pines* et la totalité des pixels pour les autres images). Remarquons également que ce type d'approche attribue à chacune des classes un label qui ne correspond pas forcément à celui de la vérité terrain. Afin d'appareiller cette dernière à la carte de classification, quelques permutations de label sont ainsi nécessaires. Soit L_{VT} l'ensemble des labels de la vérité terrain et L_{CC} l'ensemble des labels de la carte de classification. Pour chaque élément $l \in L_{CC}$, on recherche le label $l' \in L_{VT}$ tel que le nombre de pixel associé à l sur la carte de classification et à l' sur la vérité terrain soit maximal. Tous les pixels associés à l sur la carte de classification sont alors associés à l' . Ce post-traitement rend possible la mesure de précision du résultat d'une classification non supervisée.

5.2 Etude de sensibilité

Le calcul des caractéristiques spectrales et spatiales dépend de plusieurs paramètres qui peuvent avoir beaucoup d'influence sur la précision des algorithmes de classification. L'étude de sensibilité que nous présentons par la suite vise d'une part à mesurer l'impact de chaque paramètre sur la qualité des résultats, et d'autre part à identifier pour chaque jeu de données (et par extension pour toute image présentant des caractéristiques semblables) un paramétrage optimal. Nous commencerons par étudier l'influence des caractéristiques spectrales, puis celle de chaque type de caractéristique spatiale - coefficients de Haralick, spectres de texture et profils morphologiques - et nous concluerons par la comparaison de ces dernières. Dans chaque cas, nous considérons un point de vue supervisé (SVM) et non supervisé (modèle de mélange gaussien).

Pour le classifieur SVM, nous utilisons un noyau polynomial de degré raisonnable ($d = 2$) avec un paramètre de pénalisation fixé à $C = 1500$. Une validation

croisée a en effet montré que ce paramétrage était efficace sur les quatre jeux de données utilisés ainsi que les différents jeux d'échantillons d'apprentissage utilisés. Dans un premier temps, nous utilisons pour la phase d'apprentissage de l'algorithme 1% des échantillons de chacun des K ensembles associés aux images, ces échantillons étant répartis de façon aléatoire sur l'ensemble des parcelles à notre disposition (méthode SVM1). Cependant, lorsque l'on dispose d'une vérité terrain pour un jeu de données que l'on désire classifier, cette dernière est généralement limitée à un ensemble d'échantillons regroupés sur un faible nombre de parcelles (et parfois une seule). Dans un souci de réalisme, nous considérons ainsi un second cas de figure tel que la totalité des échantillons retenus pour l'apprentissage soient groupés et issus d'une parcelle unique pour chaque classe (méthode SVM2). En ce qui concerne le modèle de mélange gaussien, nous avons testé les divers types d'algorithme d'optimisation mentionnés dans la section 4.2 : EM et CEM. Les résultats obtenus avec les deux premiers algorithmes manquant de stabilité, nous avons utilisé l'algorithme CEM couplé à un modèle de type *Gaussian_pk_L_Bk* [33] considérant que la matrice de covariance associée à chaque classe est diagonale. Afin de minimiser la probabilité que la solution converge vers un minimum local, 10 répétitions de 50 itérations de l'algorithme CEM sont tout d'abord lancées. La carte de classification associée à la valeur de log-vraisemblance optimale est conservée puis utilisée en tant que point de départ d'une ultime occurrence de 200 itérations de l'algorithme CEM. En ce qui concerne la sélection du nombre de classes, parmi les trois possibilités évoquées dans la section 4.2.2, c'est l'approche où le nombre de classes est fixé qui donne les résultats les plus stables dans la totalité des cas ; c'est donc celle que nous avons retenue pour la suite.

Dans cette section, nous commençons par nous focaliser sur le cas particulier de l'image *Indian Pines*, puis nous étudions l'impact de chacun des paramètres utilisés pour la création des attributs aussi bien spectraux que spatiaux. Les tableaux 2.4 à 2.8 permettent ainsi d'étudier l'influence de ces paramètres sur la précision de la classification. Nous calculons dans chaque cas la précision moyenne obtenue lorsqu'un des paramètres est fixé, et ce pour toutes les valeurs possibles - que nous spécifions par la suite - de ce paramètre. Le tableau 2.1 montre comment sont calculées ces valeurs moyennes à partir d'un ensemble de résultats obtenus sur l'image CO avec des profils morphologiques.

Nous utilisons les acronymes suivant pour désigner les types de caractéristique :

- CS : Composantes Spectrales
- CH : Coefficients de Haralick
- PM : Profils Morphologiques
- ST : Spectres de Texture

CHAPITRE 2. CLASSIFICATION VECTORIELLE DE CARACTÉRISTIQUES SPECTRALES ET SPATIALES

Algorithme de réduction de dimension	Nombre de bandes conservées	Résultats purement spectraux	Résultats spectraux + spatiaux (PM)		
			Niveau de granulométrie		
			2	3	4
ACP	GR = 2	78,94	85,16	86,13	85,91
	CV = 3	95,52	97,29	96,93	96,53
	MV = 5	96,98	97,86	97,87	97,81
	MIN = 10	98,17	98,77	99,03	98,93
ACPS	GR = 2	79,41	85,79	85,75	85,58
	CV = 3	95,55	96,54	96,54	96,28
	MV = 5	96,37	97,65	97,38	97,3
	MIN = 10	98,37	99,1	98,29	99,15
MNF	GR = 2	91,6	96,06	96,16	96,57
	CV = 3	95,52	97,93	97,78	97,87
	MV = 5	97,92	98,89	98,89	98,6
	MIN = 10	98,54	99,04	99,19	99,11

TAB. 2.1: Ce tableau présente l'ensemble des précisions de classification obtenues sur l'image CO en fonction des différents paramétrages spectraux et spatiaux possibles pour la méthode basée sur les profils morphologiques. La précision moyenne de la classification lorsque le niveau de granulométrie est de 3, quel que soit le nombre de bandes conservées, est calculée à partir des valeurs de précisions incluses dans la colonne rouge. Parallèlement, la précision moyenne de la classification lorsque l'on ne conserve que 3 composantes, quel que soit le niveau de granulométrie, est calculée à partir des valeurs de précisions incluses dans les lignes bleues

5.2.1 Le cas *Indian Pines*

Le tableau 2.2 présente la précision moyenne des résultats de classification de l'image *Indian Pines* obtenus avec divers paramétrages des caractéristiques spectrales et spatiales (que l'on détaillera par la suite).

Algorithme de classification	Caractéristiques utilisées			
	CS	CS + CH	CS + PM	CS + ST
SVM1	63,8	75,9	73	68,7
SVM2	54,8	56,8	57,4	53,7
Mixmod	43,4	48,9	43,6	47,9

TAB. 2.2: Résultats de classification moyens (en %) obtenus pour l'image *Indian Pines* et dépendant de l'algorithme de classification employé et du ou des types de caractéristique utilisés.

Si les résultats de la classification SVM1 sont acceptables, surtout lorsque l'on ajoute des caractéristiques spatiales, nous observons une forte baisse de précision avec la classification SVM2, aussi bien pour les résultats purement spectraux que pour les résultats spectraux-spatiaux. Ce décalage montre que les classes de la vérité terrain associée à l'image *Indian Pines* ne sont pas totalement homogènes d'une parcelle à l'autre. En effet, deux parcelles appartenant à une même classe thématique ne présentent pas toujours les mêmes caractéristiques spectrales et/ou spatiales (voir figure 2.16). Typiquement, dans le cas de parcelles cultivées, des paramètres tels que le niveau de maturité peuvent changer la nature du spectre associé à la plante elle-même, tandis que différents niveaux de densité de culture impliquent différentes proportions entre cette dernière et le sol nu. Cette hétérogénéité justifie notamment la faible précision des résultats de Mixmod qui ne peut raisonnablement fusionner en une seule classe des parcelles aussi spectralement dissemblables.

Constatant la non adéquation de la vérité terrain associée à l'image *Indian Pines* pour la validation de nos méthodes de classification, nous avons décidé de ne pas utiliser ce jeu de données pour l'étude de sensibilité qui suit. Nous le conserverons cependant pour le chapitre suivant, afin de comparer nos résultats avec ceux obtenus dans la littérature.

5.2.2 Caractéristiques spectrales

L'extraction de caractéristiques spectrales est réalisée par l'intermédiaire de trois algorithmes de réduction de dimension (décrits dans la section 2) : l'ACP, l'ACP Spatiale (ACPS) et la méthode MNF. Dans chaque cas, nous considérons quatre critères de sélection du nombre de composantes à conserver : le rapport de croissance (RC), deux seuillages (VC et VM) et un nombre de composantes fixé à 10 (MAX). Le premier seuillage conserve un nombre de composantes suffisant

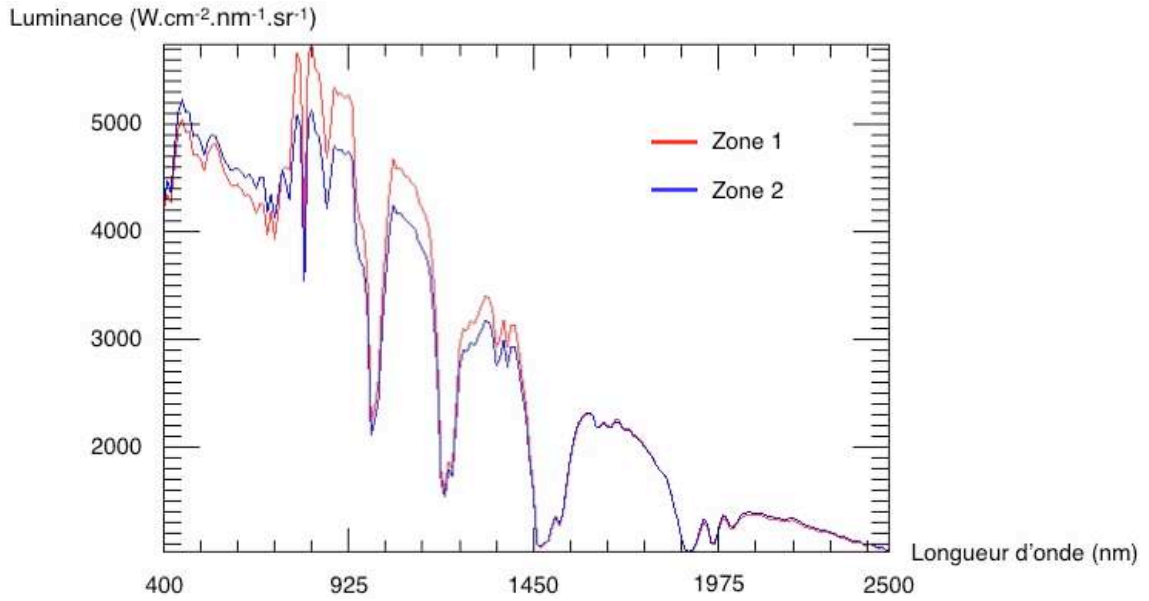


FIG. 2.16: Spectres de luminance moyens de deux parcelles appartenant à la même classe thématique.

pour contenir au minimum 99% de la variance cumulée, tandis que le second conserve toutes celles dont la variance est supérieure à 1%. Le nombre maximal de composantes a été fixé à 10, limite à partir de laquelle la complexité devient trop élevée pour GMM. Dans les tableaux qui suivent, ces quatre critères de sélection sont classés en fonction du nombre de bandes qu'ils proposent de conserver. Le critère MAX mis à part, ces nombres sont susceptibles de changer d'une image à l'autre. En général, le critère RC préconise de conserver de 2 à 3 composantes là où les critères VC et VM en recommandent respectivement de 4 à 5 et de 5 à 6.

Le tableau 2.3 montre que les classifieurs SVM1 et SVM2 donnent déjà de bons résultats (supérieur à 95% dans le meilleur des cas) lorsque seules des caractéristiques spectrales sont utilisées, et ce quel que soit le jeu de données considéré. La précision de classification est cependant moins convaincante en ce qui concerne les résultats obtenus avec GMM, notamment sur l'image CO. C'est ainsi dans un contexte non supervisé que l'information spatiale est susceptible d'être la plus nécessaire, particulièrement sur des images telles que CO dont les classes placent ces deux types d'information au moins au même niveau d'importance.

Le tableau 2.4 évalue l'impact du nombre de composantes spectrales conservées sur la précision de classification. Ce dernier est assez difficile à mesurer car il dépend non seulement de la présence ou non de connaissance a priori sur les données à classifier, mais aussi de la nature de cette connaissance. Avec un classifieur SVM1, le résultat de classification est d'autant plus précis que l'on conserve de composantes et ce, quel que soit le jeu de données considéré, alors qu'avec un classifieur SVM2, les résultats sont à la fois moins bons et moins stables. La

CHAPITRE 2. CLASSIFICATION VECTORIELLE DE CARACTÉRISTIQUES
SPECTRALES ET SPATIALES

MRD	CNC	MR			MR2			CO		
		SVM1	SVM2	GMM	SVM1	SVM2	GMM	SVM1	SVM2	GMM
ACP	RC	95,4	92,8	84,5	94,2	93,3	81,9	78,9	75,7	56,4
	VC	95,4	92,8	84,5	94,2	93,3	81,9	95,5	96,5	56
	VM	97,6	96,2	94,3	98	96,8	79,1	97	95,8	63,3
	MAX	96,3	97,5	82,6	97,1	93,1	96,2	98,2	96	59,7
ACPS	RC	95,3	92,8	73,8	93,9	93,2	86,3	79,4	78,1	56,2
	VC	95,3	92,8	73,8	96,6	95,6	81,7	95,6	96,3	56
	VM	97,3	96	83	97,8	97,5	80,1	96,4	94,2	78,8
	MAX	96,7	96,8	85,1	97	97,1	96,9	98,4	95,4	83,1
MNF	RC	96,9	96	93,3	97,9	97,2	90,9	91,6	90,5	36,7
	VC	96,9	96	93,3	98,2	95,9	91,7	95,5	94,7	49,1
	VM	97,5	97,5	94,2	97,9	97,2	90,9	97,9	96,9	45,2
	MAX	97,8	96,1	94,8	97,4	92,6	92,2	98,5	96	54

TAB. 2.3: Pourcentage de bonne classification de jeux de caractéristiques exclusivement spectrales, en fonction de la méthode de réduction de dimension (MRD), du critère de sélection du nombre de composantes (CNC), de l'image concernée ainsi que de l'algorithme de classification utilisé.

Image	Classification	CNC			
		RC	VC	VM	MAX
CO	SVM1	91	97,1	98,1	98,5
	SVM2	85,7	92,5	93,3	93,3
	Mixmod	60,4	66,6	74,7	79,8
MR	SVM1	96,4	96,4	98,1	98,3
	SVM2	87,5	87,5	92,3	89,7
	Mixmod	86,7	86,7	87,6	87,6
MR2	SVM1	96,5	96,9	97,9	97,9
	SVM2	91,8	92	93,3	90,5
	Mixmod	89,1	88,7	91,2	92,6

TAB. 2.4: Influence du critère de sélection du nombre de composantes à conserver (CNC) sur la précision de la classification.

Image	Classification	Méthode RD		
		ACP	ACPS	MNF
CO	SVM1	95,4	95,4	97,4
	SVM2	90,1	89,5	93,7
	Mixmod	75,9	74,9	60,4
MR	SVM1	97,8	97,3	98
	SVM2	87,9	89,6	90,1
	Mixmod	87,7	84,6	89,6
MR2	SVM1	97,2	97	98,1
	SVM2	91,9	91,9	91,6
	Mixmod	92,7	90,2	90,9

TAB. 2.5: Influence du choix de l’algorithme de réduction de dimension sur la précision de la classification.

précision de classification n’est plus proportionnelle au nombre de composantes utilisées et atteint la plupart du temps son maximum avec les nombres de bandes intermédiaires déterminés par les méthodes de seuillage. Les résultats de classification non supervisée suivent pour leur part la même tendance que ceux du classifieur SVM1. La solution la plus adaptée consiste ainsi à se limiter au nombre de bandes préconisé par les méthodes de seuillage, et notamment le critère de variance maximale (VM). En effet, même lorsqu’un nombre de bandes supérieur permet d’atteindre une meilleure précision de classification, le gain est faible en comparaison de l’augmentation du coût en temps de calcul induit par la présence de ces bandes supplémentaires.

La tableau 2.5 évalue pour sa part l’impact de la méthode de projection choisie sur la précision de classification. En ce qui concerne les images MR et MR2, l’ACP et la MNF donnent des résultats comparables quel que soit le type de classification, avec tout de même un léger avantage pour la MNF. Sur l’image CO en revanche, si la MNF fonctionne légèrement mieux que l’ACP dans un contexte supervisé, elle dégrade fortement les résultats de classification dans un contexte non supervisé. Remarquons également que l’ACP spatiale donne au mieux des résultats identiques à ceux de l’ACP classique. D’un point de vue général, l’ACP se révèle ainsi être l’algorithme de réduction de dimension le plus fiable.

5.2.3 Caractéristiques spatiales

Nous commencerons par traiter le paramétrage de chaque type de caractéristique spatiale séparément, puis nous évaluerons les forces et faiblesses de chacun en fonction des jeux de données étudiés.

L’extraction de coefficients de Haralick dépend de trois paramètres : le nombre de niveaux de gris N_g résultant de la quantification de l’image, la largeur L du voisinage (associé à une fenêtre de taille $L \times L$) et enfin le ou les décalages de pixels

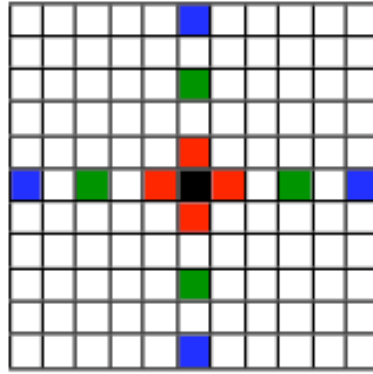


FIG. 2.17: Décalages horizontaux et verticaux de 1, 3 et 5 pixels (respectivement représentés en rouge, vert et bleu) vis-à-vis du pixel noir

Δ considérés, tous trois utilisés pour le calcul des matrices de co-occurrence. Une valeur de N_g trop élevée engendrerait des matrices de co-occurrence énormes, ce qui est problématique pour deux raisons. Tout d'abord, multiplier par deux le nombre de niveaux de gris reviendrait à quadrupler la taille de ces matrices, de même que la complexité algorithmique. Ensuite, la majorité des éléments d'une telle matrices de co-occurrence seraient nuls, ce qui diminuerait la pertinence des calculs statistiques effectués sur elle. Dans [17], Haralick préconise d'utiliser 16 niveaux de gris et beaucoup de travaux ayant par la suite utilisé ces coefficients ont suivi ce conseil. En ce qui concerne la taille du voisinage considéré pour chaque pixel, la fenêtre doit également être assez grande pour que les statistiques extraites soient robustes - minimisant ainsi l'impact du bruit - et assez petite pour diminuer le risque de chevauchement de textures. Nous avons testé trois tailles de voisinage allant d'une fenêtre de 11×11 pixels, ce qui nous paraît être un minimum pour construire des statistiques robustes, à une fenêtre de 31×31 pixels, au-delà de quoi le chevauchement de textures devient difficile à éviter, du moins sur les images que nous considérons dans cette étude. La taille intermédiaire choisie est de 21×21 pixels. Il reste enfin à choisir le ou les décalages (distance et direction) de pixels à partir desquels nous construirons la matrice de co-occurrence. A chaque fois nous considérons pour une distance donnée a les décalages de pixels des quatre directions principales : $\Delta_a = \{(a, 0), (-a, 0), (0, a), (0, -a)\}$. Afin de minimiser le nombre de coefficients, nous calculons ensuite une unique matrice correspondant à la moyenne des matrices de cooccurrence associées à chaque décalage de Δ_a (voir figure 2.17). L'étude menée par Weszka et al. [56] statue qu'une distance d'un ou deux pixels est généralement suffisante pour décrire une texture, ce que nous vérifions en testant des décalages allant de 1 à 5 pixels.

Les profils morphologiques dépendent quant à eux d'un unique paramètre, le niveau de granulométrie. Un niveau de granulométrie élevé permet d'extraire une information spatiale plus fine mais augmente le temps de calcul nécessaire à la création du profil morphologique. Nous avons testé trois niveaux de granulométrie

g allant de 2 à 4. Les profils morphologiques sont construits sur le principe de l'équation 2.21, sachant que d'un élément structurant à l'autre, le rayon de ce dernier n'augmente pas d'un pixel mais de deux, afin d'étudier un maximum d'échelles spatiales différentes sans crouler sous le nombre d'attributs spatiaux.

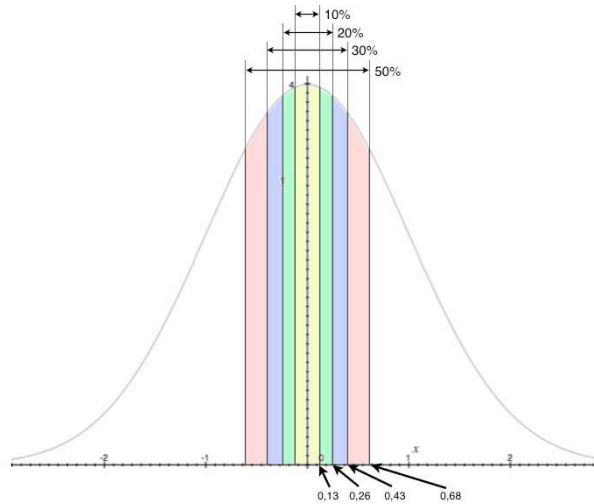


FIG. 2.18: Distribution représentant les différences entre pixels voisins. Chaque intervalle représente la probabilité qu'un pixel et son voisin soient considérés similaires en fonction de α .

Enfin, la paramétrisation des spectres de texture se fait à deux niveaux : la taille du voisinage et le seuil ψ associé à la création des unités de texture. En ce qui concerne la taille du voisinage, nous avons procédé de la même façon que pour les coefficients de Haralick. Quant au seuil ψ , au lieu de choisir une faible valeur indépendante des données comme le font He et al. [20], nous avons décidé de le rendre dépendant de l'écart-type global de l'image. Rappelons qu'une unité de texture est une valeur construite à partir des variations entre un pixel et ses huit voisins immédiats (cf. section 3.2). Pour chaque couple centre/voisin, nous considérons trois possibilités. Soit le compte numérique du centre est plus élevé que celui de son voisin, soit il est plus faible, soit les deux comptes numériques sont similaire. Cette notion de similarité est définie par l'intermédiaire du seuil ψ . Ce dernier est ainsi de la forme $\psi = \alpha\sigma$ tel que σ est l'écart-type global de l'image et α est un coefficient multiplicatif associé à la fonction de répartition de la loi normale. L'idée est de choisir α telle qu'un certain pourcentage des valeurs absolues de différences entre pixels voisins de la bande spectrale traitée lui soient inférieures. Les pourcentages que nous avons testés sont les suivants : 10%, 20%, 30% et 50%, chacun correspondant à une valeur de α respectivement égale à 0.13, 0.26, 0.43 et 0.68 (voir figure 2.18).

CHAPITRE 2. CLASSIFICATION VECTORIELLE DE CARACTÉRISTIQUES
SPECTRALES ET SPATIALES

Image	Paramètre fixe		Algorithme de classification		
			SVM1	SVM2	Mixmod
MR	L	11	98,6	91,3	84,1
		21	98,3	84,4	81,7
		31	97,9	79,2	82,8
	a	1	98,3	84	82,1
		3	98,3	85,1	83,1
		5	98,3	85,8	83,5
MR2	L	11	97,7	92,6	92,8
		21	97,5	89,5	92,7
		31	97,5	87,5	92,6
	a	1	97,6	89,8	92,6
		3	97,6	89,8	93,2
		5	97,6	89,9	92,3
CO	L	11	96,7	91,6	76,3
		21	96,9	87,1	81,3
		31	96,2	83,6	81,1
	a	1	96,6	88	79,9
		3	96,6	86,8	79,3
		5	96,7	87,5	79,5

TAB. 2.6: Influence des paramètres associés à la création des coefficients de Haralick, le rayon de la fenêtre de calcul de la matrice de co-occurrence (r) et la distance associée à chaque décalage de pixels (a), sur la précision de la classification.

Image	g	Algorithme de classification		
		SVM1	SVM2	Mixmod
MR	2	98,8	97,4	90,7
	3	98,9	97,1	92,9
	4	98,8	97	91,1
MR2	2	98,6	97,7	92,9
	3	98,5	97,5	93,7
	4	98,5	97,4	93,9
CO	2	95,8	93,1	62,8
	3	95,8	92	63,9
	4	95,8	92,2	63,1

TAB. 2.7: Influence du niveau de granulométrie (NG) des profils morphologiques, sur la précision de la classification.

CHAPITRE 2. CLASSIFICATION VECTORIELLE DE CARACTÉRISTIQUES SPECTRALES ET SPATIALES

Image	Paramètre fixe	Algorithme de classification			
		SVM1	SVM2	Mixmod	
MR	L	11	96,7	86,7	85,2
		21	96,3	80,7	87,7
		31	96,7	76,8	88,9
	ϕ	0,13 σ	96,2	80,2	88,9
		0,26 σ	96,7	83	87,2
		0,43 σ	96,8	81,7	89,5
		0,68 σ	96,5	80,6	83,5
MR2	L	11	96,4	88,5	87,4
		21	96,9	85,2	88,8
		31	96,4	80,9	89,5
	ϕ	0,13 σ	96,3	83,4	89,2
		0,26 σ	96,7	87	86,7
		0,43 σ	96,6	85,6	88,9
		0,68 σ	96,6	83	89,1
CO	L	11	97,9	94,2	85,3
		21	98,6	92,7	87,1
		31	98,4	90,4	87
	ϕ	0,13 σ	98	87,5	81,5
		0,26 σ	98,3	94,5	88,6
		0,43 σ	98,5	95,1	90
		0,68 σ	98,3	92,5	85,8

TAB. 2.8: Influence des paramètres associés à la création des spectres de texture, la taille de la fenêtre de calcul des spectres (r) et le seuil associé à la création des unités de texture (ϕ), sur la précision de la classification.

Le tableau 2.6 montre qu'en ce qui concerne les coefficients de Haralick, la taille du voisinage a peu d'impact sur la précision d'un classifieur SVM1. Il en va cependant tout autrement lorsque les échantillons sont extraits d'une parcelle unique. Dans ce cas particulier, la précision des résultats se révèle sensiblement plus élevée lorsque les coefficients de Haralick sont réalisés à partir de voisinages de petite taille (jusqu'à 12% d'amélioration vis-à-vis d'un voisinage plus étendu). Cette différence observée entre SVM1 et SVM2 est due à la localisation des échantillons utilisés pour l'apprentissage. Dans le cas de SVM1, ces derniers sont recueillis de manière aléatoire sur l'ensemble de la vérité terrain, c'est-à-dire aussi bien au niveau des frontières entre parcelles qu'au milieu. Dans le cas de SVM2, les échantillons ont été sélectionnés de façon groupée au milieu d'une des parcelles de chaque classe. Les zones frontalières ne sont ainsi pas prises en compte dans l'apprentissage de ce classifieur, ce qui peut expliquer la confusion entre classes qui y apparaît sur les résultats de classification (voir figure 2.19). Or, plus on aug-

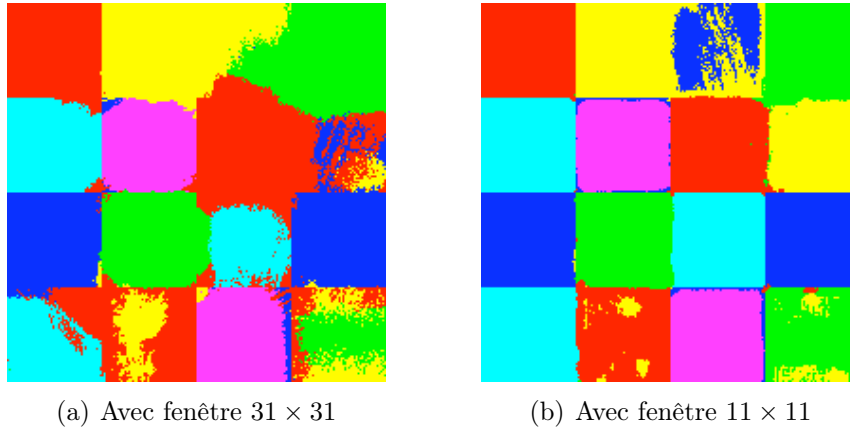


FIG. 2.19: Résultats de classification obtenus avec différentes tailles de voisinage. La carte de classification apparaît beaucoup plus incohérente au niveau des frontières lorsque le fenêtrage utilisé est trop grand.

mente la taille du voisinage d'étude lors de la création des coefficients de Haralick et plus ces derniers comporteront des pixels associés à une frontière entre deux parcelles de classe différente, des pixels que le classifieur aura beaucoup plus de mal à étiqueter. Mixmod agit pour sa part à un niveau plus global, c'est-à-dire qu'il utilise la totalité des pixels à sa disposition dans l'image pour effectuer la classification. C'est pourquoi l'impact de la taille du voisinage sur la précision de la classification sont beaucoup moins marquants, excepté pour l'image CO dont les textures de faible fréquence nécessitent - dans un cas non supervisé - de grands voisinages afin d'être décrits correctement. Comme le spécifiait Weszka et al. [56], la distance a n'a quant à elle que peu d'influence sur la précision des résultats.

Le tableau 2.7 nous montre pour sa part que le niveau de granulométrie n'a pas d'influence majeure sur la qualité des résultats de classification, ou du moins que le niveau 2 (tel que pour chaque bande de l'image, on calcule une ouverture et une fermeture par reconstruction avec des éléments structurants de taille 3×3 et 7×7) est suffisant pour mettre en valeur les textures de forte fréquence des images MR et MR2, tandis que de tels attributs spatiaux ne sont de toute façon pas adaptés à la mise en valeur des textures de faible fréquence de l'image CO.

Concernant les spectres de texture, la taille du voisinage affecte la précision des résultats de classification de façon relativement similaire aux coefficients de Haralick (voir tableau 2.8) pour SVM1 et SVM2. On peut toutefois remarquer une petite différence de tendance au niveau des résultats non supervisés qui se révèlent en effet légèrement meilleurs lorsque des voisinages de grande taille sont utilisés, probablement parce que les spectres de texture nécessitent beaucoup plus d'observations pour être statistiquement pertinents que les coefficients de Haralick. L'influence du seuil ϕ dépend elle aussi du type d'algorithme de classification utilisé. Avec SVM1, la différence de précision obtenue d'une valeur de

seuil à l'autre est encore une fois négligeable. Avec SVM2 et Mixmod, les valeurs de seuil intermédiaires et notamment $0,43\sigma$ donnent des résultats parfois bien meilleurs que les valeurs extrêmes.

5.2.4 Analyse et comparaisons des résultats

Dans cette section, nous comparons d'une part les résultats de classification purement spectrale avec les résultats de classification spatiale/spectrale et confrontons d'autre part les divers attributs spatiaux les uns aux autres. Pour ce faire, nous adoptons trois critères de comparaison : le pourcentage général de bonne classification, la carte de classification et la complexité.

Image	AC	Caractéristiques utilisées			
		CS	CS + CH	CS + PM	CS + ST
MR	SVM1	96,7	98,6	98,8	96,9
	SVM2	95,7	91,3	97,2	89,3
	GMM	85,2 ± 3,8	85,4 ± 4,4	91,8 ± 2,3	89 ± 3,4
MR2	SVM1	96,8	97,6	98,5	96,6
	SVM2	95,2	92,6	97,6	91,6
	GMM	88,5 ± 2	93,1 ± 1,6	93,7 ± 0,1	87,8 ± 1,9
CO	SVM1	93,6	96,6	95,8	98,3
	SVM2	92,2	91,6	92,4	94,7
	GMM	64 ± 1	76,4 ± 0,6	67,8 ± 0,2	86,8 ± 1,2

TAB. 2.9: Comparaison entre les résultats de classifications purement spectrales et ceux obtenus en intégrant chacun des trois types de caractéristique spatiale au processus de classification. A chaque fois les indices spatiaux ont été extraits en utilisant le paramétrage optimal estimé précédemment. Chaque valeur est une moyenne des résultats obtenus quel que soit l'algorithme de réduction et le critère de sélection du nombre de bandes utilisés. En ce qui concerne GMM, l'algorithme a été lancé 15 fois pour chaque paramétrage. Le résultat est donc la précision moyenne (en rouge) plus ou moins l'écart-type (en bleu) sur ces 15 occurrences.

Sur le tableau 2.9 nous voyons que dans un contexte supervisé, l'inclusion d'attributs spatiaux dans le processus de classification n'apporte pas une plus-value conséquente à la précision des résultats. Si le jeu d'échantillons d'apprentissage est extrait aléatoirement au sein de l'ensemble des parcelles occupées par chaque classe (classifieur SVM1), l'emploi de coefficients de Haralick ou de profils morphologique permet d'améliorer les résultats de classification d'à peine 2% dans le meilleur des cas, du moins pour des images composés de textures à forte fréquence telles que MR et MR2. Sur ces même images, lorsque l'on opte pour un jeu d'échantillons d'apprentissage plus réaliste (classifieur SVM2), seul les profils morphologiques conservent leur capacité à améliorer la classification. L'utilisa-

tion de coefficients de Haralick et de spectres de texture a même tendance à dégrader les résultats, ce qui peut s'expliquer de la même façon que la perte de précision remarquée dans la section précédente lors de l'utilisation de voisinage trop étendu (voir figure 2.19). En effet, les profils morphologiques reflètent l'information spatiale à un niveau plus local, ce qui permet de limiter les effets de contours dont souffrent les deux autres types d'attribut. Le classifieur Mixmod n'étant pas limité à un apprentissage local comme peut l'être le classifieur SVM, il est également moins pénalisé par l'effet des contours, ce qui permet dans la plupart des cas à l'ensemble des types d'attributs spatiaux d'apporter une plus-value intéressante lorsqu'ils sont ajoutés à la classification. L'image CO étant pour sa part constituée principalement de textures complexes à basse fréquence, constituées de nombreux spectres différents, les coefficients de Haralick de même que les spectres de texture permettent une meilleure discrimination des classes, surtout dans un contexte non supervisé où le classifieur ne dispose d'aucune connaissance a priori lui indiquant quels spectres rassembler au sein de la même classe. Dans ce dernier cas, la présence de l'information spatiale devient indispensable, la classification d'attributs uniquement spectraux conduisant inévitablement à une confusion entre classes. Remarquons également que dans un contexte non supervisé, les écarts-types les plus faibles sont obtenus lorsque l'on utilise les profils morphologiques, ce qui tend à signifier que ces attributs rendent la classification plus robuste.

Les figures 2.20 et 2.21 apportent une illustration aux statistiques du tableau 2.9. En observant ces résultats en image, nous pouvons remarquer que les profils morphologiques se démarquent des deux autres types d'attributs spatiaux dans ce qu'ils apportent vis-à-vis de la classification purement spectrale. Avec un classifieur SVM ou Mixmod et quel que soit le jeu de donnée, les profils morphologiques ont un simple effet de régularisation permettant de limiter l'effet poivre et sel de la classification spectrale sans modifier la forme des zones occupées par chaque classe. Ce pouvoir de régularisation est suffisant pour des images composées de textures à haute fréquence telle que MR et MR2 mais est moins efficace pour mettre en valeur des textures à fréquence plus basse comme les textures urbaine et rurale de l'image CO, surtout dans le cadre d'une classification non supervisée. C'est d'ailleurs pour des images telles que CO que les coefficients de Haralick et les spectres de texture sont le plus efficaces. Le caractère global de l'algorithme Mixmod limite les confusions au niveau des contours entre zones tandis que la considération de voisinages plus étendus caractéristiques de ces attributs permet une meilleure homogénéisation de ces mêmes zones.

Enfin, le tableau 5.2 permet de comparer le temps de calcul associée à l'extraction de chaque type d'attribut spatial. Les coefficients de Haralick et spectres de textures sont les attributs les moins contraignants à ce niveau, notamment les seconds, pour lesquels l'augmentation de la taille du voisinage d'étude affecte très faiblement la complexité. Cette particularité est due au fait que les histogrammes que constituent les spectres de texture sont bien plus rapides à calculer

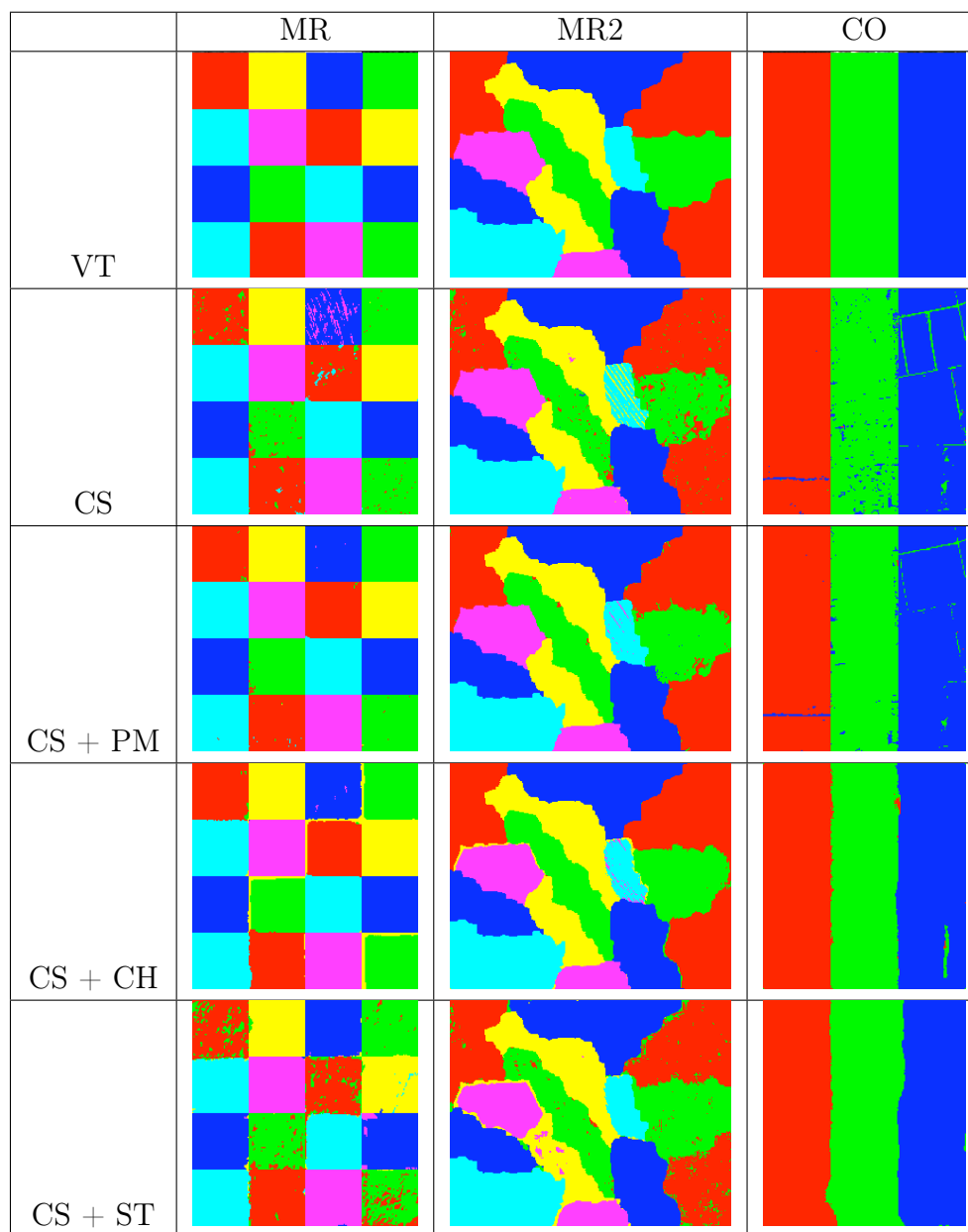


FIG. 2.20: Résultats de classification obtenus avec le classifieur SVM2

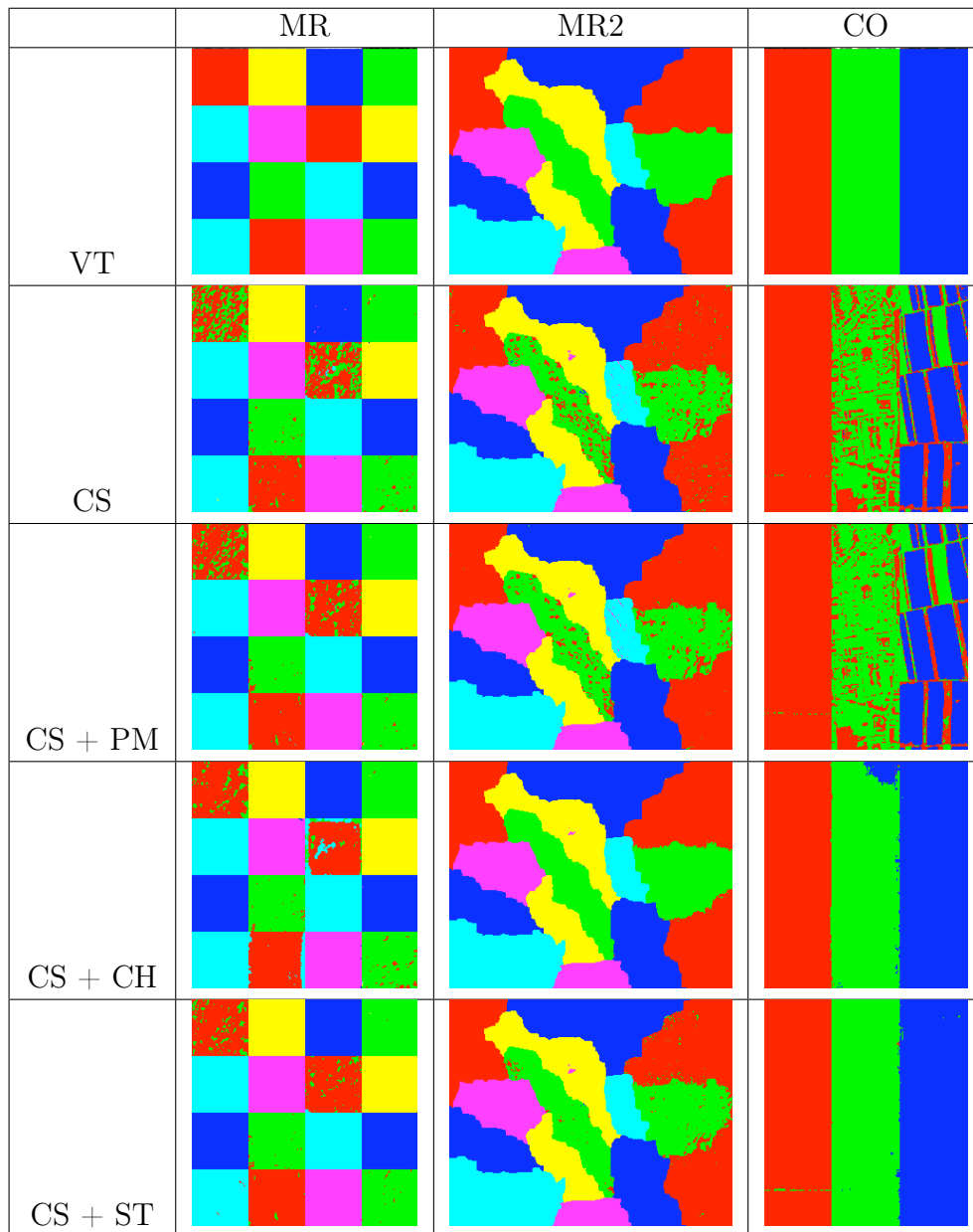


FIG. 2.21: Résultats de classification obtenus avec le classifieur GMM

Type d'attribut (paramètre)	Valeur	Nombre de bandes				
		2	3	4	5	10
CH (r)	11	0,53	0,78	1,05	1,3	2,63
	21	1,2	1,8	2,4	3,01	6,04
	31	2,3	3,45	4,6	5,78	11,51
PM (g)	2	1,1	1,91	2,29	3,26	7,48
	3	1,98	3,52	3,96	5,93	12,79
	4	2,85	5,15	5,65	8,74	18,25
ST (r)	11	0,73	1,13	1,48	1,89	3,68
	21	0,75	1,15	1,54	1,94	3,8
	31	0,76	1,13	1,59	1,99	3,93

TAB. 2.10: Temps d'extraction des attributs spatiaux (en secondes) en fonction du nombre de bandes et de la valeur du seul paramètre susceptible, pour chaque type d'attribut spatial, de modifier la complexité de l'extraction. Chaque valeur correspond au temps mis pour traiter un carré de 20×20 pixels. Les calculs ont été réalisés avec un processeur Intel(r) Xeon(r) QuadCore cadencé à 2,80 GHz et 6 Go de mémoire vive.

que les matrices de co-occurrence. Les profils morphologiques sont les attributs les plus longs à réaliser, avant tout parce que les opérateurs par reconstruction nécessitent parfois d'effectuer un grand nombre d'érosions et de dilatations pour chaque pixel de l'image.

6 Conclusion

Dans ce chapitre, nous avons présenté une méthode de classification vectorielle d'attributs spectraux/spatiaux dédiée aux images comprenant des textures. Les attributs spectraux sont extraits par projection des jeux de données initiaux sur des composantes ACP ou MNF, tandis que les attributs spatiaux sont extraits par le biais d'une étude texturale ou morphologique du voisinage de chaque pixel. A partir d'une partie des jeux de données présentés dans le chapitre précédent ainsi que de divers outils de classification, supervisés ou non, nous avons ensuite effectué une étude de sensibilité visant à déterminer dans quelle mesure l'information spatiale pouvait constituer une plus-value pour ce type d'approche de classification.

Le tableau 2.11 est une synthèse des atouts et inconvénients de chaque type d'attribut. Du point de vue des performances, il convient de séparer les jeux de données étudiés en deux groupes, ceux comportant des textures simples à forte fréquence d'une part (MR et MR2) et ceux présentant au contraire des textures complexes à faible fréquence (CO). Dans le premier cas, l'information spatiale est utile à titre de régularisation, aussi vaut-il mieux privilégier des attributs tels que

Type d'attribut		CH	PM	ST
Performances	MR/MR2	-	++	-
	CO	+	-	++
Robustesse		+	++	-
Complexité		+	-	++

TAB. 2.11: Atouts et inconvénients des divers types d'attributs en fonction de trois facteurs : performances, robustesse et complexité.

les profils morphologiques qui ne détérioreront pas la carte de classification au niveau des contours entre zones de classes différentes. Remarquons toutefois que si cette détérioration est fâcheuse d'un point de vue statistique, elle n'empêche pas les coefficients de Haralick et les spectres de texture de réaliser une régularisation parfois plus efficace que les profils morphologiques sur l'image d'étude, notamment dans un contexte non supervisé (cf. figures 2.20 et 2.21). Dans le second cas, les classes sont telles que l'information spatiale est indispensable, surtout lorsqu'on ne dispose pas de vérité terrain. Etant donnée la faible fréquence des textures, l'étude d'un voisinage plus étendu est nécessaire, ce qui rend le choix d'attributs tels que les coefficients de Haralick ou les spectres de texture plus judicieux.

En matière de robustesse, nous considérons d'une part l'impact du paramétrage de chaque type d'attribut spatial sur la précision de la classification et d'autre part l'homogénéité ou non des résultats obtenus avec le modèle de mélange gaussien. En ce qui concerne l'impact du paramétrage, les profils morphologiques sont clairement avantagés car ils dépendent d'un unique paramètre, le niveau de granulométrie, qui à sa valeur minimale testée de 2 donne des résultats similaires à un niveau de granulométrie plus élevé quel que soit le jeu de donnée considéré. C'est également ce type d'attribut qui se révèle le plus robuste dans un contexte non supervisé puisque l'écart-type de la précision obtenue avec le classifieur GMM est le plus faible, là encore quelle que soit l'image étudiée.

En termes de temps de calcul, ce sont par contre les spectres de textures qui se révèlent les plus efficaces, notamment car leur extraction fait intervenir des histogrammes dont la création et la manipulation est très rapide. A l'inverse, les opérateurs par reconstruction des profils morphologiques implique l'application itérative d'opérateurs morphologiques jusqu'à idempotence, ce qui explique le fait qu'ils soient plus long à réaliser.

Pour conclure, il paraît difficile de ne retenir qu'un seul type d'attribut spatial parmi les trois que nous avons analysé. Nous pouvons toutefois remarquer que les profils morphologiques se démarquent à la fois en terme de robustesse et de performances sur des jeux de données nécessitant simplement une régularisation spatiale. Lorsque l'utilisation de l'information spatiale est primordiale, comme dans le cadre d'images telles que CO, ce sont plutôt les spectres de texture qu'il

CHAPITRE 2. CLASSIFICATION VECTORIELLE DE CARACTÉRISTIQUES SPECTRALES ET SPATIALES

convient de privilégier car malgré une robustesse assez faible, leurs performances en termes de temps de calcul comme de précision de la classification sont les plus intéressantes.

Chapitre 3

Conditional Random Field (CRF) adapté aux données hyperspectrales

1 Introduction

Les indices spatiaux calculés dans le précédent chapitre analysent l'information contextuelle relative à chaque pixel sur de vastes voisinages, mais l'algorithme utilisé au terme du processus pour classifier l'image d'indices spectraux/spatiaux n'emploie aucun modèle contextuel pour attribuer un label aux pixels. Dans ce chapitre, nous introduisons un processus de classification par Maximum A Posteriori s'appuyant sur une modélisation markovienne des données et intégrant de ce fait les relations de voisinage à l'intérieur même de sa structure.

Nous utilisons un modèle de type *Conditionnal Random Field* (CRF) défini par Zhong et al. [58]. L'objectif est d'optimiser ce modèle vis-à-vis de l'imagerie hyperspectrale en intégrant une phase de réduction de dimension préalable et en modifiant le mode de représentation d'une relation de voisinage entre deux pixels. Ces deux modifications nous permettent de réduire drastiquement le nombre de paramètres nécessaires au fonctionnement de l'algorithme.

Enfin nous analysons l'apport de cette approche pour les divers scénarios évoqués dans le chapitre 1.

2 Modèles markoviens

Un champ de Markov est un modèle probabiliste de représentation des données (par exemple une image) intégrant la notion de voisinage entre individus. Soit une image que l'on assimile à un champ \mathbf{X} constitué d'un ensemble de variables aléatoires \mathbf{X}_s . Chaque variable aléatoire est associée à un site $s \in S$ tel que S représente la grille de pixels et s un de ses éléments. S est caractérisé par un système de voisinage permettant de définir les relations contextuelles entre sites. Dans le cas d'une grille de pixels carrés, le système utilisé sera la 4-connexité ou

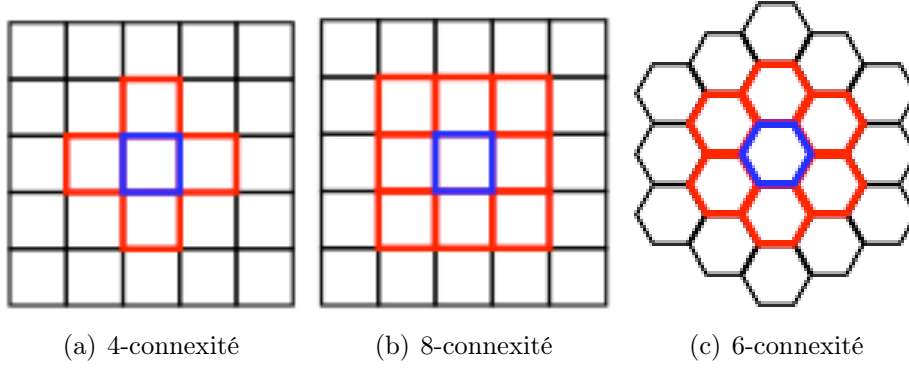


FIG. 3.1: Divers types de voisinage (en rouge, les voisins du pixel bleu)

la 8-connexité tandis qu'une grille hexagonale (d'une utilisation marginale) sera associée à un système en 6-connexité (voir figure 3.1).

Soit V_s l'ensemble des sites appartenant au voisinage de s . Le champ \mathbf{X} est un champ de Markov si et seulement si :

$$P(\mathbf{X}_s | (\mathbf{X}_t)_{t \in S, t \neq s}) = P(\mathbf{X}_s | (\mathbf{X}_t)_{t \in S, t \in V_s}) \quad (3.1)$$

Concrètement, la probabilité en un site sachant l'ensemble du champ est équivalente à la probabilité en ce même site sachant uniquement son voisinage. Cela permet de modéliser à la fois les relations contextuelles à courte et à longue portée (par transitivité). Le théorème de Hammersley-Clifford établit alors qu'un champ de Markov peut être assimilé à un champ de Gibbs de loi :

$$P(\mathbf{X} = \mathbf{x}) = \frac{1}{Z(\mathbf{x}, \Lambda)} \exp(U(\mathbf{x}, \Lambda)) \quad (3.2)$$

où \mathbf{x} est une réalisation du champ \mathbf{X} , $Z(\mathbf{x}, \Lambda)$ est une constante de normalisation appelée fonction de partition et $U(\mathbf{x}) = \sum_{c \in C} U_c(\Lambda)$. C est l'ensemble des cliques associées au type de voisinage utilisé (une clique étant un ensemble de sites dont tous les éléments sont voisins), $U_c(\Lambda)$ un potentiel dépendant de la clique c et Λ l'ensemble de paramètres du champ markovien.

Tupin [53] distingue deux familles d'applications liées aux champs de Markov : les applications de synthèse et d'analyse de texture d'un côté, celles de segmentation, de restauration et de classification de l'autre. Pour ces dernières, on considère des données incomplètes en introduisant deux champs aléatoires \mathbf{X} et \mathbf{Y} au lieu d'un. Le champ \mathbf{X} modélise les observations, c'est à dire l'image à classifier, tandis que le champ \mathbf{Y} modélise la solution, à savoir la carte de classification. Le problème se définit en utilisant la formulation de Bayes :

$$P(\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{x}) = \frac{P(\mathbf{X} = \mathbf{x} | \mathbf{Y} = \mathbf{y}) P(\mathbf{Y} = \mathbf{y})}{P(\mathbf{X} = \mathbf{x})} \quad (3.3)$$

La réalisation du champ \mathbf{X} étant connue, le dénominateur est constant et la recherche d'une solution optimale fait intervenir uniquement le terme de vraisemblance $P(\mathbf{X} = \mathbf{x} | \mathbf{Y} = \mathbf{y})$ et le terme d'a priori $P(\mathbf{Y} = \mathbf{y})$.

Un modèle simple permettant d'optimiser la probabilité a posteriori consiste à assimiler le terme d'a priori à un champ de Gibbs tout en considérant que les spectres observés de \mathbf{x} sont indépendants les uns des autres conditionnellement à leur label. Cela revient à réécrire la vraisemblance globale telle que :

$$P(\mathbf{X} = \mathbf{x} | \mathbf{Y} = \mathbf{y}) = \prod_{s \in S} P(\mathbf{X}_s = \mathbf{x}_s | Y_s = y_s) \quad (3.4)$$

La distribution a posteriori est ainsi formulée de la manière suivante :

$$\begin{aligned} P(\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{x}) &\propto P(\mathbf{Y} = \mathbf{y}) P(\mathbf{X} = \mathbf{x} | \mathbf{Y} = \mathbf{y}) & (3.5) \\ &\propto \frac{1}{Z(\mathbf{y}, \Lambda)} \exp(-U(\mathbf{y}, \Lambda)) \prod_{s \in S} P(\mathbf{X}_s = \mathbf{x}_s | Y_s = y_s) & (3.6) \end{aligned}$$

3 Conditionnal Random Field (CRF)

L'utilisation d'un champ markovien tel que décrit dans la section précédente implique une modélisation implicite de la vraisemblance $P(\mathbf{X} = \mathbf{x} | \mathbf{Y} = \mathbf{y})$. En comparaison, les CRF définis par Lafferty et al. [28] considèrent les propriétés markoviennes de \mathbf{Y} conditionnellement à \mathbf{X} et modélisent de ce fait directement la probabilité a posteriori comme une distribution de Gibbs :

$$P(\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{x}, \Lambda) = \frac{1}{Z(\mathbf{x}, \Lambda)} \exp(U(\mathbf{y}, \mathbf{x}, \Lambda)) \quad (3.7)$$

En particulier, si l'on ne considère que des cliques composées d'au maximum deux éléments :

$$P(\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{x}, \Lambda) = \frac{1}{Z(\mathbf{x}, \Lambda)} \exp \left\{ \sum_{s \in S} \phi_s(y_s, \mathbf{x}, \mathbf{w}) + \sum_{s \in S} \sum_{t \in V_s} \xi_{st}(y_s, y_t, \mathbf{x}, \mathbf{v}) \right\} \quad (3.8)$$

où $\phi_s(\cdot)$ et $\xi_{st}(\cdot)$ sont respectivement les potentiels de clique d'ordre 1 et 2, associés aux paramètres \mathbf{w} et \mathbf{v} tels que $\Lambda = \{\mathbf{w}, \mathbf{v}\}$. Afin de limiter la complexité algorithmique, nous ne considérons que les cliques d'ordre 1 (le pixel s uniquement) et les cliques d'ordre 2 associées au système de voisinage en 4-connexité.

3.1 Calcul des paramètres du modèle

Afin de modéliser les potentiels de cliques, nous utilisons un modèle de régression logistique multinomiale (MLR) [27], que Zhong et al. [58] considèrent plus à même de modéliser des données hyperspectrales qu'un modèle gaussien classique.

En effet, ce modèle ne requiert pas des données qu'elles soient distribuées selon une loi normale ni linéairement séparables vis-à-vis des classes considérées dans le processus de classification. Un potentiel de clique d'ordre 1 peut ainsi être écrit :

$$\phi_s(y_s, \mathbf{x}, \mathbf{w}) = \log P(y_s = k | \mathbf{x}, \mathbf{w}) \quad (3.9)$$

où k est une étiquette de classe et :

$$P(y_s = k | \mathbf{x}, \mathbf{w}) = \frac{\exp(\mathbf{w}_k^T \mathbf{x}_s)}{\sum_{i=1}^K \exp(\mathbf{w}_i^T \mathbf{x}_s)} \quad (3.10)$$

\mathbf{w}_k est le vecteur paramètre de la classe k de taille M , où M est le nombre de composantes de l'image à classifier. \mathbf{w} correspond donc au vecteur de taille MK produit par la concaténation des vecteurs $(\mathbf{w}_f)_{k \in \{1, \dots, K\}}$. Le vecteur de paramètres \mathbf{w}_L associé à la dernière classe étant totalement dépendant des précédents, ce dernier est fixé à $\mathbf{0}$, ce qui revient à dire que :

$$P(y_s = k | \mathbf{x}, \mathbf{w}) = \begin{cases} \frac{\exp(\mathbf{w}_k^T \mathbf{x}_s)}{1 + \sum_{i=1}^{K-1} \exp(\mathbf{w}_i^T \mathbf{x}_s)}, & k < K \\ \frac{1}{1 + \sum_{i=1}^{K-1} \exp(\mathbf{w}_i^T \mathbf{x}_s)}, & k = K \end{cases} \quad (3.11)$$

Parallèlement, les potentiels de clique d'ordre 2 sont formulés de la sorte :

$$\xi_{st}(y_s, y_t, \mathbf{x}, \mathbf{v}) = \log P(y_s = k, y_t = l | \gamma_{st}(\mathbf{x}), \mathbf{v}) \quad (3.12)$$

tels que k et l sont des étiquettes de classe, $\gamma_{st}(\mathbf{x})$ la concaténation des éléments des vecteurs pixel \mathbf{x}_s et \mathbf{x}_t et :

$$P(y_s = k, y_t = l | \gamma_{st}(\mathbf{x}), \mathbf{v}) = \frac{\exp(\mathbf{v}_{kl}^T \gamma_{st}(\mathbf{x}))}{\sum_{i=1}^K \sum_{j=1}^L \exp(\mathbf{v}_{ij}^T \gamma_{st}(\mathbf{x}))} \quad (3.13)$$

\mathbf{v}_{kl} est le vecteur paramètre de taille M du couple de classes (k, l) tel que le vecteur \mathbf{v} correspondant serait de taille MK^2 . L'image IP étant constituée de 197 bandes spectrales pour 9 classes, le nombre de paramètres nécessaire à la création du modèle s'élèverait à 15957, soit plus que le nombre d'échantillons disponibles pour l'apprentissage. Afin de limiter le nombre de paramètres, seuls $K + 1$ vecteurs \mathbf{v}_{kl} sont ainsi considérés :

- K vecteurs \mathbf{v}_{kk} pour les cas où $k = l$
- un unique vecteur $\mathbf{v}_{k \neq l}$ pour l'ensemble des cas où $k \neq l$

Pour reprendre l'exemple de l'image IP, cette simplification nous permet de passer de 15957 paramètres à 1970. C'est de plus une simplification raisonnable si l'on part de l'hypothèse que l'image à classifier se compose majoritairement de zones homogènes assez vastes. En effet dans ce cas, la plupart des couples de pixels

voisins appartiendront à la même classe et les couples de classes différentes seront marginaux. \mathbf{v} devient donc un vecteur de taille $M(K + 1)$ tel que $\mathbf{v}_{k \neq l}$ est fixé à $\mathbf{0}$ de la même façon que le vecteur \mathbf{w}_L mentionné précédemment, ce qui revient à dire que :

$$P(y_s = k, y_t = l | \gamma_{st}(\mathbf{x}), \mathbf{v}) = \begin{cases} \frac{\exp(\mathbf{v}_{kk}^T \gamma_{st}(\mathbf{x}))}{\tau + \sum_{i=1}^K \exp(\mathbf{v}_{ii}^T \gamma_{st}(\mathbf{x}))}, & k = l \\ \frac{\tau}{\tau + \sum_{i=1}^K \exp(\mathbf{v}_{ii}^T \gamma_{st}(\mathbf{x}))}, & k \neq l \end{cases} \quad (3.14)$$

tel que $\tau = K(K + 1)$, le nombre de couples (k, l) tels que $k \neq l$.

Afin d'estimer l'ensemble de paramètres optimaux Λ , l'apprentissage classique par maximisation de la vraisemblance consiste à maximiser la log-vraisemblance suivante :

$$\begin{aligned} L(\Lambda) &= \log P(\hat{\mathbf{y}} | \hat{\mathbf{x}}, \Lambda) \\ &= \sum_{c \in \hat{C}} U_c(\hat{\mathbf{y}}_c, \hat{\mathbf{x}}_c, \Lambda) - \log Z(\hat{\mathbf{x}}, \Lambda) \end{aligned} \quad (3.15)$$

$\{\hat{\mathbf{x}}, \hat{\mathbf{y}}\} = \{\hat{\mathbf{x}}_c, \hat{\mathbf{y}}_c\}_{c \in \hat{C}}$ est l'ensemble des échantillons d'apprentissage, tel que \hat{C} contient les cliques associées à cet ensemble. $\hat{\mathbf{x}}_c$ représente les échantillons appartenant à la clique c et $\hat{\mathbf{y}}_c$ les labels correspondants. La fonction de partition $Z(\hat{\mathbf{x}}, \Lambda)$ étant bien trop lourde à calculer, l'estimation exacte du Λ optimal est la plupart du temps infaisable en pratique. Il existe alors deux solutions. La première revient à approximer la fonction de partition globalement, par exemple avec des méthodes de type Monte Carlo [18] qui peuvent cependant s'avérer prohibitives en temps de calcul et difficiles à paramétrer. Une seconde possibilité consiste à estimer Λ localement par l'intermédiaire d'une pseudo-vraisemblance [5] en utilisant une fonction objectif approximée ne nécessitant pas de calculs globaux. Cette pseudo-vraisemblance est de la forme :

$$L_{PV}(\Lambda) = \sum_{c \in \hat{C}} \log \frac{\exp(U_c(\hat{\mathbf{y}}_c, \hat{\mathbf{x}}_c, \Lambda))}{\sum_{\mathbf{y}_c} \exp(U_c(\mathbf{y}_c, \hat{\mathbf{x}}_c, \Lambda))} \quad (3.16)$$

Dans cette étude, nous utilisons une variante de la pseudo-vraisemblance, introduite par Sutton et al. [46], la pseudo-vraisemblance par morceaux. L'idée de départ d'une telle variante est que si chaque potentiel $U_c(\hat{\mathbf{y}}_c, \hat{\mathbf{x}}_c, \Lambda)$ peut prédire précisément $\hat{\mathbf{y}}_c$ à partir de $\hat{\mathbf{x}}_c$ par lui-même, alors la prédiction sur l'ensemble de l'image devrait être également précise. Soit :

$$\begin{aligned} L_{PW}(\mathbf{w}, \mathbf{v}) &= \sum_{s \in C_1} \frac{\exp\{\phi_s(\hat{y}_s, \hat{\mathbf{x}}, \mathbf{w})\}}{\sum_{x_s} \exp\{\phi_s(y_s, \hat{\mathbf{x}}, \mathbf{w})\}} + \sum_{(s,t) \in C_2} \frac{\exp\{\phi_s(\hat{y}_s, \hat{y}_t, \hat{\mathbf{x}}, \mathbf{v})\}}{\sum_{x_s, x_t} \exp\{\phi_s(y_s, y_t, \hat{\mathbf{x}}, \mathbf{v})\}} \\ &= Q_{\mathbf{w}} + Q_{\mathbf{v}} \end{aligned} \quad (3.17)$$

une réécriture de $L_{PV}(\Lambda)$ où les sommes de potentiels relatives à C_1 et C_2 , les ensembles des cliques d'ordre 1 et 2, sont séparées en deux termes $Q_{\mathbf{w}}$ et $Q_{\mathbf{v}}$. Si

l'on considère que $L_{PW}(\mathbf{w}, \mathbf{v})$ est une pseudo-vraisemblance par morceau, trouver (\mathbf{w}, \mathbf{v}) qui la maximise revient à trouver \mathbf{w} qui maximise $Q_{\mathbf{w}}$ d'une part et \mathbf{v} qui maximise $Q_{\mathbf{v}}$ d'autre part. En pratique, ce procédé n'est pas optimal mais facilite la recherche des paramètres.

Prenons tout d'abord le cas de $Q_{\mathbf{w}}$. Les paramètres optimaux sont estimés au moyen de la méthode IRLS [58] (pour *Iteratively Reweighted Least Squares*). Les paramètres \mathbf{w} sont mis à jour itérativement de la façon suivante :

$$\mathbf{w}^{(r+1)} = \mathbf{w}^{(r)} - \mathbf{B}^{-1}g(\mathbf{w}^{(r)}) \quad (3.18)$$

où $\mathbf{w}^{(r)}$ est le vecteur de paramètres optimal à l'itération r et $g(\mathbf{w}^{(r)})$ le vecteur gradient de $Q_{\mathbf{w}}$ à cette même itération. \mathbf{B} est la borne inférieure au sens de l'ordre des matrices semi-définies positives de la matrice Hessienne de $Q_{\mathbf{w}}$ telle que :

$$\mathbf{B} \equiv -\frac{1}{2}[\mathbf{I} - \frac{\mathbf{1}\mathbf{1}^T}{M}] \otimes \sum_{s \in C_1} \hat{\mathbf{x}}_s \hat{\mathbf{x}}_s^T \quad (3.19)$$

\mathbf{I} étant la matrice identité de taille $K - 1$ et $\mathbf{1}$ un vecteur de taille correspondante ne comportant que des 1. Déterminer \mathbf{B} consiste ainsi à effectuer le produit tensoriel (\otimes : produit de chaque composante du premier tenseur, ici une matrice, par chaque composante du second) de deux matrices carrées, la première de taille $K - 1$ et la seconde de taille M , ce qui donne une autre matrice carrée de taille $M(K - 1)$. $g(\mathbf{w})$ correspond au vecteur gradient de $Q_{\mathbf{w}}$:

$$\mathbf{g}(\mathbf{w}) = \sum_{s \in C_1} (\hat{\mathbf{y}}'_s - \mathbf{p}_s(\mathbf{w})) \otimes \hat{\mathbf{x}}_s \quad (3.20)$$

avec $\mathbf{p}_s(\mathbf{w}) = [p_s^1, \dots, p_s^{K-1}]^T$, sachant que $p_s^k = P(\hat{y}_s = k | \hat{\mathbf{x}}, \mathbf{w})$. Ce vecteur est mis à jour à chaque itération. $\hat{\mathbf{y}}'_s$ est pour sa part un vecteur de taille $K - 1$ tel quel $\hat{y}'_s^k = 1$ si l'échantillon $\hat{\mathbf{x}}_s$ appartient à la classe k et $\hat{y}'_s^k = 0$ sinon. Pour le

calcul de $Q_{\mathbf{v}}$, le procédé est le même à quelques détails près. Tout d'abord dans l'équation 3.19, $\hat{\mathbf{x}}_s$ est remplacé par $\gamma_{st}(\hat{\mathbf{x}})$ et la matrice \mathbf{I} est désormais de taille K car il y a désormais $K + 1$ classes. Ceci nous donne une matrice B de taille $2MK$ qui peut être singulière, d'où l'utilisation dans ce cas-ci d'une fonction pseudo-inverse. Enfin, le vecteur $\mathbf{p}_s(\mathbf{w})$ utilisé dans le calcul du gradient de l'équation 3.20 est également modifié de sorte à contenir K éléments, $\mathbf{p}_s(\mathbf{w}) = [p_s^1, \dots, p_s^K]^T$.

3.2 Modes de représentation alternatifs des cliques d'ordre deux

Le modèle décrit précédemment représente un couple de pixels voisins s et t en concaténant les vecteurs \mathbf{x}_s et \mathbf{x}_t correspondant. Ce procédé est problématique car il triple le nombre de paramètres à estimer vis-à-vis d'un modèle qui ne

considérerait que les cliques d'ordre 1. Afin d'éviter cette surcharge de paramètres, nous proposons deux autres types de représentation.

Le premier, basé sur le calcul de l'angle spectral entre s et t , réduit le vecteur \mathbf{v} de $2 \times M \times K$ paramètres à K paramètres seulement. Un potentiel d'ordre 2 est alors écrit de la façon suivante :

$$\xi'_{st}(y_s, y_t, \mathbf{x}, \mathbf{v}') = \log P(y_s = k, y_t = l | \nu_{st}(\mathbf{x}), \mathbf{v}') \quad (3.21)$$

tel que $\nu_{st}(\mathbf{x}) = \frac{\mathbf{x}_s \cdot \mathbf{x}_t}{\|\mathbf{x}_s\| \cdot \|\mathbf{x}_t\|}$ et :

$$P(y_s = k, y_t = l | \nu_{st}(\mathbf{x}), \mathbf{v}') = \begin{cases} \frac{\exp(\mathbf{v}'_{kk} \nu_{st}(\mathbf{x}))}{\tau + \sum_{i=1}^K \exp(\mathbf{v}'_{ii} \nu_{st}(\mathbf{x}))}, & k = l \\ \frac{\tau}{\tau + \sum_{i=1}^K \exp(\mathbf{v}'_{ii} \nu_{st}(\mathbf{x}))}, & k \neq l \end{cases} \quad (3.22)$$

Le second mode de représentation alternatif considère quant à lui les valeurs du pixel centre s et de ses voisins dans une direction donnée (cf figure 3.2) prises sur un certains nombre de composantes principales de l'image.

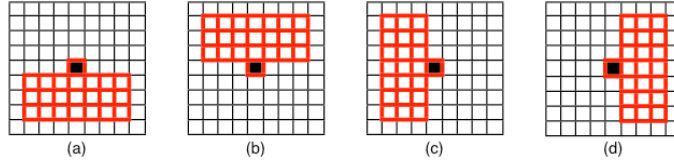


FIG. 3.2: Voisins utilisés pour les cliques d'ordre deux inférieure (a), supérieure (b), gauche (c) et droite (d).

Pour une portée de voisinage p donnée, le vecteur \mathbf{v} est cette fois réduit à $p(2p + 1) \times K$ paramètres.

3.3 Classification de l'image

Une fois les paramètres du modèle estimés, il reste à définir la carte de classification maximisant la probabilité a posteriori de l'équation 3.8, tout en gardant à l'esprit que les paramètres \mathbf{w} et \mathbf{v} ont été estimés indépendamment. Cet apprentissage par morceaux peut être la cause d'un certain nombre de problèmes relatifs à la combinaison de plusieurs modèles corrélés [21]. Afin de limiter leur impact, l'expression de la probabilité a posteriori est réécrite de cette façon :

$$P(\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{x}, \Lambda) = \frac{1}{Z} \exp \left\{ \lambda \sum_{s \in S} \phi_s(y_s, \mathbf{x}, \hat{\mathbf{w}}) + (1 - \lambda) \sum_{s \in S} \sum_{t \in V_s} \xi_{st}(y_s, y_t, \mathbf{x}, \hat{\mathbf{v}}) \right\} \quad (3.23)$$

où $\hat{\mathbf{w}}$ et $\hat{\mathbf{v}}$ sont les vecteurs de paramètres optimaux estimés pendant la phase d'apprentissage, et λ un facteur permettant de pondérer l'influence des potentiels de clique d'ordre 1 vis-à-vis des potentiels de clique d'ordre 2.

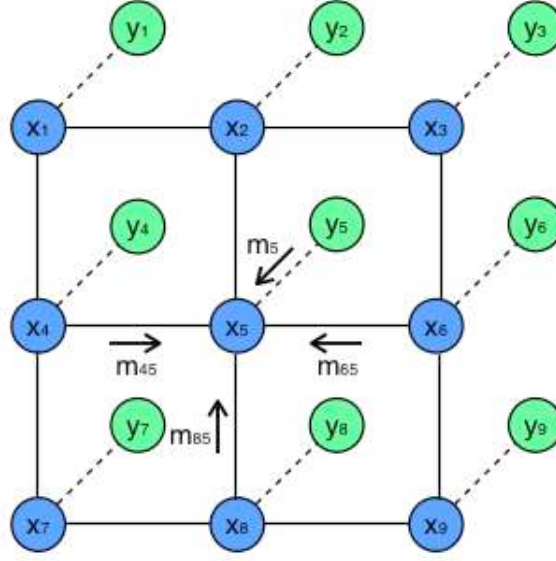


FIG. 3.3: Exemple de graphe G associé à une image 3×3 . Les noeuds verts sont associés aux données observées et les noeuds bleus aux labels.

La maximisation de l'équation 3.23 est réalisée par l'intermédiaire de l'algorithme LBP (*Loopy Belief Propagation*). L'image est tout d'abord assimilée à un graphe G qui à chaque site s associe deux noeuds, un noeud de données correspondant au pixel \mathbf{x}_s et un noeud de classe correspondant au label y_s associé à ce pixel (voir figure 3.3). L'objectif de l'algorithme est de définir la carte de classification \mathbf{y} qui maximise la probabilité a posteriori que l'on réécrit de la manière suivante :

$$P(\mathbf{y}|\mathbf{x}, \Lambda) \propto \prod_{s \in S} \left(\zeta_s(y_s, \mathbf{x}) \prod_{t \in V_s} \zeta_{st}(y_s, y_t, \mathbf{x}) \right) \quad (3.24)$$

avec :

$$\zeta_s(y_s, \mathbf{x}) = \exp\{\lambda \phi_s(y_s, \mathbf{x}, \hat{\mathbf{w}})\} \quad (3.25)$$

$$\zeta_{st}(y_s, y_t, \mathbf{x}) = \exp\{\xi_{st}(y_s, y_t, \mathbf{x}, \hat{\mathbf{v}})\} \quad (3.26)$$

L'étiquette d'un pixel est déterminée par un ensemble de "messages" : un message spectral relatif à la clique d'ordre 1 formé par ce pixel et quatre messages spatiaux relatifs aux cliques d'ordre 2 formées par ce même pixel et ses voisins en 4-connexité. Soient $m_s(y_s)$ le message envoyé par le noeud de donnée \mathbf{x}_s au noeud de label y_s associé (message spectral), $m_{st}(y_t)$ le message envoyé par le noeud de label y_t au noeud de label voisin y_s (message spatial) et $b_s(y_s)$ la valeur de confiance au noeud y_s . Ces trois variables sont vectorielles et comprennent autant de composantes qu'il y a de classes. Le processus comporte trois étapes (voir algorithme 1), l'initialisation (lignes 1 à 2), l'optimisation des messages spatiaux (lignes 3 à 9) et enfin l'étiquetage (lignes 10 à 15).

Algorithme 1 LBP

```

1: Initialisation des  $m_{st}(y_t)$  en tant que distributions uniformes
2:  $m_s(y_s) \leftarrow \zeta_s(y_s, \mathbf{x})$ 
3: pour  $t$  de 0 à  $T-1$  faire
4:   pour chaque site  $s$  faire
5:     pour chaque voisin  $t$  de  $s$  faire
6:        $m_{st}^{r+1}(y_t) \leftarrow \rho_1 \max_{y_s} \{ \zeta_{st}(y_s, y_t, \mathbf{x}) \times m_s(y_s) \times \prod_{u \in \eta_s \setminus t} m_{us}^r(y_s) \}$ 
7:     fin pour
8:   fin pour
9: fin pour
10: pour chaque site  $s$  faire
11:   pour chaque valeur de  $y_s$  possible faire
12:      $b_s(y_s) \leftarrow \rho_2 m_s(y_s) \prod_{u \in \eta_s} m_{us}(y_s)$ 
13:      $y_s^{MAP} \leftarrow \max_{y_s} b_s(y_s)$ 
14:   fin pour
15: fin pour

```

ρ_1 et ρ_2 sont des variables de normalisation et T le nombre d'itérations du processus d'optimisation des messages spatiaux. L'ensemble des messages spectraux et spatiaux permet de calculer une valeur de confiance entre un pixel et chacune des classes, la classe retenue étant celle à laquelle est associée la plus forte confiance.

4 Résultats

Nous présentons dans la section qui suit les résultats obtenus avec le processus de classification par *Conditional Random Field*. Nous commençons par nous intéresser au cas particulier de l'image IP, une des deux images utilisées par Zhong et al. [58], afin de comparer nos résultats aux leurs. Nous évaluons ensuite l'efficacité de cette seconde approche de classification sur les trois autres jeux de données considérés dans cette étude.

4.1 Comparaison à l'existant

Afin de comparer les résultats obtenus à ceux présentés dans [58], nous avons tenu à effectuer la classification de l'image IP dans des conditions similaires. Nous utilisons ainsi un paramétrage identique des algorithmes IRLS et LBP. Le nombre d'itérations de l'algorithme IRLS est notamment fixé à 200 tandis que celui de l'algorithme LBP est fixé à 20. Avant la classification, chacune des bandes de l'image est centrée puis réduite. L'apprentissage de l'algorithme IRLS est effectué sur un ensemble d'échantillons choisis aléatoirement au sein de la

vérité terrain et représentant 50% des effectifs de cette dernière, le reste étant utilisé ultérieurement pour mesurer la précision des résultats.

Nous considérons cinq types de classification. Le premier (M1) est en tout point similaire à celui utilisé dans l'article [58] : application de l'algorithme sur des données non réduites et utilisation du même mode de représentation des cliques d'ordre deux. Pour le second type de classification (M2), nous utilisons en entrée de l'algorithme une image réduite (dix premières composantes ACP). Les classifications M3 et M4 sont appliqués sur les mêmes jeux de données que M1 et M2 respectivement, mais les cliques d'ordre deux sont cette fois représentées par des angles spectraux (cf section 3.2). Enfin, le dernier type de classification (M5) utilise des données non réduites tout en considérant des cliques d'ordre deux représentées par des voisinages orientés. Dans chaque cas, nous avons testé des classifieurs LBP associés à diverses valeurs de λ ainsi qu'un classifieur MAP calculant le maximum a posteriori du seul modèle MLR associé aux cliques d'ordre 1, afin de mesurer l'impact de l'information spatiale sur le résultat de classification final. Enfin, ces résultats sont comparés à ceux d'un classifieur SVM paramétré de la même façon que dans le chapitre 2.

Méthode	Valeur de λ							MAP
	0,01	0,1	0,3	0,5	0,7	0,9	0,99	
M1	86,4	86,2	86,9	88,2	87,4	84,7	82,2	82
M2	76,1	75,8	76,9	78,9	79,2	77,3	74,4	74
M3	26,8	56,3	82	84,9	83,54	82,6	82	82
M4	26,4	48,8	68,4	72,9	74,8	74,5	74	74,02
M5	85	85,1	85,4	85,7	85,5	83,7	82,1	81,95
SVM	87,5							

TAB. 3.1: Pourcentage de bonne classification des résultats obtenus sur l'image IP avec les diverses méthodes en fonction du paramètre λ .

Le tableau 3.1 montre tout d'abord que la précision de classification obtenue avec nos travaux est sensiblement plus faible que celle présentée dans [58], aussi bien avec les classifieurs LBP (94,9%) et MAP (85,2%) qu'avec le classifieur SVM (94,2%). Cet écart est certainement dû au fait que l'apprentissage des deux algorithmes n'a pas été effectué sur le même jeu d'échantillons. Ensuite nous pouvons remarquer d'une part que la réduction de dimension du jeu de données entraîne une forte diminution de la précision de classification (jusqu'à près de 10%), et d'autre part que les deux variantes de représentation des cliques d'ordre 2 sont moins efficaces que le système de représentation originel. Toutefois, si la méthode M1 semble montrer une robustesse acceptable lorsque l'apprentissage est effectué avec 50% des échantillons de la vérité terrain, la figure 3.4 révèle que ce n'est pas le cas avec seulement 10% de ces mêmes échantillons, une limitation nécessaire qui nous permet de nous approcher plus près des conditions de classification

supervisée réelles. Dans ce cadre, on remarque en effet que la précision de classification est inversement proportionnelle au nombre d'itérations de l'algorithme IRLS permettant d'optimiser les paramètres du modèle.

Ce phénomène peut s'expliquer de deux façons. La construction d'un modèle statistique par apprentissage supervisé nécessite la sélection d'un ensemble d'échantillons connus a priori et parfaitement représentatifs des diverses classes en présence. Ne pas remplir cette condition fondamentale reviendrait à bâtir un modèle de plus en plus biaisé, apte à reconnaître des classes sensiblement différentes de celles véritablement présentes dans l'image. Or, rappelons que les classes de l'image IP sont très hétérogènes, aussi bien d'un point de vue spectral que spatial, ce qui explique qu'une forte proportion d'échantillons d'apprentissage soit nécessaire pour bâtir un modèle représentatif. Parallèlement, le grand nombre de paramètres associés au modèle, au-delà de ralentir nettement la convergence de ces derniers (cf. figures 3.5 et 3.6), limite également le potentiel de généralisation de chacune des classes considérées.

Pourtant, si ces problèmes de robustesse disparaissent en diminuant le nombre de paramètres, soit par réduction de dimension du jeu de données (méthodes M2 et M4), soit par modification de la représentation des cliques d'ordre 2 dans le modèle (méthodes M3 et M4), cela coïncide avec un affaiblissement notable de la précision de classification. Dans le cas de la réduction de dimension, cela vient probablement du fait que dans certains cas, les premières bandes ACP ne reflètent pas toute l'information utile (une hypothèse sur laquelle nous reviendrons lors de l'analyse des autres jeux de données), tandis que dans le second cas, cela prouve simplement que l'angle spectral entre spectres immédiatement voisins est moins discriminant que la concaténation des spectres eux-même, au moins en ce qui concerne les classes de l'image IP. La méthode M5 est aussi peu robuste que la méthode M1 tout en présentant des résultats sensiblement moins bons, c'est pourquoi nous n'y ferons plus référence durant la suite de l'étude.

Le paramètre λ est un facteur important du processus car il définit dans quelle mesure chaque pixel dépendra de ses voisins dans la détermination de son étiquette de classe. Ce paramètre est d'autant plus important pour les méthodes M3 et M4 que leur mode de représentation des cliques d'ordre 2 à base d'angles spectraux ne fait intervenir que l'information spatiale. Les très mauvais résultats obtenus avec ces méthodes lorsque λ est petit prouvent qu'en ce qui concerne l'image IP, l'information spatiale ne joue qu'un rôle complémentaire de l'information spectrale, cette dernière restant le moyen le plus sûr de discriminer les diverses classes en présence. Si l'on sélectionne un ensemble d'échantillons d'apprentissage assez conséquent pour garantir une certaine robustesse à l'algorithme, la valeur optimale de λ oscille alors entre 0.5 et 0.7, ce qui implique de donner légèrement plus de poids à l'information spectrale.

Notons enfin que le gain de précision de classification entraîné par la prise en compte de l'information spatiale dépend lui aussi beaucoup de la méthode utilisée. Relativement conséquent avec la méthode M1, il se montre moins marquant avec

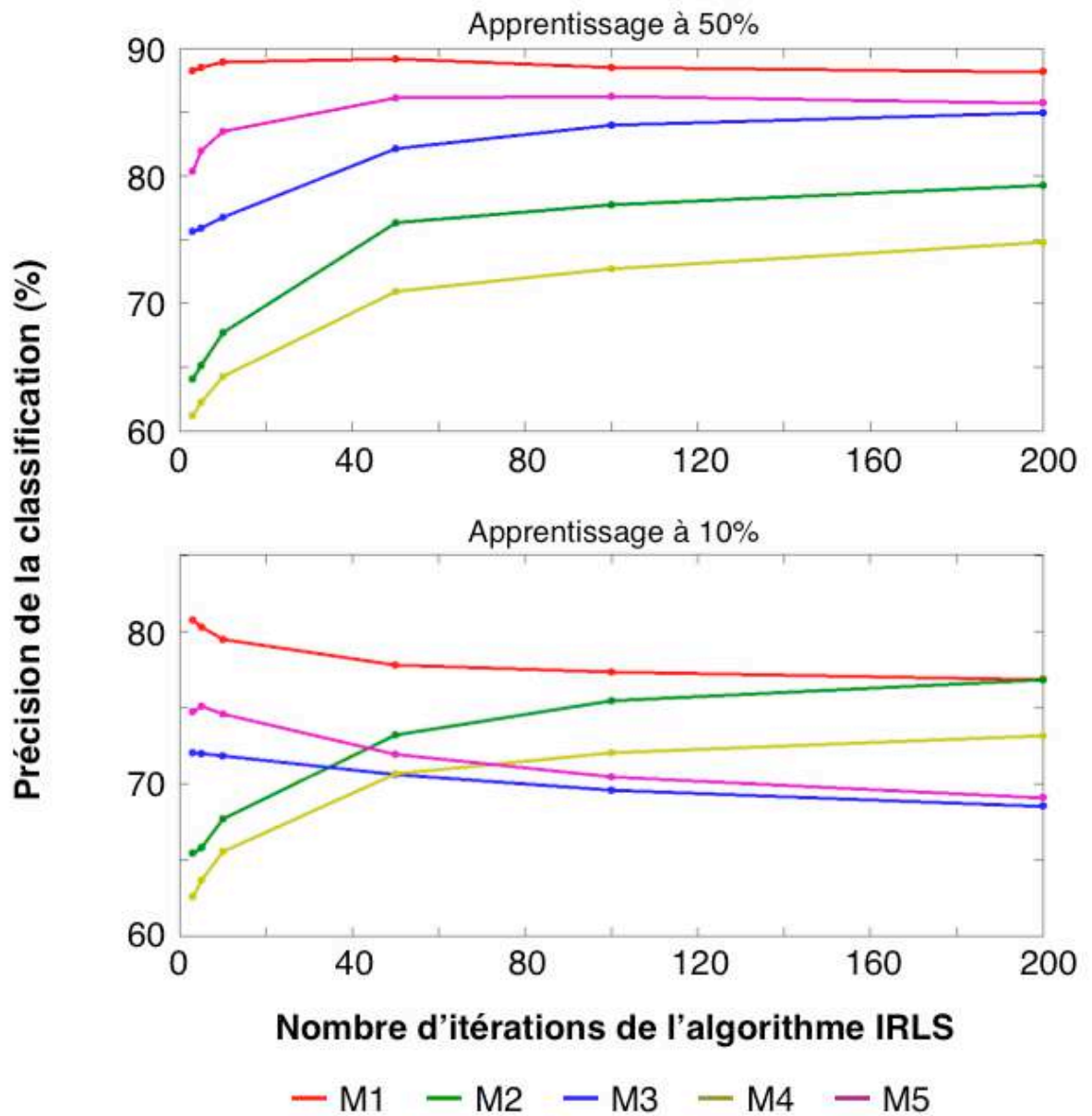


FIG. 3.4: Pourcentage de bonne classification des résultats obtenus sur l'image IP avec les diverses méthodes en fonction du nombre d'itération de l'algorithme IRLS. Dans chaque cas, le λ conduisant au résultat optimal a été choisi.

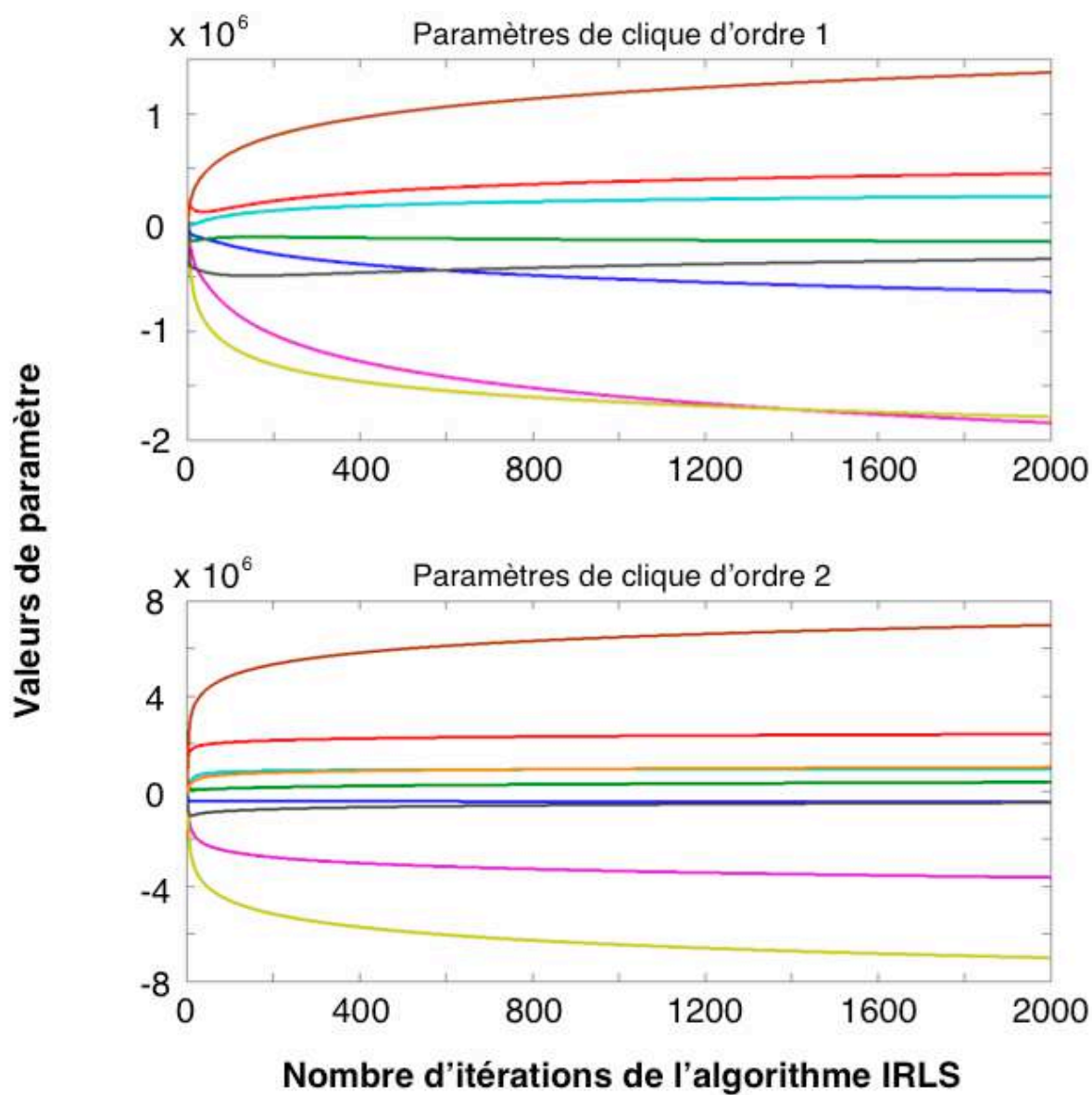


FIG. 3.5: Evolution de la valeur de certains paramètres de l'algorithme IRLS appliqué sur l'image IP en fonction du nombre d'itérations réalisées. Chaque courbe de paramètre est associée à une classe différente

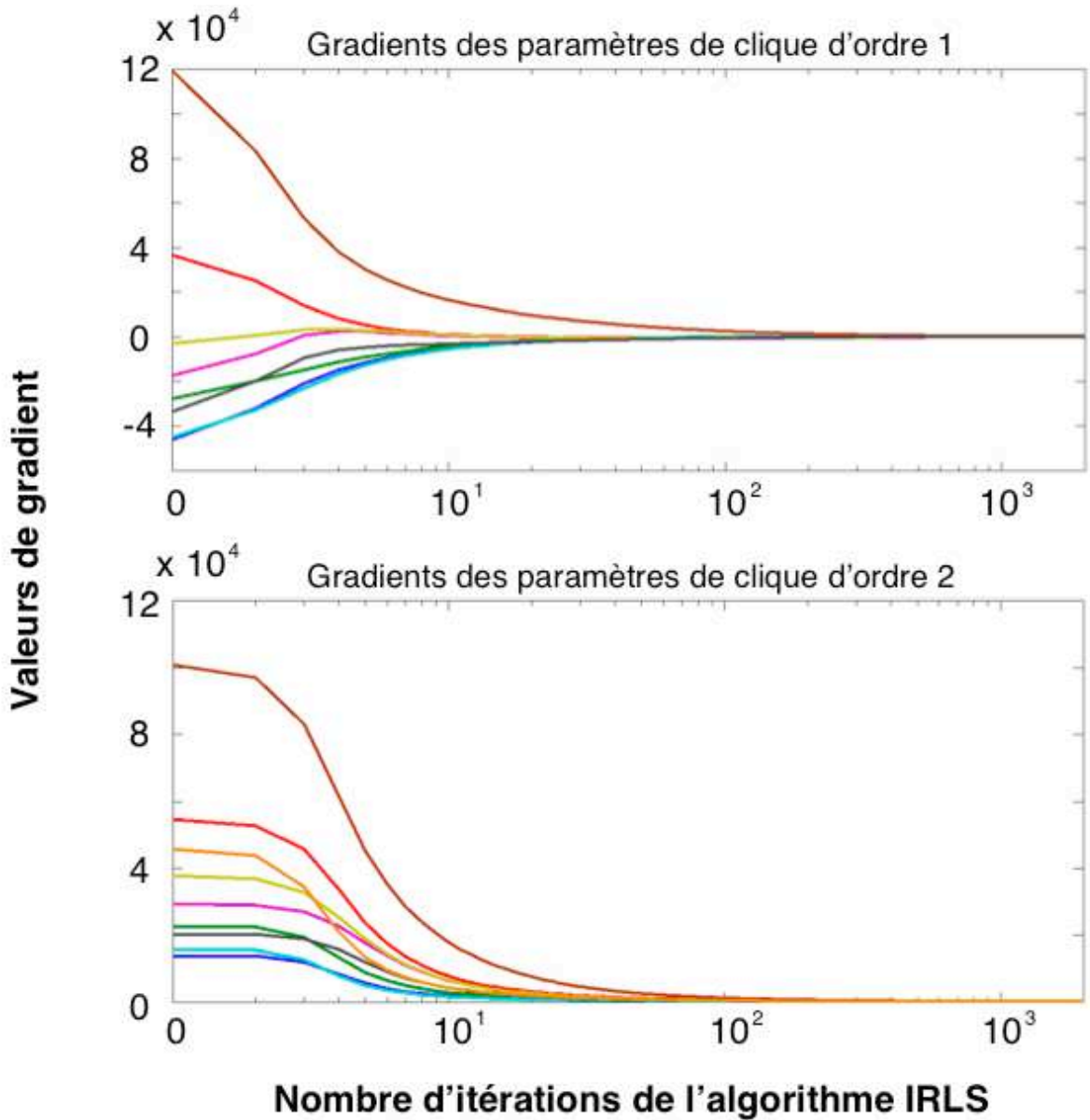


FIG. 3.6: Evolution du gradient de certains paramètres de l'algorithme IRLS appliqué sur l'image IP en fonction du nombre d'itérations réalisées. Chaque courbe de paramètre est associée à une classe différente

les méthodes M2 et M3 pour devenir presque négligeable avec la méthode M4.

4.2 Extension à l'analyse d'autres jeux de données

En ce qui concerne les jeux de données MR, MR2 et CO, l'apprentissage a été réalisé avec des proportions d'échantillons utilisées pour l'apprentissage (PEA) moindres, à savoir 10% et 1%. La figure 3.7 montre que lorsque l'apprentissage est réalisé sur 10% des échantillons de la vérité terrain, l'ensemble des méthodes étudiées est à même de bâtir un modèle robuste tel que la précision de la classification augmente avec le nombre d'itérations de l'algorithme d'optimisation IRLS, ce qui n'était pas le cas de l'image IP. Avec seulement 1% de ces échantillons, seule l'image MR2 présente les mêmes symptômes de manque de robustesse que cette même image. Ces observations tendent à signifier que les classes des images MR2, MR et CO (ces deux dernières tout particulièrement) sont mieux définies que celles de l'image IP puisqu'une proportion moindre d'échantillons d'apprentissage permet de construire un modèle représentatif capable de classer correctement le reste des pixels de l'image.

Cette figure montre également que la réduction de dimension (méthode M2 et M4) et la modification de représentation des cliques d'ordre 2 (méthodes M3 et M4) entraînent sur ces images une baisse de précision beaucoup moins préjudiciable (de l'ordre de 1%, la plupart du temps) que sur l'image IP. Si cela vient conforter l'hypothèse énoncée plus haut, selon laquelle l'efficacité de l'analyse en composantes principales pouvait dépendre du jeu de données utilisé, on ne peut toutefois pas dire que les angles spectraux soient plus efficaces sur ces images. Les tableaux 3.2, 3.3 et 3.4 montrent en effet que l'apport de l'information spatiale est systématiquement plus réduit sur les images MR, MR2 et CO que sur l'image IP. Dans le cas des méthodes M3 et M4, cet apport est même quasiment inexistant. Nous pouvons également remarquer sur ces tableaux que si les résultats obtenus avec les classifications CRF (M1) et SVM sont comparables lorsque l'on utilise 10% des échantillons de la vérité terrain pour l'apprentissage, la classification SVM est plus performante que la classification CRF sur les images MR et MR2 lorsque 1% seulement des échantillons sont utilisés, l'image CO constituant la seule exception. De manière générale, la classification SVM apparaît ainsi plus robuste que la classification CRF puisque la précision de ses résultats est moins sensible au jeu d'échantillons utilisé lors de l'apprentissage, ce qui dans un contexte supervisé constitue un atout majeur.

L'image 3.8 témoigne de la légère régularisation spatiale apportée par la classification LBP vis-à-vis des classifications MAP et SVM purement spectrales. Toutefois nous pouvons remarquer que l'algorithme LBP engendre parfois des erreurs de classification sur de larges zones, ce que l'on constate notamment sur

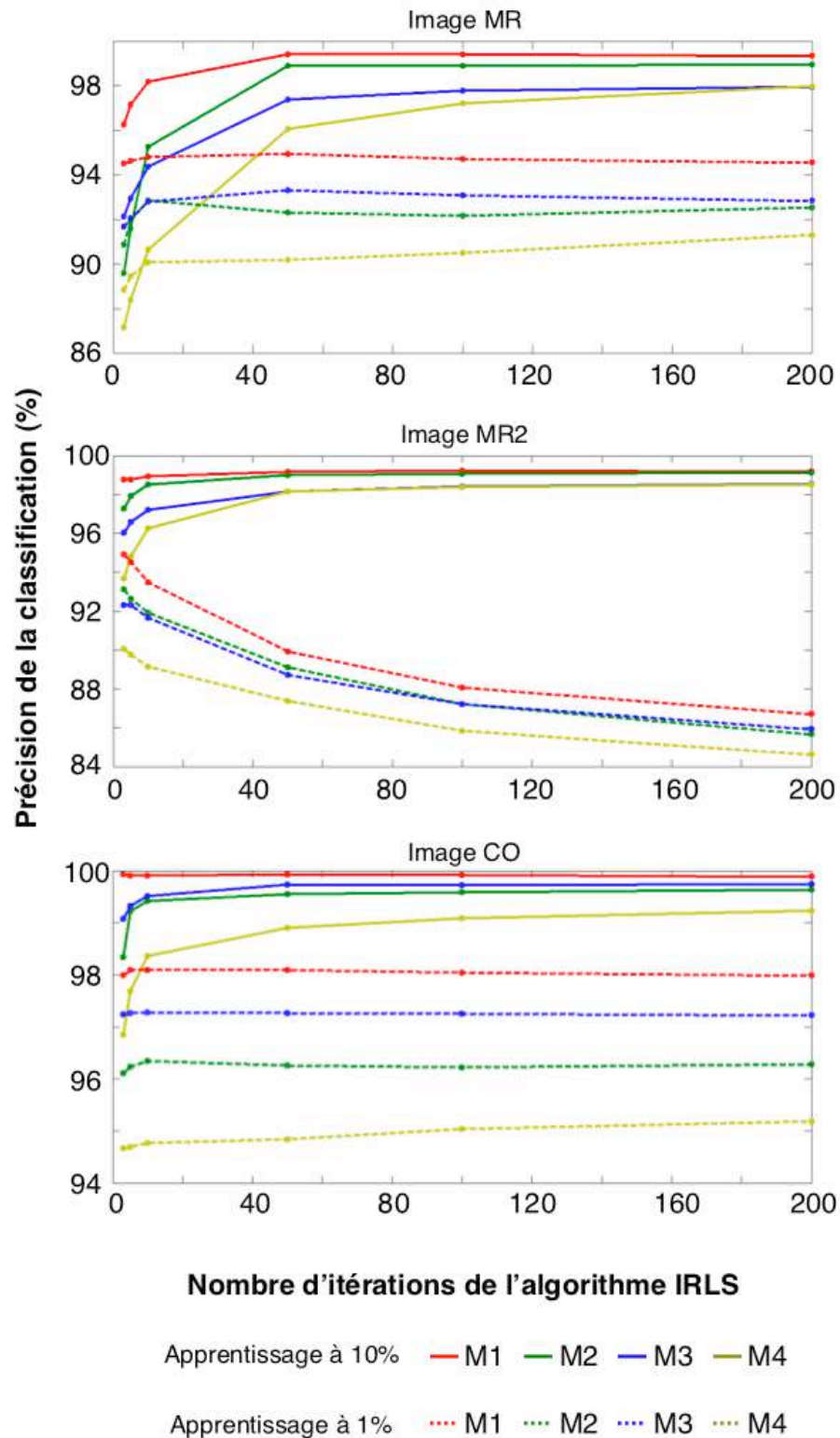


FIG. 3.7: Pourcentage de bonne classification des résultats obtenus avec les diverses méthodes en fonction du nombre d'itération de l'algorithme IRLS.

les images MR et MR2, là où les erreurs d'un classifieur SVM sont le plus souvent de l'ordre du pixel ou du petit groupe de pixels. Ce cas de figure arrive lorsque, pour une zone donnée, le modèle MLR est partagé entre plusieurs classes (cf. les cartes de classification MAP). Le LBP, censé régler le problème en utilisant l'information spatiale, n'opte pas toujours pour la classe adéquate, ce qui veut dire que les cliques d'ordre 2 associées au voisinage immédiat de chaque pixel ne sont pas assez représentatives du comportement textural de la classe considérée. Cela explique notamment la précision de classification généralement plus faible du classifieur LBP.

PEA	Méthode	Valeur de λ							
		0,01	0,1	0,3	0,5	0,7	0,9	0,99	MAP
10%	M1	98,6	99	99,3	99,1	98,7	98	97,7	97,6
	M2	97,9	98,4	98,9	98,9	98,6	98,1	97,8	97,7
	M3	19	92,9	95,4	97,8	98	97,7	97,6	97,6
	M4	41	83,8	98	97,9	97,8	97,7	97,7	97,7
	SVM	99,1							
1%	M1	94,5	94,6	94,6	94,3	93,3	92,1	91,7	91,6
	M2	91,5	91,8	92,4	92,6	92,1	91,3	90,8	90,7
	M3	18,9	61,5	92,5	92,9	92	91,7	91,6	91,6
	M4	49,9	82,6	91,3	91	90,9	90,8	90,7	90,7
	SVM	97,6							

TAB. 3.2: Pourcentage de bonne classification des résultats obtenus sur l'image MR avec les diverses méthodes en fonction du paramètre λ . La colonne MAP correspond au cas où λ est égal à 0. Dans ce cas, il est inutile d'appliquer une classification LBP et une simple classification par Maximum A Posteriori est utilisée.

CHAPITRE 3. CONDITIONAL RANDOM FIELD (CRF) ADAPTÉ AUX DONNÉES
HYPERSPÉCTRALES

PEA	Méthode	Valeur de λ							
		0,01	0,1	0,3	0,5	0,7	0,9	0,99	MAP
10%	M1	98,8	99,1	99,2	99,2	98,9	98,5	98,3	98,3
	M2	98,6	98,9	99,1	99,1	98,9	98,5	98,2	98,2
	M3	39,4	88,5	94,1	97,3	98,6	98,4	98,3	98,2
	M4	49	83,9	96,9	98,5	98,4	98,2	98,2	98,2
	SVM	98,9							
1%	M1	86,3	86,6	86,7	86,7	86,9	86,1	86	86
	M2	84,7	85,3	85,6	85,7	85,7	85,6	85,5	85,5
	M3	17,3	84,2	83,9	85,9	85,9	86	86	86
	M4	37,1	79	84,6	85,3	85,4	85,4	85,5	85,4
	SVM	97,2							

TAB. 3.3: Pourcentage de bonne classification des résultats obtenus sur l'image MR2 avec les diverses méthodes en fonction du paramètre λ .

PEA	Méthode	Valeur de λ							
		0,01	0,1	0,3	0,5	0,7	0,9	0,99	MAP
10%	M1	99,9	99,9	99,9	99,9	99,8	99,8	99,7	99,7
	M2	99,7	99,7	99,6	99,5	99,4	99,3	99,2	99,2
	M3	58	98,7	99,3	99,7	99,7	99,7	99,7	99,7
	M4	60,2	98	99	99,2	99,3	99,2	99,2	99,2
	SVM	99,5							
1%	M1	98,1	98	97,9	97,8	97,5	97,3	97,2	97,2
	M2	96,1	96,3	96,1	95,8	95,5	95,3	95,2	95,2
	M3	49,4	94,2	97,1	97,2	97,2	97,2	97,2	97,2
	M4	53,6	95,7	95,4	95,2	95,2	95,2	95,2	95,2
	SVM	96,5							

TAB. 3.4: Pourcentage de bonne classification des résultats obtenus sur l'image CO avec les diverses méthodes en fonction du paramètre λ .

	IP	MR	MR2	CO
VT				
MAP				
LBP				
SVM				

FIG. 3.8: Comparaison des résultats de classification obtenus avec les classifieurs MAP, LBP et SVM. Pour l'image IP, l'apprentissage a été effectué sur 10% des éléments de la vérité terrain, 1% pour les trois autres images.

5 Conclusion

Nous avons présenté dans ce troisième chapitre une approche de classification spectrale/spatiale basée sur un modèle CRF (*Conditional Random Field*) imaginé par Zhong et al. [58]. Nous avons proposé plusieurs variantes de l'algorithme afin notamment de diminuer le nombre de paramètres requis par son modèle. Ces variantes consistent d'une part en l'intégration d'une réduction de dimension par ACP des jeux de données à classifier et une modification de la représentation des cliques d'ordre 2 dans le modèle d'autre part.

L'étude des performances de ce processus de classification s'est tout d'abord limitée au jeu de données utilisé dans l'article [58] (l'image IP). Nous avons notamment remarqué que la robustesse de l'algorithme est fortement dépendante de la proportion d'échantillons de vérité terrain utilisée lors de l'apprentissage. Utiliser une proportion trop faible revient en effet à spécialiser le modèle sur un ensemble d'échantillons de la vérité terrain insuffisamment représentatif des classes associées au jeu de données considéré. Ce problème tend à disparaître lorsque nous mettons en place les variantes mentionnées plus haut, mais cela coïncide avec une diminution notable de la précision de classification signifiant d'une part que l'image IP réduite par ACP ne contient pas autant d'information que le jeu de données original, et d'autre part que les variantes de représentation des cliques d'ordre 2 reflètent moins le comportement textural des classes que le mode de représentation original. Sur les trois autres jeux de données, la réduction de dimension est plus efficace et diminue plus légèrement la précision des résultats de classification. Cependant, la plus-value apportée par l'information spatiale dans ces trois cas est beaucoup plus faible qu'avec l'image IP, ce qui limite l'intérêt de l'approche puisqu'une classification SVM, robuste même lorsque le nombre d'échantillons utilisés lors de l'apprentissage est limité, aboutit à des résultats généralement meilleurs. Enfin, malgré des résultats statistiquement moins bons dus à une moins bonne discrimination des classes que la classification SVM en général, la classification CRF permet tout de même une meilleure régularisation spatiale au sein de larges zones comportant des pixels appartenant à une même classe.

Chapitre 4

Classification par zones

1 Principe

Après la classification vectorielle d'attributs spectraux/texturaux et la modélisation markovienne des relations de voisinage, nous nous focalisons dans ce chapitre sur une troisième manière de tirer parti de l'information spatiale : la segmentation. Nous introduisons ainsi un nouveau processus de classification (AC, pour *Area Classification*) dont le matériel de base n'est plus le pixel mais la zone, c'est-à-dire un ensemble de pixels connexes de forme et de taille variable présentant une certaine homogénéité spectrale et/ou texturale.

Dans les grandes lignes, le processus se divise en trois étapes (voir figure 4.1). Tout d'abord, l'image est segmentée de sorte à identifier des zones connexes dont les pixels présentent des caractéristiques spectrales et contextuelles similaires. Nous utilisons pour cela l'algorithme JSEG développé par Deng et al. [12]. Cela débouche sur une carte de segmentation où chaque zone est associée à un label unique. Parallèlement, divers attributs spectraux et texturaux sont extraits de l'image hyperspectrale pour former une image d'attributs, de la même façon que dans le chapitre 2. Carte de segmentation et image d'attributs sont alors utilisées conjointement pour classifier les zones.

Là encore, nous étudierons l'intérêt de cette méthode relativement aux différents objectifs de classification que nous nous sommes fixés.

2 Segmentation JSEG

Ce processus de segmentation, introduit par Deng et al. [12], a la particularité de tenir compte de l'information contextuelle afin d'éviter les problèmes de sursegmentation dus à la présence de texture dans l'image. Il se compose de plusieurs étapes successives : la quantification, le création de l'image d'hétérogénéité et enfin la croissance de régions.

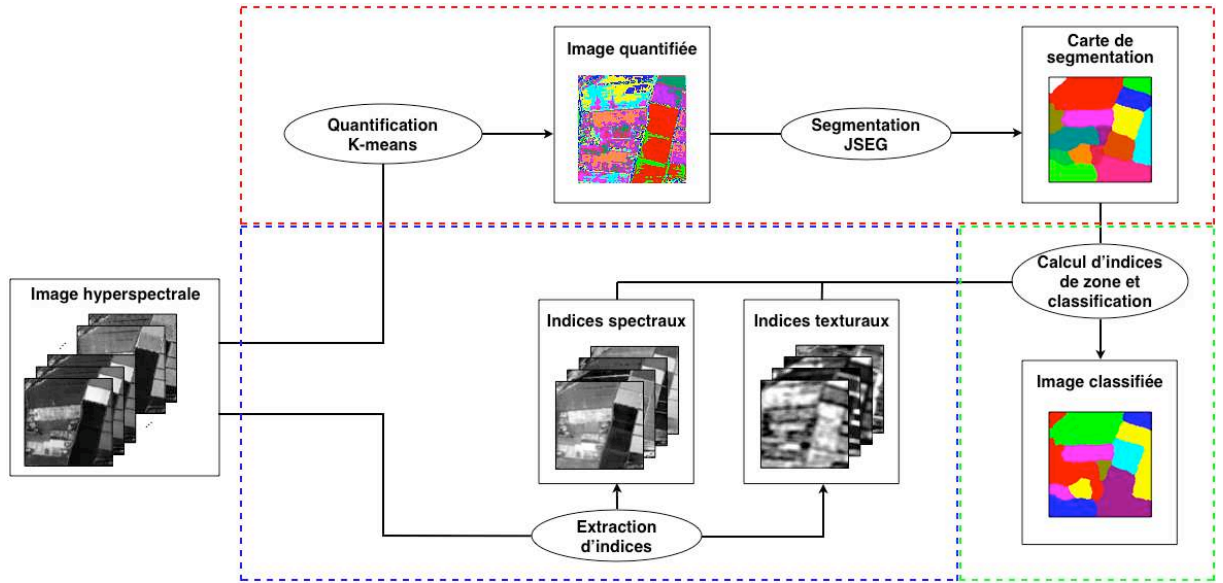


FIG. 4.1: Aperçu général du processus de classification

2.1 Quantification

L'algorithme de segmentation s'appliquant sur des données scalaires, il est nécessaire d'effectuer au préalable une quantification sur l'image vectorielle de départ. Nous proposons d'utiliser pour cela une classification K-means, chaque classe représentant un intervalle de valeurs vectorielles de l'image originale. Afin d'obtenir l'image quantifiée la plus informative possible, les centroïdes de départ sont sélectionnés successivement par l'intermédiaire d'une série de tirages aléatoires, de sorte qu'à chaque étape, le nouveau centroïde soit aussi différent que possible au sens de l'angle spectral des centroïdes déjà retenus. Plusieurs classifications par K-means sont ensuite appliquées à partir de différentes initialisations et le résultat conservé est celui maximisant la variance inter-classe tout en minimisant la moyenne des variances intra-classe.

2.2 Calcul du critère d'hétérogénéité

A partir de cette carte de classification est définie une J-image correspondant à une carte de mesure de l'hétérogénéité locale. Soient \mathbf{X} l'image hyperspectrale originale et \mathbf{X}' la carte de classification K-means correspondante. \mathbf{X} et \mathbf{X}' sont toutes deux associées à une grille de pixel de L lignes et S colonnes ou (s, l) tel que $s \in [0, \dots, S-1]$ et $l \in [0, L-1]$ représente le couple de coordonnées spatiales d'un pixel. Pour chaque pixel (s, l) (hors pixel proche des bords), nous considérons un voisinage V correspondant à une fenêtre glissante de taille $(2R+1) \times (2R+1)$ centré sur ce pixel. Dans le référentiel de ce voisinage, $z(v)$ correspond au couple de coordonnées locales d'un voisin $v \in V$ et m au couple de coordonnées moyennes

de la fenêtre (à savoir celles de son centre).

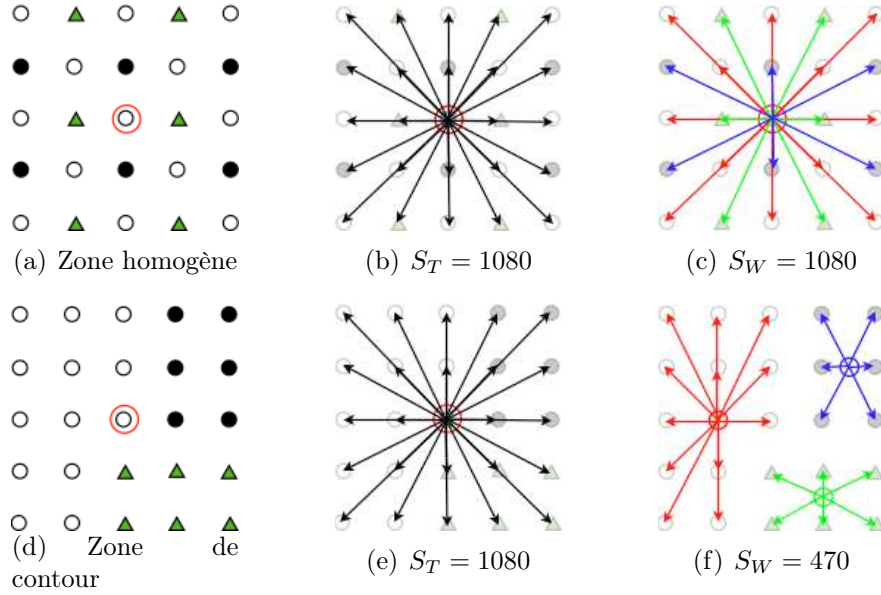


FIG. 4.2: Calcul du critère d'hétérogénéité h sur une zone homogène (a) et une zone de contour (d). Chaque flèche représente une distance euclidienne entre le couple de coordonnées d'un centroïde (entouré d'un cercle de couleur) et celui d'un pixel du voisinage. (b) et (e) représentent S_T pour l'un et l'autre voisinage. Sur la figure (c), chaque classe est répartie de façon parfaitement homogène sur l'ensemble de la fenêtre. Leur position moyenne est donc confondue avec le centre de cette dernière et $h = 0$. Sur la figure (f), les pixels de même classe se concentrent aux extrémités du voisinage. Leur position moyenne m_k est par conséquent excentrée et $S_T > S_W$. Dans ce cas ci, $h = 1.3$.

Si l'on part du principe que V est composé de pixels appartenant à K classes différentes, $V_k, k = 1, \dots, K$ contient l'ensemble des couples de coordonnées des pixels appartenant à la classe k , N_k est le cardinal de cet ensemble et m_k la moyenne de ses éléments. Soient :

$$S_T = \sum_{v \in V} \|z(v) - m\|^2 \quad (4.1)$$

et :

$$S_W = \sum_{i=k}^K \sum_{v \in V_k} \|z(v) - m_k\|^2 \quad (4.2)$$

S_T représente la somme des distances entre l'ensemble des pixels du voisinage et le pixel central. Pour une taille de fenêtre donnée, cette valeur est constante. S_W est pour sa part la somme des distances entre l'ensemble des pixels du voisinage

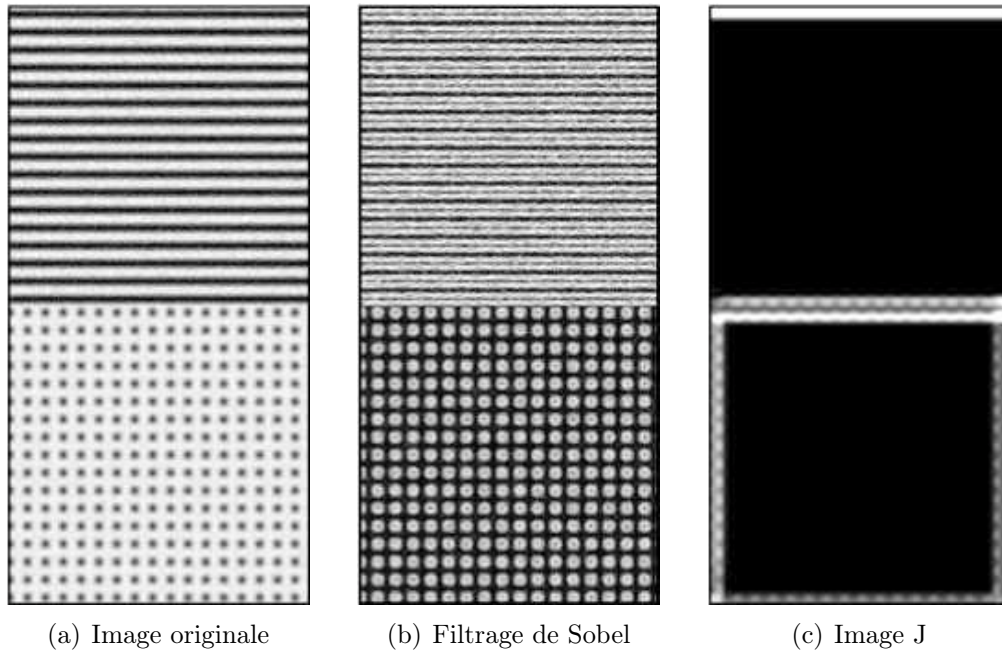


FIG. 4.3: Comparaison du filtrage de Sobel avec le calcul d'une image d'hétérogénéité

et le barycentre m_k de leur classe d'appartenance. Le critère d'hétérogénéité h d'un voisinage est alors défini par :

$$h = (S_T - S_W)/S_W \quad (4.3)$$

La valeur h , forcément positive puisque $S_T > S_W$, sera élevée si le voisinage couvre une région de contour et proche de 0 s'il couvre une région homogène, texturée ou non (voir figure 4.2). Lorsque la fenêtre glissante est assez grande, ce critère d'hétérogénéité est capable d'ignorer les micro-variations relatives aux textures pour ne souligner que les contours forts. h est calculé pour l'ensemble des pixels de l'image \mathbf{X}' de sorte à former une nouvelle image \mathbf{J} .

La figure 4.3 présente un exemple d'image d'hétérogénéité obtenue à partir d'une image synthétique composée de deux zones se démarquant uniquement par leur texture, et la compare au résultat d'un algorithme de détection de contours plus classique : le filtrage de Sobel. Là où ce dernier souligne essentiellement les contours liés aux structures internes des deux textures, l'image d'hétérogénéité ne met en valeur que l'unique contour séparant les deux zones texturées.

2.3 Algorithme de croissance de régions

La croissance de régions est un procédé de segmentation d'image consistant à identifier plusieurs ensembles de pixels connexes appelés graines et représentant chacune le coeur d'une zone, puis à les faire grossir jusqu'à ce que la totalité

des pixels de l'image soient associés à une zone. Afin de construire ces graines, nous utilisons l'image d'hétérogénéité \mathbf{J} calculée dans la section précédente, le but étant de construire des zones de texture homogène.

2.3.1 Identification des graines

Sachant que la carte de segmentation sera composée d'autant de zones qu'il y a de graines au début de l'algorithme, ces ensembles de départ doivent être définis soigneusement. La détermination automatique du jeu de graines aboutissant à la carte de segmentation optimale n'est pas un problème trivial. Deng et al. [12] définissent une valeur de seuil égale à la moyenne globale de l'image J et en dessous duquel chaque pixel sera considéré comme le membre potentiel d'une graine. La rigidité d'un tel procédé peut poser problème lorsque différentes régions, homogènes d'un point de vue textural, ne présentent pas le même niveau moyen d'hétérogénéité. Nous utilisons donc une valeur de seuil plus souple notée τ_V et proposée par Jing et al. [23]. L'ensemble des pixels candidats au statut de graine vérifie $\mathbf{J}(s, l) < \tau_V$ où τ_V est un seuil qui dépend à la fois des statistiques globales de l'image et des statistiques locales du voisinage V (défini dans la section 2.2) associées au pixel (s, l) :

$$\tau_V = \min(\max(F_V, F_{inf}), F_{sup}) \quad (4.4)$$

avec :

$$F_V = \mu_V - \alpha_1 \sigma_V \quad (4.5)$$

$$F_{inf} = \mu - \alpha_2 \sigma \quad (4.6)$$

$$F_{sup} = \mu + \alpha_3 \sigma \quad (4.7)$$

μ_V et σ_V sont respectivement la moyenne et l'écart-type des valeurs de l'image \mathbf{J} pour le voisinage V , tandis que μ et σ sont la moyenne et l'écart-type de l'ensemble de l'image \mathbf{J} . L'utilisation des statistiques locales permet d'adapter le seuil au niveau moyen du voisinage, tout en définissant des bornes supérieures et inférieures à la moyenne globale de l'image (F_{sup} et F_{inf}) à ne pas dépasser, comme nous le voyons sur la figure 4.4. F_V représente quant à lui la moyenne locale. Trois cas peuvent se présenter :

- $F_V < F_{inf} \Rightarrow \tau_V = F_{inf}$
- $F_{inf} < F_V < F_{sup} \Rightarrow \tau_V = F_V$
- $F_V > F_{sup} \Rightarrow \tau_V = F_{sup}$

Autrement dit le seuil sera égal à F_V tant que celle-ci sera comprise entre les bornes F_{inf} et F_{sup} . α_2 et α_3 fixent la distance des bornes inférieure et supérieure à la moyenne globale de l'image. α_1 permet pour sa part de régler la valeur de τ_V lorsque la moyenne locale se situe entre les deux bornes. Le choix des valeurs

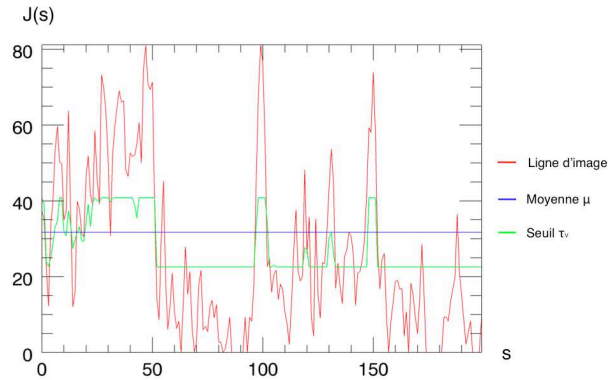


FIG. 4.4: Exemple de calcul du seuil τ_V (en vert) associé à une ligne d'image (en rouge). La ligne bleue représente quant à elle la moyenne μ de l'image.

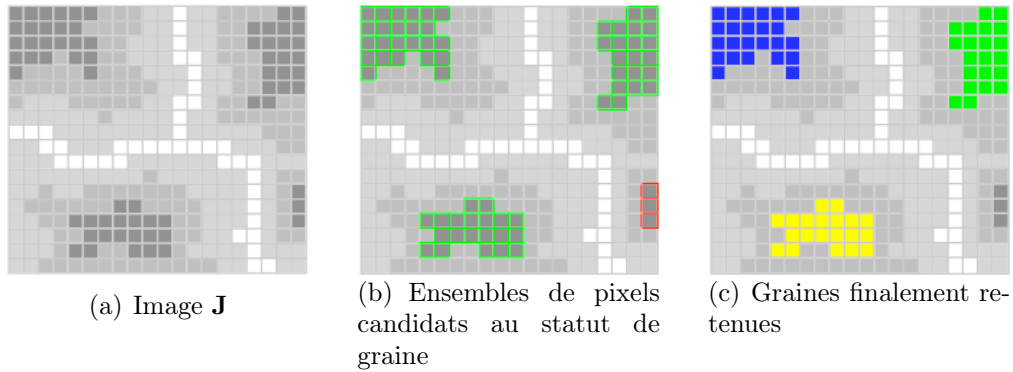


FIG. 4.5: Processus d'identification des graines d'une image \mathbf{J} . Sur l'image (b), les pixels verts représentent les futures graines et les pixels rouges les ensembles de pixels connexes trop petits pour former une graine

α_1 , α_2 et α_3 sera abordé ultérieurement dans la section présentant les résultats de ce chapitre. Une fois les pixels candidats identifiés, ils sont rassemblés en sous-ensembles connexes et conservés en tant que graines si leur cardinal est supérieur à $2R^2$, R étant la demi-largeur du voisinage considéré (voir figure 4.5).

2.3.2 Croissance de régions

Une fois les graines identifiées, une valeur d'hétérogénéité moyenne \bar{h} est calculée sur l'ensemble des pixels n'appartenant pas encore à une graine. Parmi ces derniers, ceux qui sont inférieurs à \bar{h} sont rassemblés pour former de nouveaux sous-ensembles connexes que l'on fusionne le cas échéant à la graine dont ils sont le voisin (voir figure 4.6). Ces deux étapes sont itérées jusqu'à stabilité. Les pixels n'appartenant toujours pas à une graine sont finalement traités un par un par ordre croissant de valeur d'hétérogénéité h jusqu'à ce que l'ensemble des pixels de l'image soit associé à une graine. Lorsqu'il y a ambiguïté, c'est-à-dire lors-

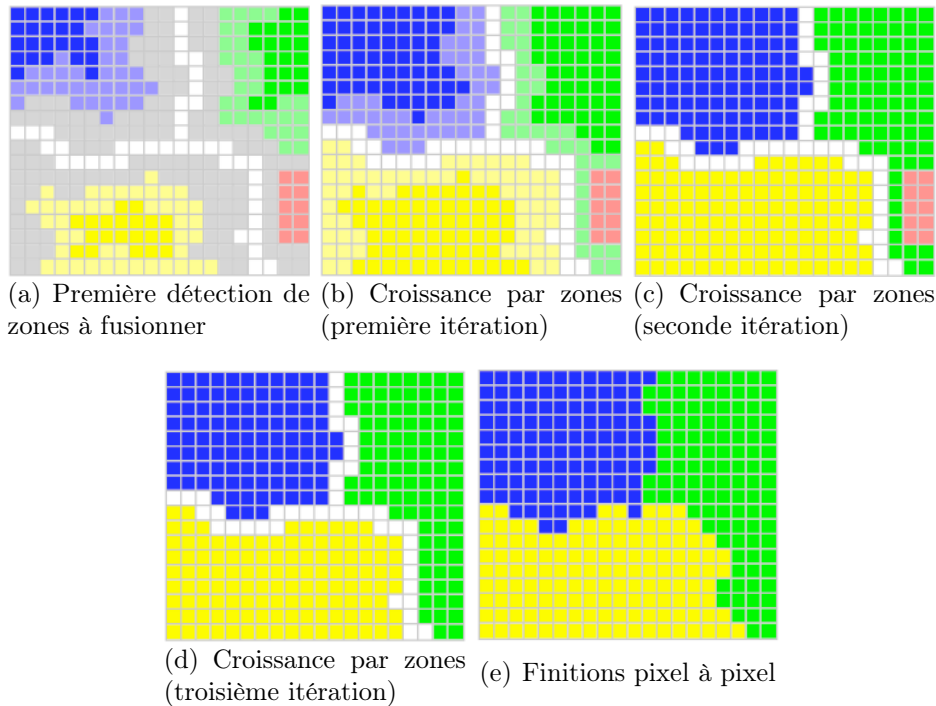


FIG. 4.6: Processus itératif de croissance de région à partir des graines. La zone rouge correspond à un ensemble de pixels connexes en dessous de la moyenne des valeurs d'hétérogénéité mais dont aucun pixel n'est voisin d'une graine. Ces cas se présentent généralement là où un ensemble de pixels trop petit n'a pas été retenue comme graine pour l'étape de croissance.

qu'un pixel est voisin de plusieurs graines, ce pixel est fusionné à la graine dont le spectre moyen est le plus proche du sien au sens de l'angle spectral.

3 Calcul d'attributs spectraux/texturaux de zone et classification

En complément de la segmentation, nous proposons une étape de classification spectrale/spatiale s'appuyant sur les zones mises en évidence par la carte de segmentation. Afin de transformer cette dernière en carte de classification, nous nous appuyons sur les mêmes attributs spectraux et texturaux que dans le chapitre 2. Rappelons que \mathbf{X} est l'image hyperspectrale originale, de taille $S \times L \times M$, où M correspond au nombre de bandes spectrales. Soit \mathbf{W} l'image d'attributs associée, de dimensions $S \times L \times C$ tel que C est le nombre d'attributs spectraux et/ou texturaux retenus. Considérons une première approche dans laquelle nous commençons par réaliser à partir de l'image \mathbf{X} et de la carte de segmentation un nouveau jeu de données \mathbf{Z} de taille $A \times C$ - où A est le nombre total de zones de la seg-

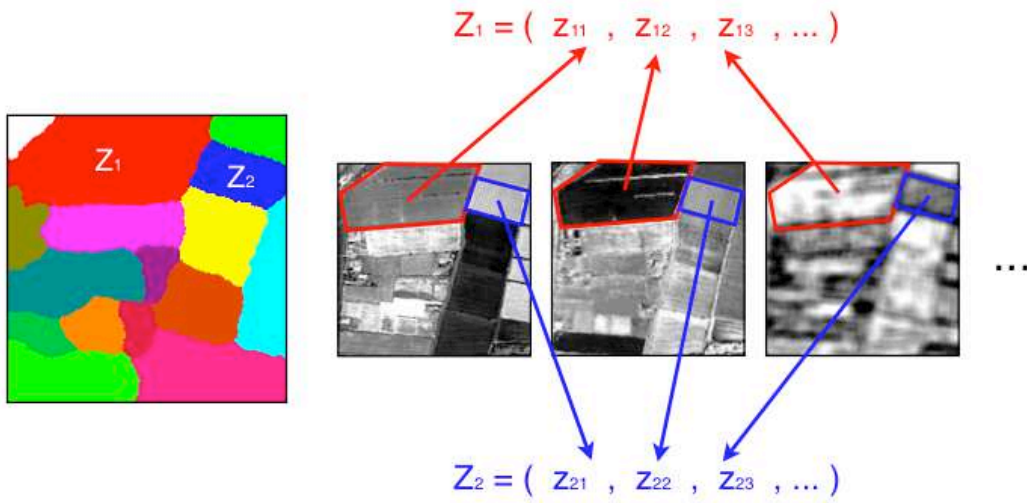


FIG. 4.7: Création des attributs spectraux/texturaux de zone

mentation - dont chaque élément $\mathbf{Z}_a = (Z_{a1}, Z_{a2}, \dots, Z_{aC}) | a \in \{1, \dots, A\}$ contient les attributs spectraux et/ou texturaux associés à une zone (voir figure 4.7). Le procédé le plus immédiat consiste à associer à chaque zone le vecteur d'attributs moyen calculé sur l'ensemble de ses pixels.

En ce qui concerne les coefficients de Haralick et les spectres de texture, nous avons également développé une variante consistant à calculer directement les attributs de zone sans passer par la phase d'extraction pixel à pixel. Matrices de co-occurrence et spectres de textures sont calculés sur l'ensemble des pixels de chaque zone et non plus sur des voisinages limités. Les attributs texturaux extraits de ces deux types de structure de données sont ainsi directement liés aux zones correspondantes. Dans le cas des spectres de texture, ce procédé permet notamment d'obtenir des histogrammes plus étoffés et donc statistiquement plus robustes et représentatifs d'une certaine texture.

Le nombre d'éléments de \mathbf{Z} étant trop réduit pour obtenir des statistiques robustes, une classification non supervisée par K-means ou modèle de mélange gaussien n'est pas appropriée. Nous utilisons donc une classification hiérarchique ascendante (CHA), un procédé itératif de fusion de classes (sachant que dans notre cas, les classes de départ correspondent aux zones de la carte de segmentation) basé sur un critère de similarité entre classes. A chaque itération, le critère de similarité est calculé pour l'ensemble des couples de zones, le couple de zone le plus semblable étant fusionné ; ces deux étapes sont répétées jusqu'à ce que le nombre de classes fixé par l'utilisateur soit atteint. Dans cette première approche, le critère de similarité utilisé est la distance de Ward :

$$D_w(\mathbf{Z}_i, \mathbf{Z}_j) = \frac{n_i n_j}{n_i + n_j} \sum_{c=1}^C (Z_{ic} - Z_{jc})^2 \quad (4.8)$$

Ce critère, correspondant à une distance euclidienne pondérée par le cardinal des zones comparées, est intéressant car très rapide à calculer. Il ne permet cependant pas de mettre en valeurs toute la complexité de l'information statistique relative à une zone.

C'est pourquoi nous considérons également un second critère de similarité : la distance de Bhattacharyya. Ce critère ne fait plus intervenir les centroïdes des zones à classifier mais leur histogramme. Le jeu de données \mathbf{Z} est ainsi remplacé par un jeu de données \mathbf{Q} de taille $A \times B^C$ où B (*bins*) représente le niveau d'échantillonnage de chacune des C composantes. Chaque zone étant associée à un histogramme multidimensionnel $\mathbf{Q}_a(e)$, la distance de Bhattacharyya entre deux zones est donnée par :

$$D_{bc}(\mathbf{Q}_i, \mathbf{Q}_j) = -\log \sum_{e \in E} \sqrt{\mathbf{Q}_i(e)\mathbf{Q}_j(e)} \quad (4.9)$$

telle que E représente une partition de l'espace divisée en B^C intervalles de valeurs (ou "classes"). Ce critère de similarité est également utilisé dans le cadre d'une classification hiérarchique ascendante, mais il est limité par la taille des histogrammes multidimensionnels lorsque la dimension de l'image à classifier devient trop grande. Prenons l'exemple de l'image MR, dont les zones comportent toutes 2500 pixels. Un ensemble de $C = 4$ attributs couplé à un niveau d'échantillonnage $B = 10$ engendrerait des histogrammes de taille 10^4 , soit quatre fois plus que le nombre de pixels de chaque zone. De tels histogrammes seraient inévitablement trop vides pour représenter correctement les zones à classifier. Le nombre d'attributs C étant le plus pénalisant en ce qui concerne la taille des histogrammes à comparer, ce type de critère ne peut être utilisé que sur des images comportant un faible nombre d'attributs.

4 Résultats

Nous allons maintenant présenter les résultats du processus de classification par zones. Ce dernier a été appliqué sur les jeux de données MR, MR2 et CO. L'image IP n'est pas utilisée pour les raisons exposées dans les chapitres précédents.

De même que la classification vectorielle présentée dans le chapitre 2, le processus de classification par zones est divisé en plusieurs étapes, chacune faisant intervenir plusieurs paramètres. Afin de déterminer la valeur optimale de ces derniers, nous avons ainsi procédé à une étude de sensibilité similaire, chaque paramètre étant testé sur un intervalle de valeurs significatif. La quantification, première étape du processus, dépend de trois paramètres : le nombre de pré-classifications par K-means effectuées pour la quantification (NKM), le nombre de tirages aléatoires de jeux de centroïdes de départ pour chaque pré-classification (NTA) et

enfin le nombre de "sous-classes" qu'elles doivent rechercher (NC). NKM et NTA sont des paramètres d'optimisation que l'on fixe à $NKM = 10$ et $NTA = 500$. NC est le paramètre crucial de cette étape car c'est de lui que dépend la capacité de l'image quantifiée à simplifier l'image originale sans supprimer totalement l'information texturale qui y réside. Nous avons testé pour NC les valeurs $2K$, $3K$ et $4K$, où K est le nombre de classes réelles de l'image étudiée. $2K$ paraît être un minimum car nous partons du principe que l'image est constituée de classes réelles texturées et comportant par conséquent chacune un minimum de 2 sous-classes, tandis qu'au-delà de $4K$ la texture risquerait d'être noyée sous le nombre de sous-classes.

La seconde étape est la segmentation, composée du calcul de l'image d'hétérogénéité sur laquelle est ensuite appliqué l'algorithme de croissance de régions. Elle dépend pour sa part de deux types de paramètres : la taille du voisinage (V) à partir duquel est calculé le critère d'hétérogénéité pour chaque pixel et les valeurs α_1 , α_2 et α_3 . Nous avons considéré des tailles de voisinages similaires à celles employées dans le chapitre 2 pour le calcul des attributs texturaux, c'est-à-dire des fenêtres carrées allant de 11×11 à 31×31 pixels de large. En ce qui concerne les valeurs de seuil associées à l'identification des pixels graines, nous avons fixé α_1 à 0.5, comme dans l'article [23]. Puis, en imposant $\alpha_2 = \alpha_3$, nous avons fait varier α_2 et α_3 afin de donner plus ou moins de poids aux statistiques locales. Trois valeurs ont été testées : 0.4, 0.7 et 1.0.

Lors de la dernière étape, la classification hiérarchique ascendante, les paramètres dont il faut tenir compte sont le nombre et le type d'attributs de zone mis à contribution ainsi que le type de distance utilisé pour comparer les zones. D'un point de vue spectral, nous avons calculé les attributs de zones sur les bandes ACP de chaque image étudiée, cette méthode de réduction de dimension ayant manifesté le meilleur rapport performance/robustesse dans le chapitre 2. Pour déterminer le nombre de composantes à retenir pour la classification, nous avons testé les quatre critères de sélection (CS) présenté dans ce même chapitre. D'un point de vue textural, seuls les coefficients de Haralick ont montré un intérêt en termes de gain de précision dans le cadre du processus de classification par zones. Rappelons que pour chaque attribut spectral, il est possible de calculer six coefficients de Haralick mais qu'il est préférable de conserver un nombre similaire d'attributs texturaux et spectraux. Dans le cadre de la classification vectorielle, nous appliquons une ACP à l'ensemble des attributs texturaux et nous conservons autant de composantes qu'il y avait d'attributs spectraux. Or, le processus de classification du présent chapitre ne fait plus intervenir des pixels mais un nombre restreint de zones, trop restreint pour justifier l'usage d'une ACP. Afin de limiter le nombre d'attributs texturaux, deux méthodes sont proposées. La première consiste à sélectionner les trois coefficients de Haralick les plus informatifs (homogénéité, second moment angulaire et entropie, par ordre décroissant) calculés sur les premières composantes ACP de l'image jusqu'à parvenir au nombre d'attributs spectraux retenus pour la classification. Par exemple pour

quatre attributs spectraux, on ne conservera que les attributs d'homogénéité, de second moment angulaire et d'entropie calculés sur la première composante ACP de l'image, plus l'attribut d'homogénéité calculé sur la seconde. La seconde méthode consiste quant à elle à calculer pour chaque attribut spectral le coefficient de Haralick le plus informatif, à savoir le coefficient d'homogénéité. Par la suite, nous considérerons trois stratégies de classification : SS (Seulement Spectral), qui ne fait intervenir que des attributs spectraux tandis que ST1 et ST2 (Spectral Textural 1 et 2) appliquent la classification hiérarchique ascendante sur des attributs à la fois spectraux et texturaux, ces derniers étant respectivement sélectionnés par l'intermédiaire des deux méthodes mentionnées plus haut.

Le tableau 4.2 confirme que la valeur des seuils α_2 et α_3 utilisée par Jing et al. [23], à savoir $\alpha_2 = \alpha_3 = 0,4$ donne les meilleurs résultats. Les valeurs plus élevées que nous avons testées rendent la sélection des pixels graines moins sévère, ce qui entraîne la fusion de zones normalement distinctes avant même la fin du processus de segmentation. Le NC optimal constaté est variable selon l'image considéré. Toutefois, ce paramètre apparaît comme le moins décisif de ce processus de classification puisque les différences en termes de précision ne sont que de l'ordre de quelques pourcents. En ce qui concerne la taille de voisinage V , il convient de distinguer deux cas, les images MR et MR2 d'une part et l'image CO d'autre part. Le premier type d'image comporte des textures de fréquence relativement élevée, ce qui justifie l'efficacité de voisinages relativement réduit. De plus, les classes associées à ces jeux de données sont telles que l'information texturale s'y révèle utile en tant que complément de l'information spectrale sans être indispensable à leur discrimination. C'est pourquoi les écarts de précision d'une taille de voisinage à l'autre sont plus resserrés qu'avec l'image CO. Cette dernière comporte pour sa part des textures à la fréquence plus faible dont la détection nécessite un voisinage d'étude plus large.

Stratégie	Image		
	MR	MR2	CO
SS	90,8	95,3	97,8
ST1	97	88,2	97,8
ST2	97	95,6	97,8

TAB. 4.1: Précision de classification obtenue avec les stratégies SS, ST1 et ST2, sachant que les paramètres α_2 , α_3 , NC, CS et V ont été fixés à leur valeur optimale pour chaque image.

Paramètre	Valeur	Image		
		MR	MR2	CO
$\alpha_2 = \alpha_3$	0,4	71,5	75,6	92,1
	0,7	71,5	72,7	89,5
	1,0	70,1	68,48	88,4
NC	2K	70,4	73,1	90,2
	3K	70,6	72,4	91,5
	4K	71,3	75	88,3
CS	RC	87,7	81,5	80,1
	VC	87,7	81,5	84,5
	VM	70,1	74	88,4
	MAX	55,4	61,2	83,1
V	11	70,7	74,2	64,6
	15	74,1	72,4	69,4
	19	72,6	76	87,7
	23	72,2	74,1	91,8
	27	68,1	71,2	94
	31	68,5	65,5	96,5

TAB. 4.2: Précision moyenne (en %) des résultats obtenus sur les jeux de données MR, MR2 et CO par l'intermédiaire du processus de classification par zones. Chaque ligne présente la précision moyenne obtenue sur ces trois images lorsque l'on fixe un des paramètres de l'algorithme : les seuils α_2 et α_3 associés à l'identification des pixels graine, le nombre (NC) de sous-classes de la quantification de départ, le critère de sélection (CS) du nombre d'attributs spectraux et/ou texturaux utilisés durant la classification finale et enfin la taille (V) du voisinage étudié pour la création des J-images. L'impact de ces paramètres étant similaire quelle que soit la stratégie de sélection d'attributs employée, les résultats présentés dans ce tableau sont également moyennés sur les stratégies SS, ST1 et ST2.

Remarquons également que dans ce cas, ce paramètre est d'autant plus important que la majorité des classes de l'image CO présentent une grande variabilité spectrale rendant la prise en compte de l'information texturale indispensable à leur discrimination.

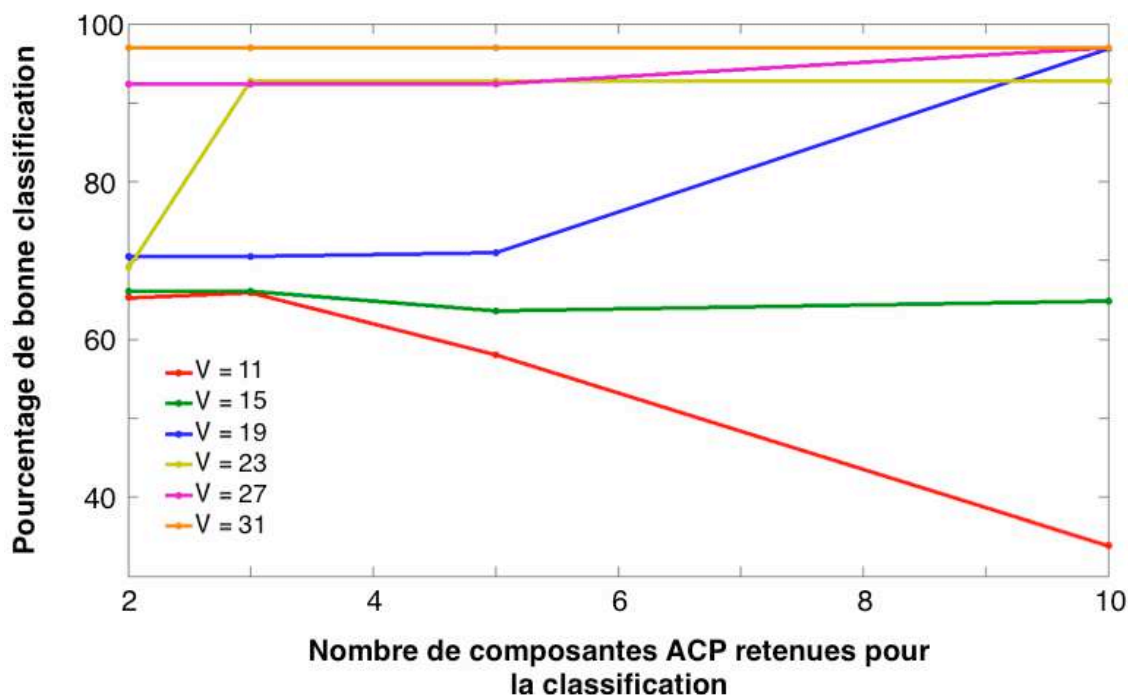


FIG. 4.8: Influence de la taille du voisinage V sur la précision de classification en fonction du nombre de composantes ACP utilisées (image CO).

Enfin, le paramètre le plus délicat à fixer est le critère de sélection du nombre de bandes à conserver pour la classification (CS). En effet si pour les images MR et MR2 les critères de sélection les plus restrictifs tels que RC et VC donnent de meilleurs résultats, la prise en compte d'un nombre trop élevé de bandes ACP diminuant très fortement la précision de la classification, le cas de l'image CO apparaît pour sa part plus complexe. La figure 4.8 montre en effet que l'impact du critère de sélection du nombre de composantes ACP sur la précision de classification est dépendant de V . Le rapprochement de ce dernier de sa valeur optimale implique d'une part une augmentation de la précision moyenne des résultats, bien sûr, mais également une diminution de la sensibilité de l'algorithme vis-à-vis du paramètre CS. Ce comportement propre à l'image CO résulte de la prépondérance de l'information texturale au sein de ce jeu de données.

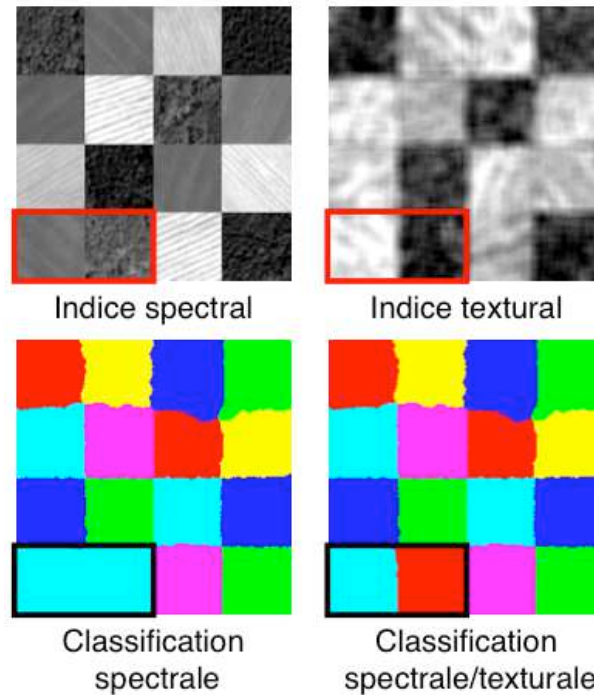


FIG. 4.9: Illustration de l'apport de l'information de texture dans le cas d'images présentant des zones spectralement proches appartenant à des classes distinctes.

Le tableau 4.1 montre que les résultats de classification obtenus avec des attributs spectraux uniquement sont déjà très bons. L'ajout d'attributs texturaux permet toutefois d'éviter certaines erreurs de classification, par exemple lorsque deux zones très proches spectralement ne sont différenciables que par l'intermédiaire de la texture (voir figure 4.9). La stratégie S2 de limitation du nombre d'attributs texturaux apparaît comme la plus efficace car non seulement elle améliore la précision de la classification lorsque l'information spectrale seule est insuffisante, mais elle ne la perturbe pas dans le cas contraire. La figure 4.10 retrace l'ensemble des étapes qui constituent le processus de classification par zone, du calcul de l'image d'hétérogénéité au résultat de classification en passant par la phase de détermination des graines et de segmentation par croissance de régions. On peut y constater que le critère d'hétérogénéité h ne peut éliminer complètement le phénomène de sur-segmentation mais que les fusions successives de la classification hiérarchique ascendante permettent, sous réserve de paramétrer correctement l'algorithme, d'obtenir au terme du processus des zones suffisamment fidèles à la vérité terrain. Les erreurs de classification se situent principalement au niveau des contours entre zones et sont dues à l'utilisation de voisinage relativement large pour la production de la J-image à partir de laquelle est effectuée la segmentation.

	MR	MR2	CO
Image			
J-image			
Graines			
Segmentation			
VT			
Classification			

FIG. 4.10: Illustration des diverses étapes du processus de classification par zones.

5 Conclusion

Dans ce chapitre, nous avons présenté une troisième approche de classification basée sur le traitement non plus de pixels mais de zones préalablement définies par l'intermédiaire d'un algorithme de segmentation par croissance de régions. Nous nous sommes inspiré d'une approche de segmentation proposée par Deng et al. [12], intéressante car elle intègre la notion de texture au processus de segmentation afin de minimiser les effets de sur-segmentation. Nous avons ensuite caractérisé chaque zone avec plusieurs attributs spectraux et/ou texturaux avant de les fusionner au moyen d'un algorithme de classification hiérarchique ascendante.

L'ensemble du processus dépendant d'un grand nombre de paramètres, nous avons procédé à une étude de sensibilité visant non seulement à mesurer la précision de la méthode mais également sa robustesse. Sur les trois images testées, la qualité des résultats s'est révélée particulièrement bonne, surtout si l'on considère que ce processus de classification est non supervisé. Il convient toutefois de préciser que nous considérons connu à l'avance le nombre de classes de l'image afin de stopper la fusion de zones durant la classification hiérarchique ascendante. D'autre part, le processus de classification fait intervenir de nombreux paramètres. Les seuils associés à la détermination des pixels graines durant l'étape de segmentation constituent des paramètres assez robustes. En effet, les valeurs proposées par Jing et al. [23] représentent un compromis efficace entre sur-segmentation et sous-segmentation. Le nombre de classes de l'étape initiale de quantification est quant à lui moins déterminant puisque les différences de précision qu'il occasionne ne sont que de l'ordre de quelques pourcents. Le critère de sélection du nombre de composantes ACP à conserver pour la phase de classification finale est de son côté plus délicat à fixer, sa valeur optimale pouvant dépendre du jeu de données à classifier. Toutefois, ce défaut disparaît si la taille du voisinage d'étude est correctement fixée. Ce dernier paramètre constitue ainsi le plus important du processus de classification. Or, si sa valeur optimale diffère d'un jeu de données à l'autre, elle peut être définie précisément en fonction de la fréquence des textures en présence, l'idéal étant de choisir un voisinage recouvrant au minimum trois répétitions de motifs d'une même texture.

Les méthodes présentées dans les trois précédents chapitres constituent trois manières bien distinctes de mêler informations spectrale et spatiale au sein d'un même processus de classification. L'objectif du chapitre suivant est de comparer ces trois méthodes afin d'identifier les avantages et inconvénients relatifs de chacune ainsi que leur contexte d'application recommandé.

Chapitre 5

Comparatif des approches de classification spectrale/spatiale

Dans les chapitres précédents, nous avons présenté trois approches différentes pour utiliser simultanément les informations spectrale et spatiale dans le cadre de la classification de données hyperspectrales. Elles nous ont permis de développer quatre algorithmes de classification, deux supervisés (SVM et CRF) et deux non supervisés (GMM et AC). Le présent chapitre revient sur chacun de ces algorithmes (appliqués aux images MR, MR2 et CO) afin d'en identifier les avantages et les inconvénients par l'intermédiaire de trois critères de comparaison : la précision des résultats, la complexité algorithmique et la robustesse de la méthode.

1 Précision des résultats

Pour mesurer la qualité des cartes de classification obtenues au terme de chaque algorithme, nous utilisons toujours le critère de précision générale présenté dans le chapitre 2. La phase d'apprentissage des algorithmes supervisés a été réalisée pour chaque jeu de données sur 1% des éléments de la vérité terrain disponible, le reste constituant le jeu d'échantillons de test. Le tableau 5.1 liste les précisions obtenues par chaque algorithme sur les trois images testées avec le paramétrage optimal. L'algorithme supervisé le plus efficace est sans conteste la classification vectorielle SVM d'attributs spectraux/spatiaux, la classification CRF ne parvenant au même niveau de précision qu'avec l'image CO. Les algorithmes non supervisés donnent pour leur part des résultats relativement proches, parvenant même la plupart du temps à surpasser les résultats de la classification CRF sans aucune information a priori si ce n'est le nombre de classes.

Les cartes de classification de la figure 5.1 reflètent bien les différences de précision remarquées entre les résultats des deux processus de classification supervisée. En revanche, bien que présentant des précisions globales similaires, les

Algorithme	Image		
	MR	MR2	CO
SVM	99,6	99,1	98,4
CRF	94,6	86,7	98
GMM	96,2	96,8	97,1
AC	97	95,6	97,8

TAB. 5.1: Précision (en %) optimale obtenue avec chaque algorithme sur les trois jeux de données. SVM et GMM (*Gaussian Mixture Model*) sont les deux algorithmes de classification vectorielle présentés dans le chapitre 2. CRF est la classification à base de champs markovien du chapitre 3 et AC (*Area Classification*) la classification par zones du chapitre 4.

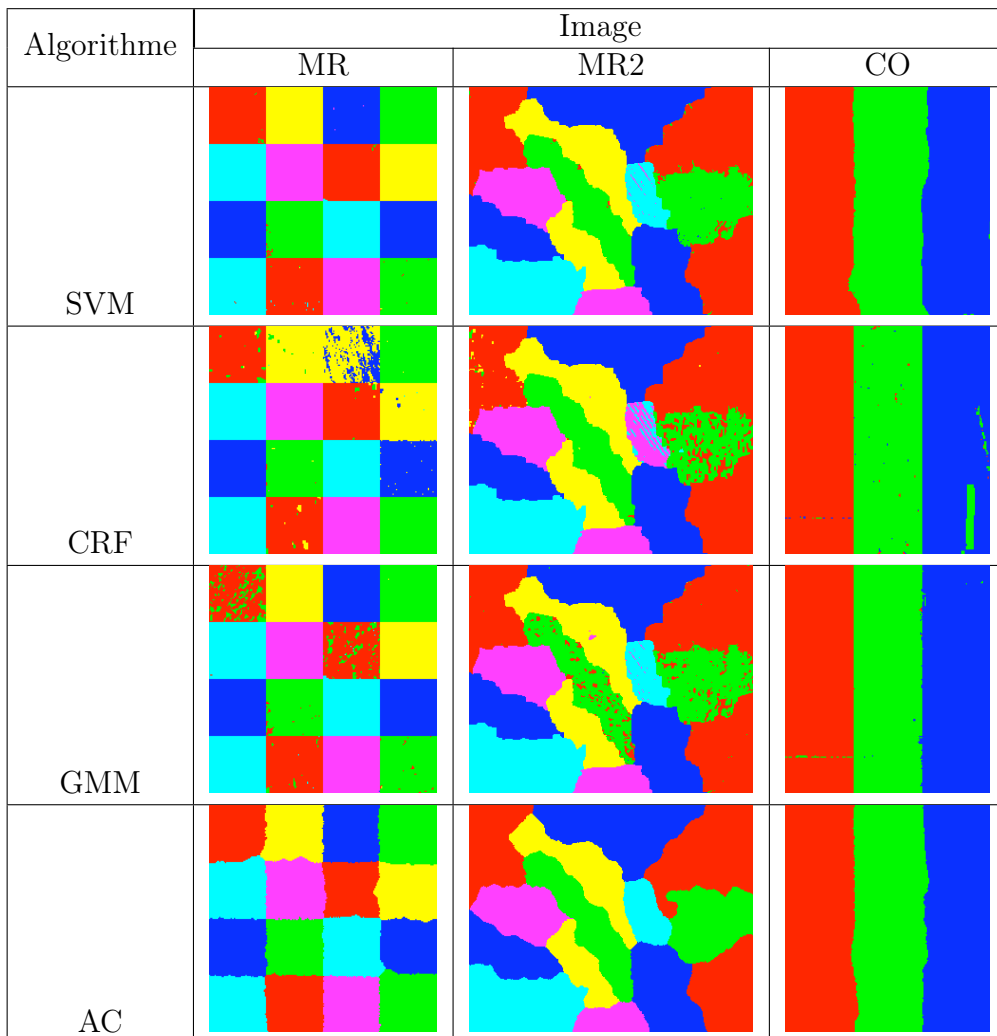


FIG. 5.1: Cartes de classification associées aux statistiques présentées par le tableau 5.1

méthodes non supervisées aboutissent de leur côté à des cartes de classification sensiblement différentes. En effet, si l’algorithme GMM dessine des contours entre zones de classes différentes plus précis, c’est au prix d’un effet "poivre et sel" plus ou moins prononcé selon la classe. A l’inverse, les zones identifiées à l’issue de la classification par zone sont beaucoup plus homogènes mais leurs frontières sont moins précises.

2 Complexité algorithmique

Le second critère de comparaison des algorithmes est le temps de calcul. Pour les algorithmes SVM, GMM et AC, nous avons utilisé une taille de voisinage intermédiaire : 21×21 . Les attributs texturaux de ces mêmes algorithmes ont quant à eux été calculés sur les cinq premières composantes ACP dans chaque cas

Algorithme	Image								
	MR			MR2			CO		
SVM	ST	CH	PM	ST	CH	PM	ST	CH	PM
	149	239	327	296	437	808	126	226	340
CRF	M1		M2		M1		M2		
	300		252		526		475		114 86
GMM	ST	CH	PM	ST	CH	PM	ST	CH	PM
	197	311	415	404	614	1021	168	286	414
AC	127			240			79		

TAB. 5.2: Complexité algorithmique (en secondes) des quatre algorithmes. Pour SVM et GMM nous avons distingué trois scénarios, un pour chaque type d’attribut spatial : les spectres de texture (ST), les coefficients de Haralick (CH) et les profils morphologiques (PM). Nous présentons deux scénarios pour l’algorithme CRF. L’algorithme est appliqué aux données non réduites dans le premier (M1) et réduites dans le second (M2), avec des cliques d’ordre 2 représentées par des concaténations de vecteurs voisins.

Le tableau 5.2 compare sur les images CO, MR et MR2 la complexité des diverses méthodes présentées durant cette étude. Les algorithmes SVM et GMM sont composés de deux étapes : l’extraction des attributs spectraux/texturaux et la classification du jeu de données résultant. Dans les deux cas, c’est la première étape qui est la plus longue. Si l’analyse en composantes principales peut être réalisée de manière très efficace au moyen de calculs matriciels, la création des trois types d’attributs texturaux nécessite d’analyser le voisinage de chaque pixel, ce qui ne peut être fait qu’en parcourant l’intégralité de l’image. La complexité d’une telle démarche dépend donc à la fois du nombre de pixels de l’image, de la taille du

voisinage à étudier et du nombre de composantes principales à partir desquelles on désire extraire les attributs. Le calcul des spectres de texture est le plus rapide car l'analyse de voisinage repose sur la création d'histogrammes, une structure de données plus simple à bâtir qu'une matrice de co-occurrence. La complexité plus élevée relative à la production des profils morphologiques s'explique quant à elle par les opérations de reconstruction qu'elle met en oeuvre, la répétition jusqu'à idempotence des opérations d'ouverture et de fermeture morphologique se révélant relativement longue. En ce qui concerne l'étape de classification, l'algorithme SVM est de loin le plus rapide. La complexité de ce dernier repose en effet essentiellement sur le nombre d'échantillons de vérité terrain considérés pour la phase d'apprentissage, sachant que nous n'en utilisons ici qu'une très faible proportion (1%). La complexité plus élevée de l'algorithme GMM s'explique par le caractère paramétrique de la méthode impliquant une phase d'optimisation onéreuse en temps de calcul, malgré le modèle simplifié utilisé.

L'étape d'apprentissage de l'algorithme CRF consiste également à optimiser une série de paramètres mais contrairement à la méthode précédente, elle s'appuie sur un ensemble d'échantillons supposé représentatif des classes à discriminer. La complexité de cette étape dépend ainsi à la fois de la taille de cet ensemble d'échantillons, du nombre de paramètres à estimer - proportionnel au nombre de bandes de l'image à classifier - et du nombre d'itérations nécessaire à la convergence de ces derniers. Les temps de calcul présentés dans le tableau 5.2 ont été obtenus avec un nombre d'itération égal à 200, comme dans l'article de Zhong et al. [58]. Nous verrons toutefois dans la section portant sur la robustesse des algorithmes pourquoi ce choix est problématique. Afin de propager l'information spatiale d'un pixel à l'autre, l'étape de classification LBP utilise un système de messages spectraux et spatiaux qu'elle estime pour l'ensemble des pixels de l'image de manière itérative. Cela explique sa complexité plus élevée vis-à-vis de l'étape d'apprentissage qui se focalise sur une faible portion des pixels de la vérité terrain.

Enfin, l'algorithme AC se divise en quatre étapes : la quantification, la segmentation, le calcul d'attributs de zone puis la fusion de ces dernières par classification hiérarchique ascendante. La quantification par K-means est l'étape la plus onéreuse en temps de calcul. Elle dépend à la fois du nombre de classes à rechercher et des nombres de pixels et de bandes dans l'image. De plus, la dépendance du résultat à l'initialisation de l'algorithme nous contraint à en lancer plusieurs occurrences (une dizaine, en ce qui nous concerne) pour ne conserver que la meilleure. L'étape suivante réalise successivement la création de l'image d'hétérogénéité, la détermination des pixels graines et la croissance de régions. Bien que la création de l'image d'hétérogénéité nécessite l'analyse du voisinage de l'ensemble des pixels de l'image, elle implique une complexité plus faible que les attributs texturaux mentionnés plus haut car le calcul de l'indice d'hétérogénéité associé à chaque pixel met en oeuvre un nombre moindre d'opérations arithmétiques. Les deux sous-étapes suivantes ont quant à elle une complexité

linéaire vis-à-vis du nombre de pixels de l'image et sont par conséquent réalisées très rapidement. Une fois la carte de segmentation obtenue, le calcul d'attributs et la classification hiérarchique ascendante qui suit sont très rapides car elle ne dépendent plus du nombre de pixels de l'image originale mais du nombre de zones, beaucoup plus réduit, identifiées par la segmentation.

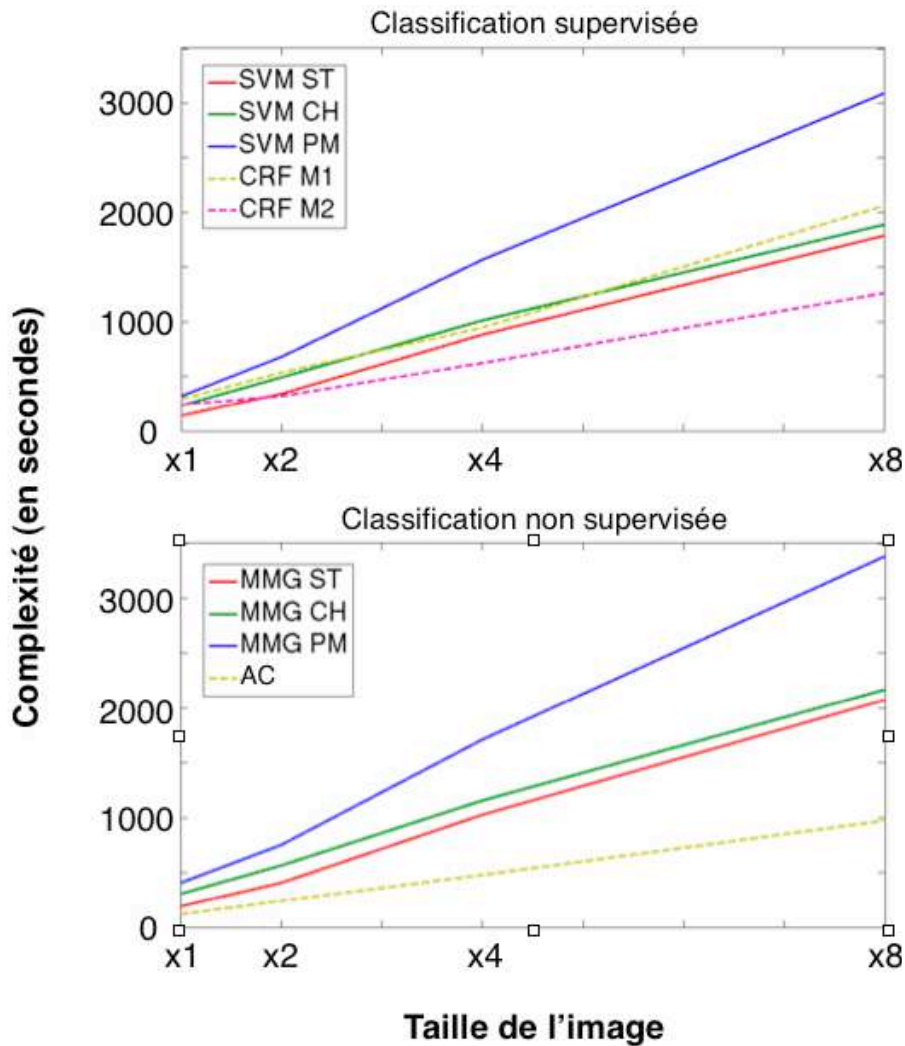


FIG. 5.2: Evolution des temps de calculs en fonction de la taille des jeux de données. Pour effectuer cette étude, nous avons appliqué chaque algorithme sur l'image MR multiplié par un coefficient 1, 2, 4 et 8.

La figure 5.2 montre que pour l'ensemble des algorithmes développés durant cette étude, les temps de calculs évoluent linéairement en fonction du nombre de pixels présents dans l'image. Pour conclure, s'il existe effectivement des méthodes plus rapides que d'autres, la différence n'est pas assez importante pour que le critère de complexité algorithmique passe avant celui de précision des résultats.

3 Robustesse

La robustesse d'un algorithme se mesure à l'aune de la dépendance existant entre la qualité de ses résultats et son paramétrage. Il faut cependant distinguer deux types de paramètres : les paramètres externes et internes. Les paramètres externes sont fixés une fois pour toutes par l'utilisateur et n'évoluent pas au fur et à mesure de l'exécution du programme. Un algorithme est considéré robuste vis-à-vis de ce type de paramètre dans deux cas : soit la valeur du paramètre n'impacte la précision du résultat que de manière superficielle (le nombre de sous-classes de l'étape de quantification de l'algorithme AC), soit elle est aisément définissable, quitte à requérir une analyse préalable des données par l'utilisateur (comme la taille de la fenêtre utilisée pour l'analyse de textures dans les algorithmes SVM, GMM et AC).

Par ailleurs, certaines étapes des algorithmes décrits durant cette étude reposent sur l'élaboration d'un modèle paramétrique représenté par une série de paramètres internes optimisés de manière itérative par le programme. Ces paramètres ne sont donc pas fixés par l'utilisateur mais doivent être initialisés. Un algorithme est considéré robuste vis-à-vis de ce type de paramètres s'il est indépendant de leur initialisation et si les paramètres en question convergent vers leur valeur optimale en un temps raisonnable. En fonction du caractère supervisé ou non des algorithmes, un facteur supplémentaire doit être considéré dans le cadre d'une estimation de la robustesse : l'influence du nombre d'échantillons utilisés pour l'apprentissage sur la qualité des résultats de classification dans un contexte supervisé et le fait que le nombre de classes est connu à l'avance dans un contexte non supervisé.

Le processus de classification vectorielle, supervisé ou non, repose sur une phase d'extraction d'attributs spectraux et texturaux dépendant de nombreux paramètres externes. Nous avons montré dans le chapitre 2 que la plupart de ces paramètres sont relativement faciles à fixer même si l'estimation de certains d'entre eux, dépendants du type de texture présent dans l'image à classifier, nécessite une analyse préalable de l'image par l'utilisateur. D'un point de vue supervisé, la classification SVM bénéficie de deux avantages. Tout d'abord, le noyau polynomial de degré deux utilisé s'est montré particulièrement efficace sur les trois jeux de données testés. De plus, ce type d'algorithme donne de bons résultats même lorsque l'on dispose d'un faible nombre d'échantillons d'apprentissage - ce qui est généralement le cas dans le contexte d'une application de classification réaliste - tant que ces derniers sont assez représentatifs des classes de l'image. La classification SVM est ainsi la plus robuste parmi les quatre méthodes développées dans cette étude. En revanche en ce qui concerne la classification GMM, plusieurs stratégies d'initialisation des paramètres internes du modèle ont été expérimentées mais dans tous les cas, nous avons observé des précisions globales de classification pouvant varier sensiblement d'une initialisation à une autre.

Comparée aux deux méthodes précédentes, la classification CRF présente un

nombre relativement restreint de paramètres externes, ce qui la rend beaucoup plus simple à paramétrer. La phase d'apprentissage de son modèle de représentation des données nécessite toutefois un nombre bien plus conséquent d'échantillons que l'algorithme SVM pour bâtir un modèle robuste capable de classer correctement le reste de l'image. La création de ce modèle implique de plus un grand nombre de paramètres internes qui nécessitent bien plus de 200 itérations pour converger. Les deux variantes proposées pour diminuer ce nombre de paramètres minimisent l'impact de ces deux défauts au prix d'une faible diminution de la précision de classification. La robustesse du modèle n'est cependant pas toujours parfaite et la plus-value apportée par l'information spatiale est généralement plus mesurée, quand elle n'est pas presque négligeable, avec ces méthodes alternatives.

Enfin, la classification AC comporte un nombre de paramètres externes similaire à celui des processus de classification vectorielle mais contrairement à l'algorithme GMM, il ne repose sur aucun modèle paramétrique de représentation des données. Ses résultats sont ainsi moins susceptibles d'être détériorés par une initialisation hasardeuse des paramètres d'un tel modèle. Les résultats de l'algorithme K-means utilisé pour effectuer la quantification peuvent être sensibles aux conditions initiales de ce dernier mais cette étape intervenant en tant que pré-traitement de la phase de segmentation, elle ne présente qu'un impact très limité sur l'ensemble du processus de classification. En contrepartie, le mode de fonctionnement de l'algorithme AC basé sur la fusion de larges zones rend les erreurs de classification plus pénalisantes que pour un algorithme de classification pixel à pixel.

4 Conclusion

Dans cet ultime chapitre, nous avons évalué chacun des quatre algorithmes développés durant cette étude à la lumière de trois critères de comparaison. Le tableau 5.3 présente un résumé global des avantages et inconvénients de ces méthodes. D'un point de vue supervisé, la classification vectorielle SVM apparaît à la fois plus précise et plus stable que la classification CRF. Quoique légèrement moins efficace en termes de complexité algorithmique, sa capacité à fonctionner de façon robuste avec un faible nombre d'échantillons d'apprentissage tout en donnant les meilleurs résultats parmi les quatre méthodes testées nous permet de l'utiliser efficacement dans un contexte supervisé réaliste où de tels échantillons sont disponibles en nombre restreint.

Les algorithmes de classification non supervisée sont pour leur part plus difficiles à départager. Ils aboutissent à des précisions globales similaires mais à des cartes de classification sensiblement différentes, ce qui nous empêche d'écarter complètement l'algorithme GMM malgré des temps de calcul beaucoup plus élevés et une moins bonne robustesse due à la construction de son modèle paramétrique. Ces deux méthodes représentent en effet chacune une philosophie

Critère	Algorithme			
	SVM	CRF	GMM	AC
Précision	++	-	+	+
Complexité	-	+	-	++
Robustesse	++	-	+	+

TAB. 5.3: Récapitulatif des avantages et inconvénients de chaque algorithme

de classification bien distincte, l'algorithme GMM permettant un dessin précis des contours entre zones de classes différentes tandis que les cartes de classification obtenues avec l'algorithme AC présentent des zones très homogènes et peu bruitées.

Conclusion et perspectives

Au cours de cette thèse, nous avons étudié la complémentarité des informations spectrale et spatiale dans le cadre de l'utilisation de données hyperspectrales à des fins de classification. D'un point de vue spatial, nous nous sommes intéressés tout particulièrement à l'information texturale et à la façon dont elle permet de discriminer les classes spectralement trop similaires. Quatre jeux de données présentant des propriétés texturales bien distinctes ont été utilisés pour tester trois scénarios de classification, chacun incarnant une manière différente de tirer parti du contexte spatial.

Le premier scénario développé est un processus de classification vectorielle de divers attributs spectraux et spatiaux, respectivement extraits de l'image originale par réduction de dimension et analyse texturale ou morphologique du voisinage de chaque pixel. Les jeux d'attributs ainsi créés sont ensuite classifiés de manière supervisée avec un algorithme SVM ou non supervisée avec un modèle de mélange gaussien (GMM). Une étude de sensibilité des paramètres associés aux divers attributs spatiaux nous a permis d'identifier les forces et les faiblesses de chacun, ainsi que le type de classes qu'ils sont le plus à même de discriminer. Coefficients de Haralick et spectres de texture se sont ainsi révélés plus aptes à discriminer des classes présentant des textures de faible fréquence car ils favorisent l'analyse de larges voisinages, tandis que les profils morphologiques donnent de meilleurs résultats avec des textures de fréquence plus élevée, tout en conservant un dessin précis des contours entre classes. Robuste et performant en termes de précision de classification, l'algorithme SVM présente de plus des propriétés de généralisation très puissantes. En effet, la proportion de connaissance a priori nécessaire à son bon fonctionnement est très réduite. La classification GMM donne également d'excellents résultats, surtout si l'on considère son caractère non supervisé, mais le modèle de représentation des données sur lequel elle repose comporte des paramètres dont l'initialisation est délicate et peut donner lieu à des différences de précision de classification allant jusqu'à une dizaine de pourcents d'une initialisation à l'autre. Afin de gagner en robustesse, il conviendrait ainsi de développer un système d'initialisation minimisant les risques de converger vers un minimum local. De plus, cet algorithme supposant le nombre de classe final connu par l'utilisateur, il serait pertinent de définir un critère permettant de le fixer de

manière automatique. Celui proposé par Mixmod est pertinent dans le cadre d'une classification purement spectrale mais aboutit généralement à un nombre de classes trop élevé lorsque les classes sont plus complexes. Précisons enfin qu'il existe de nombreux autres types d'attribut spatial qu'il serait possible de substituer aux trois testés durant cette étude. Outre les attributs texturaux associés au modèle markovien gaussien présenté par Rellier [41], mentionnons également les attributs LBP (*Local Binary Pattern*), LEP (*Local Edge Pattern*) utilisés dans le cadre d'une classification non supervisée par Vatsavai et al. [54] ainsi que les histogrammes de gradients orientés [11].

Le second scénario introduit quant à lui un processus de classification intégrant directement l'information contextuelle existant entre pixels voisins dans son modèle de représentation des données. Une fois construit par optimisation itérative, le modèle en question - de type *Conditional Random Field* - est utilisé comme base d'une classification par Maximum A Posteriori. L'algorithme original proposé par Zhong et al. [58] présente deux défauts. D'une part la construction d'un modèle robuste avec de bonnes propriétés de généralisation nécessite l'utilisation d'un nombre trop conséquent d'échantillons de vérité terrain, bien plus que la classification SVM. D'autre part, la convergence des paramètres du modèle durant la phase d'apprentissage est beaucoup trop longue, conséquence d'un nombre de paramètres énorme, proportionnel à la fois au nombre de classes de l'image et à sa dimension spectrale. Afin de diminuer le nombre de ces paramètres, nous avons proposé deux modifications - utilisables simultanément - de l'algorithme. La première consiste à réduire la dimension des jeux de données à classifier au préalable et la seconde à modifier la manière dont l'algorithme modélise les relations de voisinages entre pixels. Ces modifications parviennent effectivement à renforcer la robustesse de l'algorithme mais cette dernière reste tout de même inférieure à celle de l'algorithme SVM mentionné dans le paragraphe précédent. De plus, elles ont tendance à diminuer la plus-value apportée par l'information spatiale dans la classification. De tels résultats n'invalident en rien l'approche markovienne pour la représentation des interactions spatiales entre pixels voisins mais nous encourage à étudier d'autres modèles, plus efficaces aussi bien en termes de robustesse que de précision de classification.

Enfin, le troisième scénario délaisse l'approche de classification pixel à pixel pour se focaliser sur une méthode bien différente : la classification par zones. L'information texturale y est utilisée lors d'une phase de segmentation préalable de l'image originale visant à mettre en évidence l'existence de zones spectralement et texturalement homogènes tout en minimisant l'effet de sur-segmentation. Les zones ainsi obtenues sont alors associées à un vecteur d'attributs spectraux et/ou texturaux avant d'être fusionnées les unes aux autres par classification hiérarchique ascendante. De même que les méthodes développées dans le cadre du

premier scénario, la classification par zones dépend d'un grand nombre de paramètres. L'étude de sensibilité réalisée sur ces derniers a révélé que le plus crucial est la taille du voisinage d'étude, sa valeur optimale dépendant de la fréquence des textures présentes dans l'image à classifier. Les résultats acquis sur les trois jeux de données évalués se sont montrés à la fois précis et rapides à obtenir, surtout si l'on considère la nature non supervisée de l'algorithme. Du fait de sa nature itérative, l'étape finale de fusion de zones par classification hiérarchique comporte toutefois le risque, si une erreur de fusion a lieu dès le début du processus, de propager cette erreur jusqu'à son terme. Il serait ainsi judicieux à l'avenir de substituer à cette étape un algorithme capable de considérer plusieurs scénarios de fusion simultanément, puis de déterminer le meilleur par l'intermédiaire d'un critère à définir.

Comparés les uns aux autres, la plupart des algorithmes dérivés de ces trois scénarios se démarquent dans un cas de figure bien spécifique. Lorsque des échantillons de vérité terrain sont disponibles, l'algorithme SVM est de loin le plus efficace, aussi bien en termes de robustesse que de précision de classification. L'utilisation d'attributs texturaux calculés sur de vastes voisinages lui permet de mettre à profit plus efficacement l'information spatiale, là où la faible portée spatiale de l'algorithme CRF peine à mettre en évidence les structures texturales de faible fréquence. Les algorithmes de classification non supervisée GMM et AC, sont pour leur part adaptés aux situations pour lesquelles aucune vérité terrain n'est disponible. Les précisions globales obtenues avec l'une ou l'autre étant très similaires, ces deux méthodes se distinguent principalement par le type de carte de classification produite. Malgré la grande portée spatiale de certains de ses attributs texturaux, l'approche de classification pixel à pixel de l'algorithme GMM permet de dessiner des contours plus précis entre zones de classes différentes là où la phase de segmentation par croissance de région de l'algorithme AC résulte sur des zones aux contours moins bien définis. Toutefois, les cartes de classification obtenues par l'intermédiaire du premier algorithme présentent en certains endroits un aspect poivre et sel que l'on ne peut pas aisément supprimer avec des méthodes de type vote majoritaire. Au contraire, les cartes de classification identifiées par l'étape de segmentation par croissance de régions de l'algorithme AC sont parfaitement homogènes.

D'un point de vue plus global, la plus-value apportée par l'information spatiale dans le cadre de la classification thématique de données varie beaucoup selon l'application envisagée. Considérons tout d'abord MR et MR2, deux jeux de données représentatifs d'une application visant à discriminer plusieurs types de culture représentés par un faible nombre de signatures spectrales différentes ainsi qu'un motif textural propre. Pour ce type d'application, l'information spatiale devient déterminante lorsque les classes ne sont différenciables les unes des autres que par l'intermédiaire de leur motif textural. Or, les six classes de MR et MR2 sont cha-

cune constituées de spectres exclusivement représentatifs - c'est-à-dire que chaque type de spectre peut être automatiquement associé à une classe unique - ce qui implique qu'une classification purement spectrale est tout à fait capable de faire la distinction entre elles, notamment dans un contexte supervisé. L'information spatiale joue alors le rôle d'une information d'appoint favorisant la régularisation de la carte de classification obtenue à l'issue du processus. Afin de mesurer son véritable potentiel dans le cadre de ce type d'application, il conviendrait ainsi de tester les divers algorithmes développés durant cette étude sur des images comportant des classes plus à même de la mettre en valeur. Le cas du jeux de données CO est beaucoup plus intéressant. Ce dernier est l'exemple type d'une application de classification visant à distinguer plusieurs formes d'aménagement du territoire, les classes considérées étant représentées par un éventail extrêmement large de signatures spectrales différentes - avec le risque pour certaines d'être communes à plusieurs classes. Pour cette catégorie d'application poussée à l'extrême, l'information texturale devient plus importante que l'information spectrale, ce qui implique qu'un algorithme de classification purement spectrale est incapable d'atteindre l'objectif fixé. Afin de discriminer ces classes très complexes d'un point de vue sémantique, l'intervention de l'information texturale s'avère ainsi absolument cruciale.

Bibliographie

- [1] S. C. Ahn and A. R. Horenstein. Eigenvalue ratio test for the number of factors, 2009.
- [2] P. Bajcsy and P. Groves. Methodology for hyperspectral band selection. *Photogrammetric Engineering and Remote Sensing*, 70 :793–802, 2004.
- [3] J. A. Benediktsson, J. A. Palmason, and J. Sveinsson. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3) :480–491, 2005.
- [4] J. A. Benediktsson, J. A. Palmason, J. R. Sveisson, and J. Chanussot. Decision level fusion in classification of hyperspectral data from urban areas. In *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings*, pages 73–76, 2004.
- [5] J. Besag. Statistical analysis of non-lattice data. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 24(3) :179–195, 1975.
- [6] F. Bovolo and L. Bruzzone. A context-sensitive technique based on support vector machines for image classification. In Sankar Pal, Sanghamitra Bandyopadhyay, and Sambhunath Biswas, editors, *Pattern Recognition and Machine Intelligence*, volume 3776 of *Lecture Notes in Computer Science*, pages 260–265. Springer Berlin / Heidelberg, 2005.
- [7] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.*, 2 :121–167, 1989.
- [8] J. Chanussot. *Approches vectorielles ou marginales pour le traitement d'images multi-composantes*. PhD thesis, Université de Savoie, 1998.
- [9] R. Cossu. Segmentation by means of textural analysis. *Pixel*, 1(2) :21–24, 1988.
- [10] T. M. Cover and P. E. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, pages 21–27, 1967.
- [11] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 886–893, 2005.

- [12] Y Deng and B S Manjunath. Unsupervised segmentation of color-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8) :800–810, 2001.
- [13] R. Dianat and S. Kasaei. Dimension reduction of optical remote sensing images via minimum change rate deviation method. *IEEE Transactions on Geoscience and Remote Sensing*, 48(1) :198–206, 2010.
- [14] D. Dunn, W. E. Higgins, and J. Wakeley. Texture segmentation using 2-d gabor elementary functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(2) :130–149, 1994.
- [15] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson. Spectral and spatial classification of hyperspectral data using svms and morphological profiles. *IEEE Transactions on Geosciences and Remote Sensing*, 46(11) :3804–3814, 2008.
- [16] J. A. Gualtieri and S. Chettri. Support vector machines for classification of hyperspectral data. In *Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International*, volume 2, pages 813–815, 2000.
- [17] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6) :610–621, 1973.
- [18] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1) :97–109, 1970.
- [19] D. C. He and L. Wang. Texture unit, texture spectrum, and texture analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 28(4), 1990.
- [20] D. C. He and L. Wang. Texture features based on texture spectrum. *Pattern Recognition*, pages 1187–1195, 1991.
- [21] R. Hess, A. Fern, and E. Mortensen. Mixture-of-parts pictorial structures for objects with variable part sets. In *Proc. IEEE Int. Conf. Computer Vision*, pages 1–8, 2007.
- [22] Q. Jackson and D. Landgrebe. Adaptive bayesian contextual classification based on markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11) :2454–2463, 2002.
- [23] Feng Jing, Mingjing Li, Hong-Jiang Zhang, and Bo Zhang. Unsupervised image segmentation using local homogeneity analysis. In *Circuits and Systems, 2003. ISCAS '03. Proceedings of the 2003 International Symposium on*, volume 2, pages 456–459, may 2003.
- [24] Z. Kato and T. Pong. A markov random field image segmentation model for color textured images. *Image and Vision Computing*, 24 :1103–1114, 2006.
- [25] T. Kohonen. *Self-organization and associative memory*. Third ed. New York, Berlin : Springer-Verlag, 1989.

-
- [26] B. Krose and P. V. D. Smagt. An introduction to neural network, 1996.
- [27] S. Kumar and M. Hebert. A hierarchical field framework for unified context-based classification. In *Proc. IEEE Int. Conf. Computer Vision*, pages 1284–1291, 2005.
- [28] J. Lafferty, A. McCallum, and F. Pereira. Conditional random fields : Probabilistic models for segmenting and labeling sequence data. In *Proc. Int. Conf. ML*, pages 282–289, 2001.
- [29] D. Landgrebe. *Information extraction principles and methods for multispectral and hyperspectral image data*. C. H. Chen, 2000.
- [30] L. Lepisto, I. Kunttu, J. Autio, and A. Visa. Classification method for colored natural textures using gabor filtering. In *Proceedings of the 12th International Conference on Image Analysis and Processing*, pages 397–401, 2003.
- [31] L. Lepisto, I. Kunttu, J. Autio, and A. Visa. Rock image classification using non-homogenous textures and spectral imaging. In *WSCG SHORT PAPERS Proceedings*, 2003.
- [32] Y. Linde, A. Buzo, and R. M. Gray. An algorithm for vector quantizer design. *IEEE Trans. Commun.*, pages 84–95, 1980.
- [33] *MIXMOD Statistical Documentation*. <http://www.mixmod.org/>, May 2008.
- [34] G Moser and S. B. Serpico. A markovian generalization of support vector machines for contextual supervised classification of hyperspectral images. *Opera*, pages 7–10, 2010.
- [35] J. Alirezaie N. N. Kachouie. Optimized multichannel filter bank with flat frequency response for texture segmentation. *EURASIP Journal on Applied Signal Processing*, pages 1834–1844, 2005.
- [36] J. A. Palmason, J. A. Benediktsson, J. R. Sveinsson, and J. Chanussot. Classification of hyperspectral data from urban areas using morphological preprocessing and independent component analysis. In *International Geoscience And Remote Sensing Symposium*, volume 1, pages 176–179, 2005.
- [37] D. K. Panjwani and G. Healey. Markov random field models for unsupervised segmentation of textured color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17 :939–954, 1995.
- [38] 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*, 39(2) :309–320, 2001.
- [39] W. K. Pratt. *Digital image processing*. New York : Wiley, 2nd edition edition, 1991.
- [40] A. R. Rao. *A taxonomy for texture description and identification*. New York : Springer-Verlag, 1990.

- [41] G. Rellier. *Analyse de textures dans l'espace hyperspectral par des méthodes probabilistes*. PhD thesis, Université de Nice Sophia-Antipolis, 2002.
- [42] S. B. Serpico and L. Bruzzone. A new search algorithm for feature selection in hyperspectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 39 :1360–1367, 2001.
- [43] J. Sklansky. Image segmentation and feature extraction. *IEEE Trans. Syst.*, 8 :237–247, 1978.
- [44] P. Soille. *Morphological Image Analysis, Principles and Applications*. Germany : Springer-Verlag, 2003.
- [45] A. H. Strahler. The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sensing of Environment*, pages 135–163, 1980.
- [46] C. Sutton and A. McCallum. Piecewise pseudolikelihood for efficient training of conditional random fields. In *Proc. Int. Conf. Machine Learning*, pages 863–870, 2007.
- [47] C. C. J. Kuo T. Chang. Texture analysis and classification with tree-structured wavelet transform. *IEEE Trans. Image Processing*, 2(4) :429–441, 1993.
- [48] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot. Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. *IEEE Transactions on Geoscience and Remote Sensing*, 47(8) :2973–2987, 2009.
- [49] Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton. Multiple spectral-spatial classification approach for hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 48(11) :4122–4132, 2010.
- [50] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson. Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers. *IEEE Transactions on Systems, Man, and Cybernetics*, 40(5) :1267–1279, 2010.
- [51] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson. Segmentation and classification of hyperspectral images using watershed transformation. *Pattern Recogn.*, 43 :2367–2379, 2010.
- [52] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson. Svm- and mrf-based method for accurate classification of hyperspectral images. *Geoscience and Remote Sensing Letters, IEEE*, 7(4) :736–740, 2010.
- [53] F. Tupin. Champs de markov sur graphes pour le traitement des images radar, 2007.
- [54] R. R. Vatsavai, A. Cheriyyadat, and S. Gleason. Unsupervised semantic labeling framework for identification of complex facilities in high-resolution

- remote sensing images. In *IEEE International Conference on Data Mining Workshops*, pages 273–280, december 2010.
- [55] L. Wang and D. C. He. A new statistical approach for texture analysis. *Photogrammetric Engineering and Remote Sensing*, 56 :61–66, 1990.
- [56] J. S. Weszka, C. R. Dyer, and A. Rosenfeld. A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-6(4) :269–285, 1976.
- [57] A. Zare and P. Gader. Hyperspectral band selection and endmember detection using sparsity promoting priors. *Geoscience and Remote Sensing Letters, IEEE*, 5(2) :256–260, 2008.
- [58] P. Zhong and R. Wang. Learning conditional random fields for classification of hyperspectral images. *IEEE Transactions on Image Processing, Vol. 19, No 7*, 2010.

BIBLIOGRAPHIE

Développement et évaluation de nouvelles méthodes de classification spatiale-spectrale d'images hyperspectrales

L'imagerie hyperspectrale, grâce à un nombre élevé de bandes spectrales très fines et contiguës, est capable d'associer à chaque pixel d'une image une signature spectrale caractéristique du comportement réflectif du matériau ou du mélange de matériaux présents dans ce pixel. La plupart des algorithmes de classification tirent profit de cette grande profusion d'information spectrale mais exploitent très peu l'information contextuelle existant entre les pixels appartenant à un même voisinage. L'objectif de cette thèse est de réaliser de nouveaux algorithmes utilisant simultanément les informations spectrale et spatiale à des fins de classification et d'étudier la complémentarité de ces deux types d'information dans divers contextes. Dans cette optique nous avons développé trois scénarios de classification sensiblement différents, chacun étant adapté à un type d'application particulier.

Nous avons tout d'abord développé un procédé d'extraction puis de classification vectorielle d'un ensemble de caractéristiques spectrales et spatiales. Les caractéristiques spectrales sont extraites au moyen de méthodes visant à réduire la dimension des images hyperspectrales tout en conservant une majorité de l'information utile. Les caractéristiques spatiales sont quant à elles produites par l'intermédiaire d'outils de caractérisation de la texture (matrices de co-occurrence et spectres de texture) ou de la forme (profils morphologiques). Nous nous sommes ensuite intéressés à la modélisation markovienne et avons entrepris d'adapter un algorithme de classification de type Conditional Random Field à un contexte hyperspectral. Notre troisième et dernière approche s'appuie sur une segmentation préalable de l'image afin de réaliser une classification par zones et non plus par pixels.

L'information spectrale pure permet de regrouper efficacement des pixels présentant des signatures spectrales similaires et suffit généralement dans le cadre de problèmes de classification ne faisant intervenir que des classes sémantiquement très précises, liées à un unique type de matériau. Les classes plus générales (utilisées par exemple pour des applications d'aménagement des sols) se composent en revanche de plusieurs matériaux parfois communs à plusieurs classes et agencés selon des motifs qui se répètent. Caractérisables à la fois spatialement et spectralement, ces classes sont susceptibles d'être plus complètement décrites par une utilisation simultanée de ces deux types d'information. Pour conclure cette étude, nous avons effectué une comparaison des trois méthodes d'intégration de l'information spatiale au processus de classification selon les trois critères sont la précision de classification, la complexité algorithmique et la robustesse.

Mots-clés : Classification, hyperspectral, information spatiale, texture, Champs de Markov, segmentation

Development and evaluation of new spatial-spectral classification methods of hyperspectral images

Thanks to a high number of thin and contiguous spectral bands, the hyperspectral imagery can associate to each pixel of an image a spectral signature representing the reflective behaviour of the materials composing the pixel. Most of the classification algorithms use this great amount of spectral information without noticing the contextual information between the pixels that belong to the same neighborhood. This study aims to realize new algorithms using simultaneously the spectral and spatial informations in order to classify hyperspectral images, and to study their complementarity in several contexts. For this purpose, we have developed three different classification scenarios, each one adapted to a particular type of application.

The first scenario consists in a vectorial classification process. Several spectral and spatial characteristics are extracted and merged in order to form a unique data set, which is classified using a Support Vector Machine method or a Gaussian Mixture Model algorithm. The spectral characteristics are extracted using dimension reduction method, such as PCA or MNF, while the spatial characteristics are extracted using textural characterization tools (co-occurrence matrices and texture spectra) or morphological tools (morphological profiles). For the second scenario, we adapted a Conditional Random Field algorithm to the hyperspectral context. Finally, the last scenario is an area-wise classification algorithm relying on a textural segmentation method as a pre-processing step.

The spectral information is generally sufficient to deal with semantically simple classes, linked to a unique type of material. Complex classes (such as ground use classes) are composed of several materials which potentially belong to more than one class. Those classes can be characterized both spectrally and spatially, which means that they can be more completely described using both spectral and spatial informations. To conclude this study, we compared the three spectral/spatial classification scenarios using three criteria : classification accuracy, algorithmic complexity and strength.

Keywords: Classification, hyperspectral, spatial information, texture, Markov Random Fields, segmentation