

Classification de textures hyperspectrales fondée sur un modèle markovien et une technique de poursuite de projection

Hyperspectral texture classification using a markov model and a projection pursuit technique

par Guillaume RELIER, Xavier DESCOMBES, Frédéric FALZON, Josiane ZERUBIA

résumé et mots clés

Nous considérons le problème de l'analyse de textures d'images hyperspectrales. De manière à tirer parti de la finesse de la discrétisation spectrale inhérente à ce type d'image, nous réalisons une modélisation de texture qui intègre simultanément les données spatiales et spectrales en employant une approche vectorielle dans un modèle markovien. Ce modèle est utilisé conjointement avec un algorithme de poursuite de projection, permettant de déterminer un sous-espace optimal dans lequel on projette les données. Ceci permet de diminuer les effets néfastes d'une trop grande dimensionnalité des données, connus sous le nom de *phénomène de Hughes*. Ce modèle est testé dans le cadre de la classification supervisée de zones urbaines.

Image hyperspectrale, classification, texture, réduction de la dimension, poursuite de projection, champs de Markov.

abstract and key words

In this paper we tackle the problem of hyperspectral texture analysis. In order to take advantage of the images fine spectral sampling, we perform a joint spatial and spectral texture modeling using a vectorial approach within a Markovian framework. This model is used jointly with a projection pursuit algorithm, for the determination of an optimal subspace on which the data are projected. This is done to limit the effect of the high dimensionality of the data (known as *Hughes phenomenon*, or *curse of dimensionality*). This model is tested for urban areas supervised classification.

Hyperspectral image, classification, texture, dimension reduction, projection pursuit, Markov random fields.

1. introduction

Une image de télédétection est une image échantillonnée spatialement, prise par un capteur survolant une scène. Les images monospectrales, ou de niveaux de gris, fournissent, pour chaque échantillon spatial (pixel), une valeur scalaire correspondant au rayonnement électromagnétique incident sur le capteur intégré sur la bande de longueur d'onde mesurée. Les images hyperspectrales, par opposition aux images monospectrales, sont des images ayant la particularité de fournir à la fois un échantillonnage spatial et un échantillonnage spectral. Pour chaque échantillon spatial ou pixel $s = (s_1, s_2)$ de la grille spatiale \mathcal{S} de dimension $N_h \times N_v$, elles ne fournissent plus une seule valeur scalaire, mais un ensemble de valeurs qui peut être considéré comme un vecteur $x_s = (x_s^0, \dots, x_s^{N_b-1})^t$ de dimension N_b , la dimension spectrale. Chacune des valeurs x_s^k correspond à une mesure du rayonnement faite sur une des bandes de longueurs d'onde. Ces images constituent par conséquent un « pavé » dans lequel chaque point $x_{(s_1, s_2)}^k$ est défini par trois coordonnées : ses deux coordonnées spatiales s_1 et s_2 , et sa coordonnée spectrale k .

En général, la dimension N_b est supérieure à 10. Ce découpage spectral très fin autorise des mesures radiométriques précises et facilite la caractérisation du matériau présent sur le terrain. Ceci autorise, en particulier, des applications liées à la reconnaissance de sols en géologie, à l'aide à l'agriculture de précision, etc... À l'inverse, ces mesures très précises spectralement sont moins précises spatialement. Par exemple, pour le capteur aéroporté AVIRIS, on dispose pour chaque pixel s de 224 mesures spectrales x_s^k correspondant chacune à une fine bande de longueur d'onde allant du visible à l'infrarouge, et la résolution au sol est de l'ordre de 20 mètres.

Dans cet article, on s'intéresse à une analyse des textures présentes dans les images hyperspectrales en tirant parti à la fois de relations spatiales et spectrales, visant ainsi à ajouter l'information spectrale aux méthodes classiques d'analyse de textures scalaires qui ne prennent en compte que les variations spatiales des niveaux de gris. Cette analyse a pour but de caractériser les textures présentes dans les images hyperspectrales pour effectuer une classification des pixels dans les différentes classes de textures. Les méthodes d'analyse de textures couramment utilisées en traitement d'image, comme les matrices de cooccurrence [Har79] ou les filtres de Gabor [JF91], ne peuvent être adaptées simplement au cas des images multicomposantes (multispectrales ou hyperspectrales) qu'au prix d'une complexité algorithmique qui peut s'avérer rédhibitoire si l'on veut tirer parti des relations spectrales. Une autre adaptation possible consiste à faire une analyse scalaire bande par bande, les unes indépendamment des autres, suivie par une étape nécessaire de fusion de données en classification. Pour mettre en œuvre l'analyse de textures spatiale et spectrale, une méthode vectorielle semble mieux adaptée

que la précédente, car les méthodes de ce type prennent en compte les relations inter-bandes, qui, on le verra plus tard, sont porteuses d'information. Des approches de ce type ont déjà été considérées, en détection d'anomalies [SM00], par exemple.

L'application dans laquelle s'inscrit ce travail est la classification des images de zones urbaines en différents types d'occupation du sol. Ces zones sont constituées de nombreux matériaux, ce qui est à l'origine de deux phénomènes rendant en général insuffisante l'étude des réponses radiométriques $\{x_s\}$ (ou réflectances) au niveau des pixels indépendamment les uns des autres. Le premier phénomène est la grande variabilité des classes car les zones au sol correspondant à différents pixels d'une même classe sont constituées de matériaux différents. Le deuxième est que les pixels eux-mêmes peuvent être un mélange de plusieurs matériaux. De plus, une partie de ces matériaux sont aussi caractéristiques de certains paysages naturels, donnant à certains pixels urbains les mêmes signatures que ces paysages naturels. Il est donc nécessaire, pour étudier les zones urbaines, d'ajouter à la radiométrie une information de texture.

Lorsque l'on considère l'analyse de données de grande dimension par des méthodes statistiques, on est souvent confronté aux problèmes liés à la « malédiction de la dimension » de Bellman [Bel61], aussi appelé phénomène de Hughes [Hug68], se caractérisant par la difficulté de réaliser de bonnes estimations dans des espaces qui se trouvent être pratiquement « vides ». C'est pourquoi on effectue en général une projection de l'espace initial de façon à obtenir un espace de dimension plus réduite, facilitant ainsi l'estimation. Dans le paragraphe 2, on présente plus en détails ce problème, ainsi qu'un algorithme de poursuite de projection, méthode de recherche de projection moins sensible au phénomène de Hughes que certaines méthodes classiques, et peu restrictive quant au choix du critère utilisé pour évaluer la qualité de la projection. La poursuite de projection représente un cadre général de travail pour la recherche de sous-espaces qui a été adapté à la classification dans un cadre paramétrique par Jimenez et Landgrebe sous l'appellation de poursuite de projection paramétrique [JL99]. Cette mise en œuvre utilise un critère de discrimination entre les différentes classes parmi lesquelles se répartissent les pixels pour déterminer le sous-espace de projection. On propose une variante de cet algorithme à la fin de ce paragraphe. Au paragraphe 3, on propose une modélisation des images hyperspectrales prenant en compte la texture, fondée sur les champs de Markov, ainsi qu'une extension de la poursuite de projection paramétrique adaptée à cette modélisation. Enfin, dans le dernier paragraphe, on présente quelques résultats de classification de textures obtenus par le couplage de la poursuite de projection et de la modélisation par champs de Markov. Ces résultats sont comparés à ceux que l'on obtient avec une approche par modélisation gaussienne non markovienne des classes, ainsi que ceux issus d'une approche utilisant une méthode classique de réduction de la dimension de l'espace, l'analyse en composantes principales.

2. réduction de la dimension

2.1. pourquoi réduire la dimension des données ?

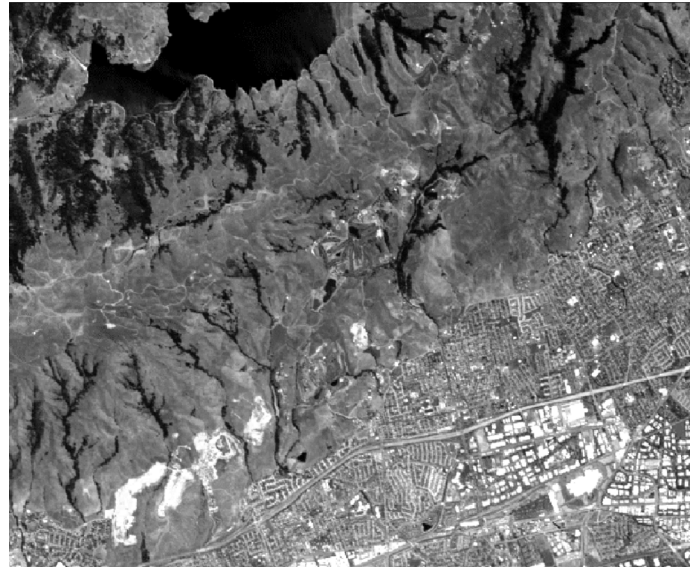
Lors d'un traitement statistique paramétrique, on passe en général par une ou plusieurs étapes d'estimation de paramètres. Notre application de l'analyse de textures consiste à effectuer une classification des pixels d'une image qui intègre la notion de texture. Dans ce cas, il est nécessaire d'estimer les paramètres des lois de probabilité des différentes classes. Or, pour un nombre fixe d'échantillons d'apprentissage, la qualité de l'estimation de certains paramètres statistiques, en particulier les paramètres matriciels, diminue quand la dimension des données augmente, car le nombre de paramètres inconnus croît plus vite que le volume de données. C'est ce qu'on appelle la « malédiction de la dimension » [Bel61]. Dans notre cas, on va chercher à déterminer une loi de probabilité sur les vecteurs aléatoires X_s , dont les valeurs x_s sont des réalisations. Or, ces vecteurs sont de dimension N_b , égale à 224 dans le cas de données AVIRIS, ce qui est beaucoup si l'on considère des ensembles de quelques centaines d'échantillons d'apprentissage.

D'autre part, plusieurs aspects caractéristiques des espaces de grande dimension en général, et des images hyperspectrales en particulier, font que cette réduction de la dimension peut être obtenue en minimisant la perte d'information. Certains de ces aspects sont recensés dans [JL98], on en rappelle les principaux ci-dessous.

Une caractéristique des espaces de grande dimension est que les projections linéaires sur des sous-espaces ont tendance à rendre les distributions gaussiennes [HL93]. Ceci permet d'utiliser des algorithmes paramétriques fondés sur une hypothèse de normalité pour réaliser la classification. En particulier, si l'on utilise une classification par maximum de vraisemblance, les équations à résoudre sont très simples dans le cas gaussien.

Un autre aspect concerne la redondance entre bandes. La figure 1 (b) représente l'image de la matrice de covariance inter-bande de la scène AVIRIS de la figure 1 (a). Le niveau de gris de chaque pixel (i, j) de la figure 1 (b) représente la covariance $c_{ij} = E\{x_s^i x_s^j\}$. La matrice montre de fortes corrélations entre bandes adjacentes, par blocs. Les matrices de covariance des différentes classes d'occupation du sol présentent également la même structure très corrélée, par blocs. Cette structure par blocs vient du fait que les propriétés de réflexion de l'onde électromagnétique des matériaux sont communes par groupes de longueurs d'ondes proches, et changent radicalement pour certains paliers, comme par exemple les longueurs d'ondes d'absorption du rayonnement par certaines molécules de l'atmosphère (eau, oxygène...). D'autre part les valeurs propres de ces matrices présentent de grandes disparités. Les réalisations des vecteurs aléatoires pouvant être représentées par des points distribués dans un

espace euclidien de dimension N_b , cette dernière propriété traduit une forte anisotropie des nuages formés par ces points, qui sont étirés dans certaines directions privilégiées, alors que la dispersion est faible dans d'autres directions. Ces propriétés traduisent l'existence d'une grande redondance entre les différentes bandes, rendant inutile l'utilisation des mesures vectorielles x_s dans toute leur dimension.



(a)



(b)

Figure 1. – a) 3 bandes parmi celles du spectre visible d'une scène AVIRIS, b) image de la matrice de covariance inter-bande de la scène a) : l'intensité en un point (i, j) de l'image représente la covariance $c_{ij} = E\{x_s^i x_s^j\}$.

2.2. méthodes de réduction de la dimension

Les propriétés précédentes suggèrent qu'il peut être intéressant d'utiliser pour le traitement des images hyperspectrales, non pas les mesures initiales de grande dimension, mais une projection de celles-ci dans un espace de plus petite dimension. Le choix de ce sous-espace n'est, bien sûr, pas indifférent, et certains sous-espaces ont de bien meilleures caractéristiques que d'autres.

On peut distinguer, en matière de réduction de la dimension des espaces, les approches générales des approches spécifiquement destinées à la classification. Les premières ont été développées dans un contexte non supervisé, alors que les secondes l'ont été dans un contexte supervisé.

2.2.1. approches générales

Dans ces méthodes, on considère un critère de variance calculé sur l'image dans son ensemble, et non un critère relatif à la séparabilité des classes. Contrairement aux méthodes que l'on verra plus tard, ces méthodes ne sont pas optimales par rapport à des critères de classification, mais elles sont simples et donnent tout de même de bons résultats.

Parmi les méthodes existantes, la méthode la plus simple, mais qui reste efficace, est l'analyse en composantes principales (ACP) [Dun89, JR99]. Le principe de cette méthode est de déterminer les « composantes principales » (CP), qui sont les vecteurs propres de la matrice de covariance des données définie comme $\{c_{ij}\}$, avec $c_{ij} = E\{x_s^i x_s^j\}$, et de projeter les données x_s dans l'espace de dimension n généré par les n premières CP correspondant aux n plus grandes valeurs propres. Les composantes principales sont un ensemble de vecteurs ayant la particularité de former une base orthogonale de l'espace de départ, et d'être toutes statistiquement décorréliées les unes des autres. Cette méthode assure des résultats minimaux en termes d'erreur quadratique, mais elle a deux inconvénients majeurs : son temps de calcul est élevé et le critère de variance peut ne pas être adapté au traitement envisagé qui, dans notre cas, est la classification des zones urbaines.

La transformation appelée *Minimum Noise Fraction* (MNF) est une variante de l'ACP, ajustée au bruit [GBSC88]. Cette transformation consiste à effectuer une rotation de la base afin d'obtenir une nouvelle base orthogonale où les vecteurs sont ordonnés suivant le rapport signal sur bruit (S/B) existant dans leur direction, et non suivant la variance comme pour l'ACP. Le problème de cette méthode est qu'il faut avoir une connaissance *a priori* sur le bruit, ce qui n'est pas forcément possible si l'on ne connaît pas les caractéristiques du capteur ou les conditions d'acquisition des données. Il est intéressant de noter que, dans le cas où l'on considère que le bruit est décorrélié et de même variance dans chaque bande, la transformation par MNF est équivalente à une ACP.

2.2.2. analyse discriminante

Pour la classification, d'autres méthodes peuvent être envisagées, en ne considérant plus un critère d'erreur quadratique comme avec l'ACP, mais en prenant en compte la séparabilité des classes.

On se place, pour ces approches, dans le cadre de la classification supervisée car ces méthodes nécessitent une définition préalable des statistiques des classes. Or, du fait de l'effet de Hughes, le calcul des statistiques devient moins fiable quand la dimension des observations x_s augmente.

Deux méthodes nous ont paru particulièrement intéressantes : l'*analyse discriminante* et la *poursuite de projection*, qui sera présentée au prochain paragraphe.

L'*analyse discriminante* [Fuk90] consiste à maximiser un critère lié à la dispersion inter-classe et à la dispersion intra-classe. La dispersion inter-classe est la dispersion des centres des classes que l'on doit maximiser, alors que la dispersion intra-classe est la dispersion d'une même classe autour de son centre, qui doit être minimisée.

On définit donc deux matrices, la matrice de dispersion intra-classe :

$$S_w = \sum_i P(\omega_i) S_i \quad (1)$$

et la matrice de dispersion inter-classe :

$$S_b = \sum_i P(\omega_i) (M_i - M_0)(M_i - M_0)^t \quad (2)$$

où $M_0 = \sum_i P(\omega_i) M_i$. M_i , S_i et $P(\omega_i)$ sont respectivement la moyenne, la matrice de covariance et la probabilité *a priori* de la classe ω_i .

On définit $S_w^d = d^t S_w d$ et $S_b^d = d^t S_b d$, comme étant les transformations des matrices S_w et S_b par une matrice de projection orthogonale d .

On peut construire différents critères à partir de ces matrices, comme ceux recensés dans [Fuk90]. En pratique, on utilise le plus souvent le critère de la trace :

$$J = \text{tr}((S_w^d)^{-1} S_b^d) \quad (3)$$

qu'il faut maximiser par rapport à d .

L'optimisation de ce critère conduit à un espace dont la dimension est limitée par le nombre de classes, car $S_w^{-1} S_b$ est de rang égal au nombre de classes moins 1. Par exemple, si l'on a deux classes, la méthode fournit un seul vecteur. D'autre part, si les moyennes des classes ne sont pas assez uniformément réparties dans l'espace initial, notamment si deux classes ont des moyennes très proches (ce qui conduit à une très faible contribution à la matrice de dispersion inter-classe), les résultats de classification obtenus en utilisant la base déterminée sont peu satisfaisants. Cependant, la méthode est intéressante pour sa

rapidité. Mais elle dépend fortement des statistiques calculées dans l'espace initial, c'est-à-dire qu'elle peut être sujette à l'effet de Hughes.

2.2.3. poursuite de projection

La poursuite de projection [Hub85] est une méthode visant à rechercher un sous-espace dans lequel un indice de projection est maximisé.

Cette méthode est intéressante pour deux raisons :

- le phénomène de Hughes est atténué du fait que tous les calculs sont faits dans le sous-espace de projection,
- l'indice de projection peut être adapté au traitement envisagé.

Le choix de l'indice de projection est bien entendu le point crucial de l'algorithme, puisqu'un bon choix dépend presque intégralement de l'application. Par exemple, un choix d'indice pour la compression pourrait être la variance totale, ce qui ramène à faire une analyse en composantes principales.

Initialement, la poursuite de projection était utilisée dans des contextes non paramétriques, afin de trouver des angles de projection pour lesquels les données seraient « bien » représentées [Hub85]. Les indices de projection étaient alors des mesures liées à l'histogramme de la projection du nuage des points x_s .

Plus récemment, cette méthode a été utilisée dans un contexte paramétrique dans le cadre de la classification d'images hyperspectrales [JL99]. Ayant choisi un modèle Gaussien pour la distribution des vecteurs aléatoires associés à l'ensemble des pixels, les auteurs proposent d'appliquer une approximation de la poursuite de projection afin de trouver le sous-espace maximisant le minimum de la distance de Bhattacharyya entre paires de classes définies au préalable à travers des échantillons d'apprentissage, car celle-ci est reliée à une borne supérieure de la probabilité d'erreur [Fuk90,LC00], cf. le terme de gauche dans l'équation (4). Cette distance, que l'on note d_B , se définit comme suit, dans le cas de deux classes équiprobables, hypothèse faite en l'absence de connaissance *a priori* :

$$\int \sqrt{P_0(x)P_1(x)}dx = e^{-d_B} \quad (4)$$

Le paragraphe 4 est consacré au choix d'un indice de projection adapté au problème d'analyse de textures qui nous occupe.

Une mise en œuvre approchée de la poursuite de projection, dite séquentielle, est décrite dans [JL99] (cf. algorithme 1). Cet algorithme est appliqué à la réduction de la dimension de l'espace où les pixels x_s prennent leur valeur. La luminance d'un pixel est considérée comme un vecteur qui prend sa valeur dans l'espace généré par les N_b bandes spectrales. Cet espace est divisé en P sous-espaces, P étant la dimension de l'espace final de projection, déterminée arbitrairement. Chacun de ces sous-espaces est généré par un groupe de bandes adjacentes, c'est-à-dire correspondant à des bandes de longueurs d'onde adjacentes. Un vec-

teur de base optimal du sous-espace de projection, estimé itérativement, est extrait de chacun des P sous-espaces, par combinaison linéaire des vecteurs de base du sous-espace. Les bandes adjacentes présentant en général une grande corrélation, il semble logique de chercher à extraire des combinaisons de ces bandes. Cette méthode de recherche donne une solution sous-optimale, puisque l'ensemble des espaces que l'on peut obtenir est très fortement réduit par le découpage en groupes de bandes. Néanmoins, la perte en terme d'indice de projection de l'espace initial est raisonnable, et le gain en temps de calcul est très grand. Dans la mise en œuvre originale, les P groupes de bandes sont déterminés en divisant le spectre à N_b bandes en groupes de même nombre de bandes. L'information n'étant pas forcément équirépartie entre les bandes, nous proposons une modification de cet algorithme, dans laquelle les groupes de bandes adjacentes n'ont pas forcément le même nombre de bandes et sont choisis conformément à un critère que l'on décrit ci-dessous.

Algorithme 1 Poursuite de projection séquentielle

L'ensemble des bandes est divisé en P groupes de bandes adjacentes $\{g_p = \{b_{p_0}, \dots, b_{p_{n_p}}\}\}_{p=1, \dots, P}$, où b_i désigne la bande numéro i .

Chaque groupe de bandes g_p génère un espace vectoriel G_p .

$n \leftarrow 0$, *arret* $\leftarrow 0$

pour $p = 1$ à P **faire**

Initialiser arbitrairement le vecteur de base $v_p = v_p^0 \in G_p$

fin pour

$n \leftarrow n + 1$

tant que *arret* $\neq 1$ **faire**

pour $p = 1$ à P **faire**

Soit $v(x) = (v_1^n, \dots, v_{p-1}^n, x, v_{p+1}^{n-1}, \dots, v_P^{n-1})$, avec

$x \in G_p$.

$v_p^n \leftarrow \arg \max_x I(x)$, $I(x)$ étant un indice de projection à définir

$v^n = (v_1^n, \dots, v_P^n)$

fin pour

si $\max_p \|v_p^n - v_p^{n-1}\| < \epsilon$ **alors**

arret $\leftarrow 1$

fin si

fin tant que

La base de l'espace de projection final est $\hat{v} = v^n$, et les vecteurs qui la constituent sont les $\hat{v}_p = v_p^n$.

Poursuite de projection séquentielle avec groupes de bandes variables

Nous proposons une variante de la poursuite de projection séquentielle, pour laquelle on cherche un découpage du spectre

en groupes de bandes plus adapté que le découpage à pas constant.

Étant donné que chaque groupe est utilisé pour extraire un vecteur de base, un critère pour le découpage peut être l'homogénéité de la répartition de l'indice de projection dans les différents groupes. Avoir des groupes de bandes homogènes permet de ne pas extraire un vecteur de base d'un groupe de bandes trop peu informatif par rapport aux autres et, au contraire, d'éviter que seulement un vecteur de base soit extrait pour représenter des groupes où l'indice de projection est très important.

Soit $z = \{z_0, \dots, z_P\}$ les bornes définissant un découpage du spectre en P sous-ensembles de bandes, avec $z_0 = 0$ et $z_P = N_b$, N_b étant le nombre total de bandes, par exemple $N_b = 224$ pour les images AVIRIS. Un groupe de bande p , duquel on va extraire le vecteur \hat{v}_p de la base de l'espace final, sera donc constitué des bandes du spectre comprises entre les bornes z_p et z_{p+1} (z_p inclus).

On cherche à trouver le découpage \hat{z} tel que :

$$\hat{z} = \arg \min_z \sigma_I(z) \quad (5)$$

avec

$$\sigma_I(z) = \frac{1}{P} \sum_{p=0}^{P-1} (I(z_p, z_{p+1}) - m_I(z))^2 \quad (6)$$

où $I(z_p, z_{p+1})$ est la valeur de l'indice correspondant à une projection sur l'espace généré par le groupe des bandes comprises entre z_p et z_{p+1} (z_p inclus) et où la moyenne $m_I(z)$ est définie par :

$$m_I(z) = \frac{1}{P} \sum_{p=0}^{P-1} I(z_p, z_{p+1}) \quad (7)$$

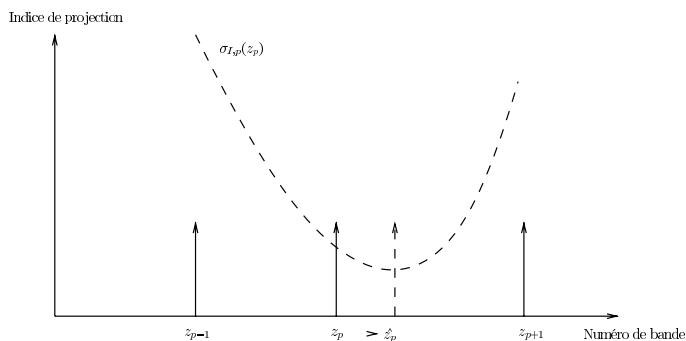


Figure 2. – Illustration de l'algorithme d'ajustement des bornes : pour z_{p-1} et z_{p+1} fixés, la valeur optimale \hat{z}_p de z_p est celle minimisant la fonction $\sigma_{I,p}(z_p)$

Or, le terme de droite de l'équation (6) se simplifie, et l'on a :

$$\sigma_I(z) = \frac{1}{P} \sum_{p=0}^{P-1} I^2(z_p, z_{p+1}) - m_I^2(z) \quad (8)$$

Pour trouver une estimation de \hat{z} qui minimise σ_I , on examine les bornes z_p , $p = 1, \dots, P - 1$ itérativement. Quand toutes les bornes sont fixes, excepté la borne de rang p , σ_I peut être considéré comme une fonction de la position de cette borne, notée x , qui reste comprise entre z_{p-1} et z_{p+1} . On note cette fonction $\sigma_{I,p}(x)$ (cf. 2). On détermine la nouvelle valeur de z_p , qui minimise la fonction convexe $\sigma_{I,p}(x)$, en faisant varier la position x jusqu'à atteindre un minimum de $\sigma_{I,p}(x)$, puis l'on passe à la borne suivante. L'algorithme est itéré jusqu'à la convergence, obtenue quand aucune borne ne bouge durant une itération.

Cet algorithme ne garantit pas de trouver le minimum global de la variance, mais il est plus rapide que les algorithmes d'optimisation globale et aboutit à de bons résultats si l'initialisation est bonne. On choisit d'initialiser les bornes z_p en les séparant par

un pas constant $k = \frac{N_b}{P}$, c'est à dire que l'on considère le découpage utilisé dans [JL99].

La figure 3 montre les évolutions comparées de l'indice optimal (le choix de l'indice à maximiser est explicité au paragraphe 4) en fonction du nombre de bandes conservées pour des données AVIRIS, pour les mises en œuvres à pas constant et à pas variable de la poursuite de projection séquentielle. Il apparaît que l'ajustement du découpage permet d'obtenir un indice de projection final supérieur pour des dimensions inférieures à 20. En revanche, lorsque l'on considère des dimensions finales supérieures à 20, l'ajustement n'est pas nécessaire, puisque dans les deux cas, les indices optimaux sont assez proches.

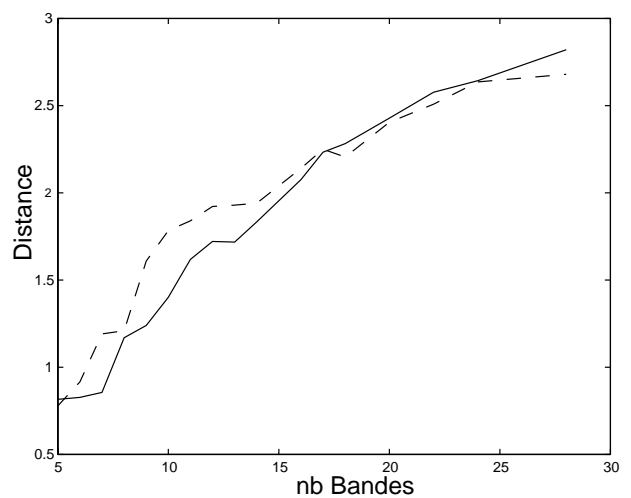


Figure 3. – Indice optimal estimé en fonction du nombre de bandes pour les deux mises en œuvre de la poursuite de projection séquentielle, dans le cas de deux textures d'une même scène AVIRIS. tirets : mise en œuvre avec groupes variables, trait plein : mise en œuvre avec groupes fixes.

On discute au paragraphe 4 du problème de la détermination d'un bon indice de projection.

Le paragraphe suivant présente le modèle de texture dont on fait l'hypothèse, qui servira à définir l'indice de projection utilisé dans notre application.

3. modélisation

Les approches en matière d'analyse de textures sont nombreuses et il serait difficile d'en faire une revue exhaustive. Parmi les plus populaires, et sans doute les plus efficaces en analyse de textures d'images à niveaux de gris, on trouve les approches utilisant les matrices de cooccurrence pour extraire des paramètres caractéristiques de la texture observée [Har79]. Cette approche donne de bons résultats pour les images monospectrales, mais son adaptation aux images hyperspectrales paraît irréalisable. En effet, une matrice de cooccurrence est une matrice de dimension $N \times N$, où N est le nombre de niveaux de gris quantifiés de l'image. Pour adapter cette méthode aux images hyperspectrales, il faudrait soit travailler bande par bande de façon scalaire, auquel cas l'information inter-bande serait perdue, soit travailler sur le pavé image. Mais ce dernier cas est pratiquement impossible, car il impliquerait de travailler sur des matrices gigantesques $N_q N_b \times N_q N_b$, où N_q est le nombre de niveaux de luminance, et N_b le nombre de bandes. Or les images AVIRIS ont 224 bandes et sont codées sur 16 bits !

L'approche par filtrage de Gabor [JF91] est plus rapide, mais l'extension aux données hyperspectrales sans perte d'information n'est pas évidente, pour les mêmes raisons que précédemment concernant la prise en compte de l'information inter-bande. La modélisation par champs de Markov, que l'on présente dans les paragraphes suivants, est intéressante du fait qu'elle concerne des champs de vecteurs aléatoires. Le fait d'avoir une méthode d'analyse fondée sur une modélisation vectorielle permet une analyse globale de l'ensemble des bandes conjointement.

3.1. modèles markoviens

Les champs de Markov sont des outils très utilisés en traitement d'image. Leur intérêt est principalement de pouvoir modéliser des interactions entre pixels de façon simple, à travers un modèle probabiliste. Un champ de Markov X est un champ aléatoire présentant la propriété suivante :

$$P(X_s = x_s | X^s = x^s) = P(X_s = x_s | \{x_t\}, t \in V_s) \quad (9)$$

où X_s désigne la variable aléatoire au site s , X^s le champ privé de cette variable aléatoire, et V_s le voisinage de s .

Ceci signifie que la probabilité que la variable aléatoire X_s en un site s prenne la valeur x_s conditionnellement au reste du champ est égale à la probabilité en ce site connaissant uniquement ses voisins. Cette modélisation permet d'introduire explicitement des interactions à l'échelle du voisinage et induit implicitement des interactions à plus longue échelle par « transitivité ».

C'est ainsi que ces champs sont efficacement utilisés en segmentation [KH95], classification [DMZ97] ou en restauration [GG84], pour régulariser les images segmentées ou de niveaux de gris.

Mais les champs de Markov peuvent aussi être utilisés pour la modélisation des images, en particulier, des textures fines [DES93, Win95, Des97, Gim99]. Il existe de nombreux modèles markoviens de texture, comme par exemple le modèle auto-binominal [Bes74], le modèle de Derin-Elliott [DE87] ou le modèle markovien gaussien aussi appelé auto-normal [CCB85]. Ce dernier est particulièrement bien adapté au problème d'analyse de textures de zones urbaines.

3.2. modèle markovien gaussien

Le modèle markovien gaussien a beaucoup été utilisé sur les images de niveaux de gris pour l'analyse de textures, notamment urbaines [DSP99, LDZ00]. Mais il est aussi possible d'utiliser cette modélisation dans le cadre de l'imagerie multispectrale. Deux approches peuvent être distinguées dans ce cas. Dans la première [SM00, PH95], les données couleur ou hyperspectrales sont considérées comme un ensemble de sites dans un modèle à trois dimensions : deux dimensions spatiales et une dimension spectrale. Chacun des sites est indexé par deux coordonnées spatiales, correspondant à sa place dans la grille spatiale, et une coordonnée spectrale, c'est-à-dire le numéro de la bande correspondante. En chaque site $s = (i, j, k)$, on a une mesure scalaire, la luminance au pixel (i, j) mesurée dans la $k^{\text{ème}}$ bande de longueur d'onde. La deuxième approche, utilisée notamment dans [Haz00], consiste à définir un modèle à deux dimensions. Chaque site $s = (i, j)$ est défini par ses deux coordonnées spatiales, et les X_s sont des vecteurs aléatoires de dimension égale au nombre de bandes N_b .

Les méthodes présentées dans cet article sont reliées à une modélisation de ce dernier type, que l'on appelle MGMRF (*Multivariate Gaussian Markov Random Field*).

3.3. le MGMRF

3.3.1. loi conditionnelle

Soit une image x de dimensions spatiales $N_h \times N_v$ et ayant un nombre de bandes N_b . Cette image est supposée être une réalisation du champ X . Comme on l'a vu au paragraphe précédent,

on suppose que le champ considéré, censé modéliser une texture homogène, est un champ à 2 dimensions (les dimensions spatiales) composé de variables vectorielles de dimension égale au nombre de bandes (les vecteurs de luminance). La modélisation markovienne implique que la probabilité qu'un vecteur X_s prenne la valeur x_s en un site s connaissant le reste de l'image est égale à la probabilité en ce site connaissant ses voisins. La modélisation markovienne gaussienne ajoute une contrainte sur cette probabilité, qui est supposée être de la forme :

$$P(x_s|\{x_t\}, t \in V_s) = \frac{1}{Z} \exp \left\{ -\frac{1}{2} \|x_s - \sum_{t \in V_s} \theta_{t-s} x_t\|_{\Sigma}^2 \right\} \quad (10)$$

avec $\|a\|_{\Sigma}^2 = a^t \Sigma^{-1} a$, les θ_{t-s} étant des matrices d'interaction, Σ la matrice de covariance conditionnelle, Z une constante de normalisation appelée fonction de partition et V_s le voisinage spatial du pixel s .

Cette loi est donc une loi gaussienne dont la moyenne dépend du voisinage, mais pas la covariance.

Cette expression est valable si l'on considère des données centrées, c'est-à-dire de moyenne nulle. Dans le cas de moyennes non nulles, l'expression est légèrement différente :

$$P(x_s|\{x_t\}, t \in V_s) \quad (11)$$

$$= \frac{1}{Z} \exp \left\{ -\frac{1}{2} \|x_s - \sum_{t \in V_s} \theta_{t-s} x_t - (I_{N_b} - \sum_{t \in V_s} \theta_{t-s}) \mu\|_{\Sigma}^2 \right\}$$

où I_{N_b} est la matrice identité de dimension N_b égale à celle du vecteur x_s .

Par la suite, on considérera, pour faciliter l'écriture, que la texture homogène considérée a été centrée, la généralisation ne présentant pas de difficulté.

La formulation (10) permet d'écrire :

$$x_s = \sum_{t \in V_s} \theta_{t-s} x_t + u_s \quad (12)$$

où u_s est un bruit blanc gaussien vectoriel, centré, de matrice de covariance Σ .

De plus, on suppose que le champ est stationnaire. On doit donc avoir $\theta_{t-s} = \theta_{s-t}$, ce qui conduit à :

$$x_s = \sum_{k=0}^K n_k \theta_k m_s^k + u_s \quad (13)$$

où k désigne chaque type de voisin (horizontal, vertical, ...), K étant le nombre de types de voisins, m_s^k le vecteur-moyenne des n_k voisins de type k du pixel s et n_k le nombre de voisins de type k dans le système de voisinage considéré. Un *type* de voi-

sin désigne l'ensemble des voisins t d'un pixel s ayant la même contribution dans le modèle, ce qui se traduit par la même matrice θ_{s-t} . À titre d'exemple, dans un système de voisinage en 4-connexité isotrope, les 4 voisins sont considérés comme étant du même type, tandis que dans un système voisinage en 8-connexité anisotrope, on peut considérer quatre types de voisins : les voisins horizontaux, verticaux, et les deux types de voisins correspondant aux deux diagonales.

3.3.2. loi globale

Pour obtenir la formule de la loi globale du champ X , on reformule l'équation (12) :

$$u_s = x_s - \sum_{t \in V_s} \theta_{t-s} x_t \quad (14)$$

Soit x l'image entière réorganisée en un vecteur colonne, $x = (x_{(0,0)}^t \dots x_{(N_v, N_h)}^t)^t$, où $x_{(0,0)}^t$ est le vecteur qui contient les niveaux du pixel de coordonnées spatiales $(0, 0)$. De même $u = (u_{(0,0)}^t \dots u_{(N_v, N_h)}^t)^t$. L'équation précédente se réécrit comme suit :

$$u = Ax \quad (15)$$

où la matrice A est une réorganisation des θ_{t-s} .

Le vecteur u suivant une loi normale, le vecteur x suit une loi du même type, ce qui permet d'écrire :

$$P(x) = \frac{1}{Z(S)} \exp \left(-\frac{1}{2} x^t S^{-1} x \right) \quad (16)$$

avec $S = E\{xx^t\} = A^{-1} E\{ux^t\}$ et où $Z(S)$ est une constante de normalisation.

Or, d'après le principe d'orthogonalité, on doit avoir :

$$E\{u_i x_j^t\} = \delta_{i,j} \Sigma, \quad \forall i, j \quad (17)$$

d'où :

$$E\{ux^t\} = I_{N_h} \otimes I_{N_v} \otimes \Sigma \quad (18)$$

où \otimes désigne le produit matriciel de Kronecker¹ [Jai89] et I_n la matrice identité de dimension n . L'équation (16) peut alors être

1. Le produit de Kronecker entre deux matrices $A = \{a_{ij}\}$ de dimensions (n_1, n_2) et $B = \{b_{ij}\}$ de dimensions (m_1, m_2) est une matrice $A \otimes B = C = \{c_{ij}\}$ de dimensions $(m_1 n_1, m_2 n_2)$ telle que $c_{im+k, jm+l} = a_{i,j} b_{k,l}$.

réécrite :

$$P(x) = \frac{1}{Z(S)} \exp \left(-\frac{1}{2} x^t (I_{N_h} \otimes I_{N_v} \otimes \Sigma^{-1}) Ax \right) \quad (19)$$

3.3.3. caractéristiques des espaces de grande dimension

Après avoir présenté le modèle markovien gaussien multivariable MGMRF dans sa forme générale, on s'intéresse plus particulièrement à son application dans le cadre de la classification des pixels d'images hyperspectrales AVIRIS.

Le processus de classification des pixels tel qu'on l'envisage est constitué dans un premier temps d'une phase durant laquelle les échantillons d'apprentissage des classes d'occupation du sol à distinguer sont définis par un opérateur, et où celles-ci sont caractérisées par les valeurs estimées de leurs paramètres.

Le nombre de paramètres du modèle markovien gaussien est variable, selon l'ordre du voisinage, son anisotropie, ou les contraintes appliquées aux matrices θ_{s-t} . Dans le cas, par exemple, d'un modèle anisotrope à 4 voisins, il est nécessaire de calculer les deux matrices θ_v et θ_h correspondant aux voisinages vertical et horizontal, ainsi que la matrice de covariance conditionnelle Σ , soit, si le vecteur d'observation est de dimension N_b , $3N_b^2$ paramètres. Dans le cas de données AVIRIS à 224 bandes, l'estimation d'un nombre aussi élevé de paramètres est très difficile, étant donné que le nombre d'échantillons disponibles déterminés par l'opérateur ne peut excéder quelques milliers. Or, plus le nombre de paramètres est élevé, plus le phénomène de Hughes [Hug68] est présent.

Deux solutions peuvent être apportées à ce problème. D'une part, comme mentionné au début de cet article, il est possible de choisir des critères et des méthodes permettant de réduire la dimension des observations de façon à améliorer l'estimation et la classification. D'autre part, il est possible de simplifier le modèle de manière à réduire substantiellement le nombre de paramètres à estimer.

Dans notre application, on effectue une simplification découlant des hypothèses suivantes sur les textures analysées :

- les interactions entre pixels n'existent qu'entre les valeurs de luminance aux mêmes longueurs d'ondes,
- les interactions sont identiques dans les différentes longueurs d'onde du spectre.

Dans ces conditions, les matrices d'interaction sont diagonales, telles que :

$$\theta_i = \text{diag}(a_i) = a_i I_{N_b} \quad (20)$$

et la probabilité conditionnelle prend la forme simple :

$$P(x_s | \{x_t\}, t \in V_s) = \frac{1}{Z} \exp \left\{ -\frac{1}{2} \left\| x_s - \sum_{t \in V_s} a_{t-s} x_t \right\|_{\Sigma}^2 \right\} \quad (21)$$

Les relations inter-bandes n'entrent donc pas en jeu dans les relations entre pixels, mais sont reflétées dans la matrice de covariance conditionnelle Σ .

3.3.4. domaine de validité des paramètres

Pour que le champ X ayant la loi définie par l'équation (16) existe, il faut que la matrice $S = E\{xx^t\}$ soit définie positive. Par conséquent il faut que A soit définie positive si l'on considère que Σ l'est. Cette dernière hypothèse est vraisemblable, car si l'on considère un nombre d'échantillons suffisamment élevé, l'estimée de Σ par maximum de vraisemblance est définie positive.

Par la suite, on utilisera la formulation suivante de la probabilité conditionnelle :

$$P(x_s | \{m_s^k\}) = \frac{1}{Z} \exp \left\{ -\frac{1}{2} \left\| x_s - \sum_k n_k a_k m_s^k \right\|_{\Sigma}^2 \right\} \quad (22)$$

où les $\{m_s^k\}$ sont les vecteurs-moyennes des voisins de type k du pixel s et les $\{n_k\}$ les nombres de voisins de type k .

Par un calcul similaire à celui mené dans [Haz00], on montre qu'une condition nécessaire pour l'existence du champ markovien gaussien de paramètres $\{a_k\}$ est que $\sum_{k=0}^{K-1} n_k a_k < 1$, où K désigne le nombre de types de voisins.

3.3.5. estimation des paramètres

Considérons une image hyperspectrale x de taille $N_v \times N_h$, ayant un nombre de bandes égal à N_b , de moyenne $E\{x_s\}$ nulle, représentant une texture homogène, supposée avoir été générée par le modèle décrit par l'équation (21), et dont les paramètres sont inconnus. On désire estimer ces paramètres d'une façon rapide et précise.

Un critère très utilisé pour l'estimation de paramètres est le Maximum de Vraisemblance (MV). Il consiste à retenir le vecteur de paramètres $\hat{\phi}$ (dans le cas présent $\phi = (\{a_k\}, \Sigma)$), tel que :

$$\hat{\phi} = \arg \max_{\phi} (P(x|\phi)) \quad (23)$$

En utilisant l'équation (16), qui définit une loi gaussienne, et en explicitant la constante de normalisation associée à cette loi, il vient :

$$P(x|\phi) = \frac{(2\pi)^{-\frac{N_v N_h N_b}{2}}}{|S|} \exp \left(-\frac{1}{2} x^t S^{-1} x \right) \quad (24)$$

où $|S|$ représente le déterminant de la matrice S .

La log-vraisemblance $LV(\phi)$ s'écrit alors :

$$LV(\phi) = \text{Cte} - \ln |S| - \frac{1}{2} x^t S^{-1} x \quad (25)$$

La maximisation de cette expression est possible, mais nécessite des algorithmes d'optimisation numériques coûteux comme avec la méthode du gradient conjugué, ou des approximations dans certaines variantes du modèle [SM00, BK99] pour pouvoir effectuer la minimisation algébriquement.

Un critère permettant une estimation plus rapide est celui du Maximum de Pseudo-Vraisemblance (MPV) [Bes74]. Ce critère est fondé sur une hypothèse d'indépendance des probabilités conditionnelles et mène à l'approximation suivante concernant la probabilité que le champ X prenne la valeur x :

$$P(x) \approx \prod_{s \in \mathcal{S}} P(x_s | \{x_t\}, t \in V_s) \quad (26)$$

L'estimateur de Maximum de Pseudo-Vraisemblance est :

$$\hat{\phi} = \arg \max_{\phi} \left(\prod_{s \in \Omega} P(x_s | \{x_t\}, \phi) \right) \quad (27)$$

$$\hat{\phi} = \arg \min_{\phi} \left(\sum_s \left\| x_s - \sum_k n_k a_k m_s^k \right\|_{\Sigma}^2 + N_v N_h \ln |\Sigma| \right) \quad (28)$$

Une fois de plus, la minimisation ne se ramène pas à une formule simple, notamment parce que la matrice Σ dépend des $\{a_k\}$. En considérant la matrice Σ indépendante de ces coefficients, la minimisation selon les $\{a_k\}$ est ramenée à un calcul matriciel :

$$\begin{pmatrix} \langle (m_s^0)^t \Sigma^{-1} m_s^0 \rangle & \dots & \langle (m_s^0)^t \Sigma^{-1} m_s^K \rangle \\ \vdots & & \vdots \\ \langle (m_s^K)^t \Sigma^{-1} m_s^0 \rangle & \dots & \langle (m_s^K)^t \Sigma^{-1} m_s^K \rangle \end{pmatrix} \begin{pmatrix} n_0 a_0 \\ \vdots \\ n_K a_K \end{pmatrix} = \begin{pmatrix} \langle x_s^t \Sigma^{-1} m_s^0 \rangle \\ \vdots \\ \langle x_s^t \Sigma^{-1} m_s^K \rangle \end{pmatrix} \quad (29)$$

Cette hypothèse d'indépendance entre les a_k et Σ n'est pas vérifiée et n'est qu'une commodité de calcul. Néanmoins, on s'approche de cette hypothèse si le champ est peu corrélé spatialement.

On calcule alors dans un premier temps la matrice de covariance conditionnelle comme étant la matrice de covariance du champ ou la matrice identité, ce qui ramène le problème à une simple estimation aux moindres carrés, de façon à pouvoir calculer les paramètres en deux étapes rapides : d'abord les $\{a_k\}$, ensuite Σ .

Une fois les $\{a_k\}$ calculés, le calcul de Σ est un simple calcul de covariance :

$$\Sigma = \langle u_s u_s^t \rangle = \left\langle \left(x_s - \sum_k n_k a_k m_s^k \right) \left(x_s - \sum_k n_k a_k m_s^k \right)^t \right\rangle \quad (30)$$

4. application de la modélisation proposée à la réduction de la dimension de l'espace pour la classification

Le modèle de texture ayant été présenté, on peut à présent revenir sur le rapport entre celui-ci et la définition de l'indice de projection qui sera utilisé dans la poursuite de projection.

4.1. mise en relation de la règle de classification et de l'indice de projection

Soit un problème de classification à L classes C_0 et C_{L-1} . Dans le cas de la modélisation présentée précédemment, la règle de décision correspondant au maximum de vraisemblance (MV) est d'affecter l'étiquette $\hat{\lambda}_s \in \{0 \dots L-1\}$ au pixel s si :

$$\hat{\lambda}_s = \arg \max_{l \in \{0 \dots L-1\}} P_l(x_s | \{m_s^k\}) \quad (31)$$

où $P_l(\cdot)$ est la notation allégée de $P(\cdot | \lambda_s = l)$, λ_s étant l'étiquette réelle du pixel s .

Dans la modélisation que l'on propose, les probabilités conditionnelles sont gaussiennes. Or, pour ces distributions, la distance de Bhattacharyya a une expression exacte simple. On peut donc définir une distance de Bhattacharyya en fonction des $\{m_s^k\}$. Pour simplifier l'écriture, considérons un modèle à un seul type de voisinage, par exemple un voisinage 4-connexe isotrope. On a donc :

$$x_s \sim \mathcal{N}(4a_l m_s + b_l, \Sigma_l) \quad (32)$$

si x_s appartient à la classe C_l , m_s étant la moyenne du vecteur de luminance des 4 voisins de s , et a_l le seul paramètre d'interaction de la classe l (les indices correspondant au type de voisinage sont omis, car il n'y en a qu'un).

On a le vecteur $b_l = (1 - 4a_l)\mu_l$, où $\mu_l = E_l\{x_s\}$, c'est-à-dire la moyenne de la texture représentée par la classe C_l . L'apparition de ce vecteur vient du fait que l'on ne peut plus considérer les données comme centrées, car cela reviendrait à se restreindre au cas très particulier où toutes les textures ont la même moyenne. On travaille donc dans le cadre de données non centrées, ce qui ne change en rien le processus d'estimation.

La distance de Bhattacharyya $d_B(m_s)$ entre les classes C_0 et C_1 s'écrit donc :

$$d_B(m_s) = \frac{1}{8} \|\mu_0(m_s) - \mu_1(m_s)\|_{\Sigma_{01}} + \frac{1}{2} \ln \left(\frac{|\Sigma_{01}|}{\sqrt{|\Sigma_0| |\Sigma_1|}} \right) \quad (33)$$

$$\text{avec } \mu_i(m_s) = a'_i m_s + b_i \text{ et } \Sigma_{01} = \frac{\Sigma_0 + \Sigma_1}{2}.$$

Étant donné la règle de décision (31), la probabilité d'erreur connaissant m_s est liée à $d_B(m_s)$. En maximisant $d_B(m_s)$, on minimise cette probabilité d'erreur. On peut construire deux indices de projection liés à la distance de Bhattacharyya. La mise en œuvre de la poursuite de projection consistera à rechercher la projection pour laquelle l'un ou l'autre de ces indices sera maximisé.

Indice 1

$$I_1 = E\{d_B(m_s)\} = \int d_B(m_s) f(m_s) dm_s \quad (34)$$

où nous estimons la densité de probabilité $f(m_s)$ du vecteur m_s empiriquement.

Indice 2

$$I_2 = \min_{m_s} d_B(m_s) \quad (35)$$

Le problème de l'indice I_1 est que l'on n'a accès à la loi $f(x)$ qu'à travers l'histogramme de l'image, ce qui peut augmenter le temps nécessaire à l'intégration. Par ailleurs, l'indice I_2 est lié à une borne supérieure de la probabilité d'erreur ce qui lui donne un second avantage en plus du temps de calcul. Mais, l'indice I_1 , en moyennant les contributions, limite l'effet des contributions de m_s peu représentatives, qui peuvent résulter d'une irrégularité du champ. Le choix entre les deux indices n'est donc pas trivial. Toutefois, de façon empirique, on remarque que les résultats de classification obtenus lorsque l'on utilise l'indice I_1 pour la réduction de la dimension de l'espace sont en général meilleurs que lorsque l'on utilise l'indice I_2 . Les résultats présentés par la suite sont donc obtenus avec le premier indice.

4.2. problème à plusieurs classes

Le cas où le nombre de classes est supérieur à 2 est beaucoup plus vraisemblable dans le cadre de la télédétection. Or, la distance de Bhattacharyya ne peut se calculer qu'entre 2 distributions. Une méthode pour construire l'indice correspondant à un problème de ce type consiste à prendre la plus petite des distances entre classes :

$$I = \min_{i,j} I_{i,j} \quad (36)$$

$I_{i,j}$ représentant un des deux indices proposés en 4.1 et correspondant au problème de classification entre i et j .

On peut même aller plus loin. En effet, si l'on est intéressé par la classification des zones urbaines par exemple, il suffit de ne chercher ce minimum que dans l'ensemble des couples de classes dont l'une d'elles au moins est une classe urbaine. Les distances entre classes non urbaines ne nous intéressent pas, car on ne cherche pas à les classifier.

4.3. algorithme global de classification

On a donc à notre disposition un algorithme permettant de réduire la dimension des données ainsi qu'une modélisation des textures. Lors du choix de l'indice, fondé sur la distance de Bhattacharyya, il est implicitement fait un autre choix, celui du critère de classification. En effet, cette distance est liée à la probabilité d'erreur de classification selon le critère du maximum de vraisemblance. Les résultats que l'on montre sont donc obtenus en utilisant le critère suivant, similaire à celui décrit par l'équation (31) :

$$\hat{\lambda}_s = \arg \max_{l \in \Lambda} P_l(x_s | \{m_s^k\}) \quad (37)$$

où Λ est l'ensemble des classes.

Les images des étiquettes $\hat{\lambda}_s$ sont très irrégulières. On va chercher à obtenir des segmentations où les zones représentant une même classe sont plus homogènes, on appelle cela la régularisation. Pour cela, on introduit un modèle *a priori* de Potts [Wu82] sur le champ d'étiquettes, ce qui pénalise les situations où les étiquettes $\hat{\lambda}_s$ et $\hat{\lambda}_t$, $t \in V_s$, sont différentes.

Si l'on récapitule ce que l'on vient d'exposer, on aboutit à l'algorithme global de classification des pixels d'images hyperspectrales texturées décrit par la figure 4, où les deux parties du traitement sont illustrées, à savoir la réduction de dimension de l'espace et la classification. On remarque que les méthodes d'estimation des paramètres par maximum de vraisemblance et maximum de pseudo-vraisemblance sont utilisées à deux étapes différentes. En effet, lors de l'optimisation de l'indice de projection, ce dernier est évalué un grand nombre de fois, et il est nécessaire d'avoir une méthode d'estimation rapide. En revanche, l'apprentissage des classes ne se fait qu'une seule fois, il n'est donc pas très pénalisant d'utiliser le maximum de vraisemblance, qui donne des estimées plus fiables.

5. résultats

5.1. collage de textures urbaines

Avant d'examiner les résultats sur des scènes réelles, les méthodes de réduction de dimension et de classification sont appliquées à la segmentation d'un collage de différentes textures

Classification de textures hyperspectrales fondée sur un modèle markovien

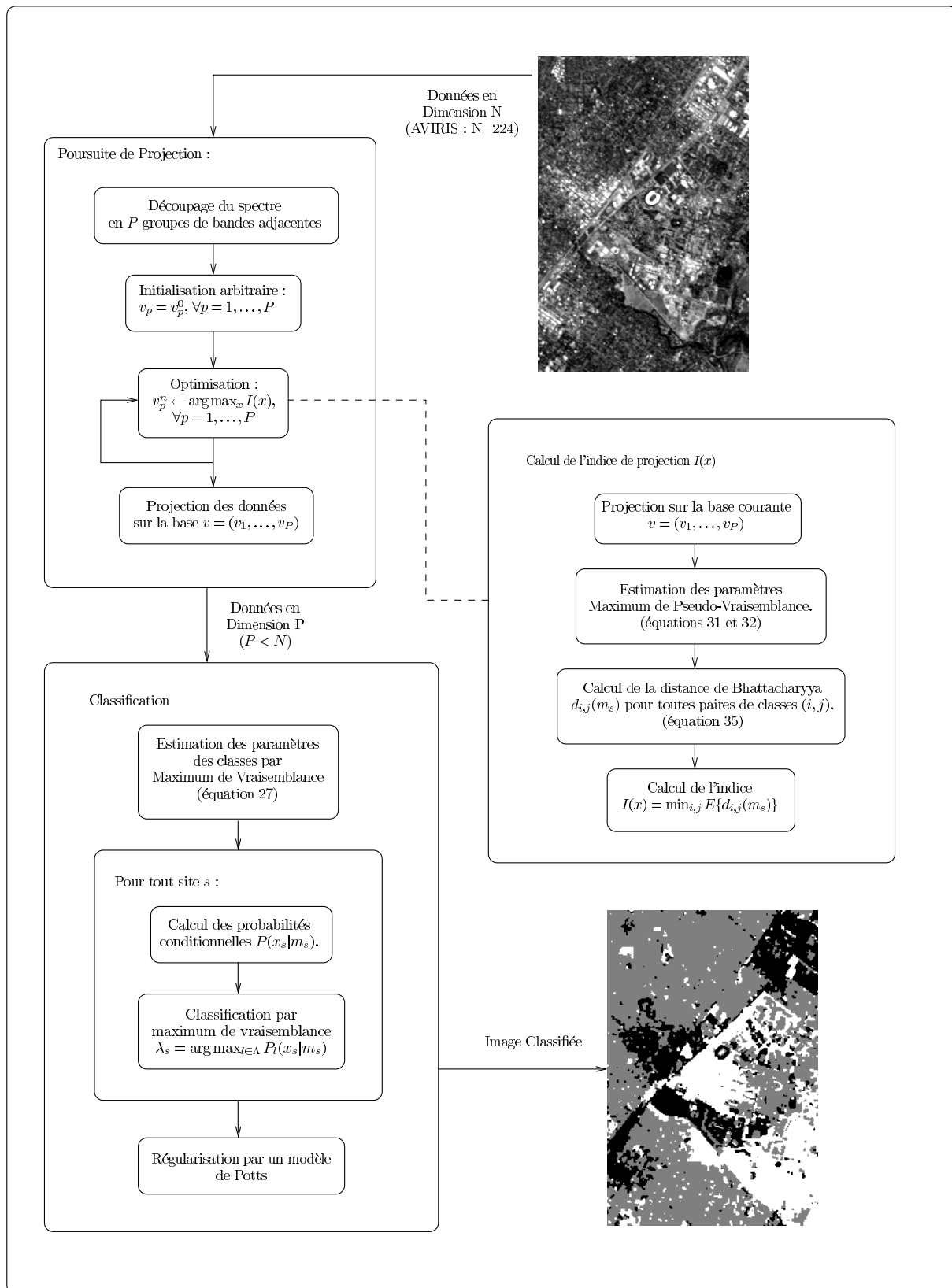


Figure 4. – Algorithme global de classification d'images hyperspectrales texturées.

extraites d'une image AVIRIS prise en 1997 au-dessus de la région de Moffett Field en Californie. Les résultats sur ces collages sont plus facilement quantifiables que sur de véritables images de télédétection, puisque l'on connaît la vérité terrain. La figure 5 représente un affichage des bandes numéro 30, 20 et 10 de ce collage. L'image est constituée de quatre carrés de petites dimensions (64×64) constitués chacun d'une textures différente, représentant des tissus urbains différents, et qui sont donc les quatre classes à distinguer. Les échantillons utilisés pour l'apprentissage sont les pixels du carré central de chaque sous-image, de dimension 19×19 , soit 361 pixels par texture.

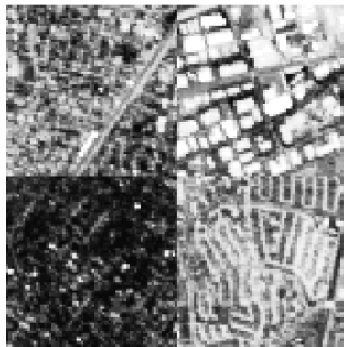


Figure 5. – Affichage des canaux {30,20,10} d'un collage de textures extraites d'une scène AVIRIS prise au-dessus de Moffett Field en Californie.

On compare les résultats de classification donnés par plusieurs méthodes et diverses modélisations dans la figure 6 pour différentes dimensions de l'espace de projection. L'analyse en composantes principales, notée ACP, est suivie d'une classification par maximum de vraisemblance fondée sur un modèle gaussien où les X_s sont indépendants. La poursuite de projection fondée sur le même modèle gaussien, proposée par Jimenez et Landgrebe [JL98], où l'indice de projection provient de la distance de Bhattacharyya calculée sur ce modèle, est notée PP-G, et le critère de classification est aussi le maximum de vraisemblance associé au modèle gaussien. La poursuite de projection utilisant un modèle markovien gaussien est notée PP-MG, et la classification associée repose sur le critère de maximum de vraisemblance associé au modèle markovien gaussien. Les résultats sont régularisés, par un algorithme de recuit simulé [GG84], en introduisant un modèle *a priori* de Potts [Wu82].

Les résultats montrent le pouvoir discriminant des méthodes testées, dans le cas de textures bien homogènes, et pour des ensembles d'apprentissage réduits. On constate que, dans le cas de ces collages d'images urbaines, utiliser une classification fondée sur la texture n'apporte pas forcément de meilleurs résultats de classification pour toutes les dimensions de l'espace final testées. En effet, le taux de bonne classification dans le cas du PP-MG est altéré par les transitions très abruptes entre textures qui créent des effets de bords pour les pixels limitrophes dont certains voisins n'appartiennent pas à la même texture. Les

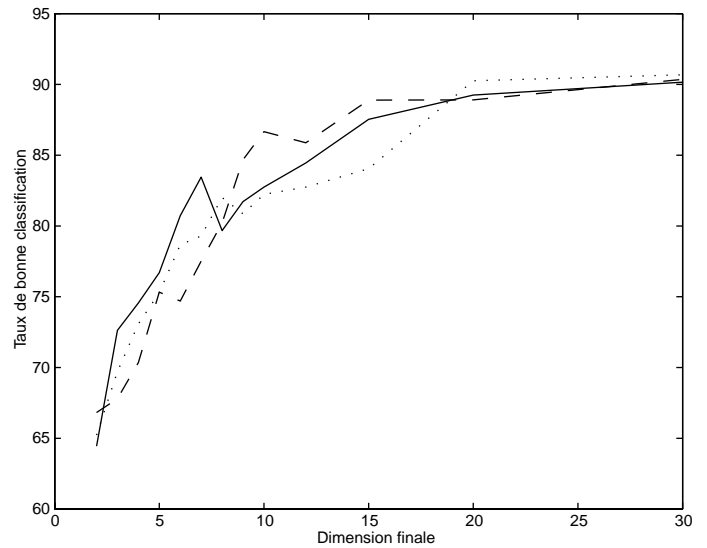


Figure 6. – Taux de bonne classification du collage de textures par différents modèles et méthodes de réduction d'espace, pour différentes dimensions finales. Trait plein : PP-MG, trait discontinu : PP-G, pointillés : ACP.

résultats sont assez similaires et aucune des méthodes testées ne semble supérieure aux autres dans tous les cas. Pour des textures à faible variabilité, les modèles pixéliques sont donc capables de saisir la majorité de l'information présente.

5.2. résultats sur une scène AVIRIS

Dans un deuxième temps, on effectue des tests de classification sur l'image réelle AVIRIS de Moffett Field.

Dans ces tests, on compare les résultats de classification pour les méthodes PP-G et PP-MG.

La figure 7 montre un affichage des bandes numéro 30, 20 et 10 du spectre de l'image de test. Celle-ci est approximativement divisée d'une part, en une zone de végétation et d'eau, et d'autre part, en une zone urbaine, que nous considérons avec plus d'attention. Elle est elle-même divisée en deux zones : une zone de petits bâtiments d'habitations et une zone de grands bâtiments assez longs et larges, dont les toits sont plus réfléchissants que ceux des habitations.

Dans toutes les figures, les résultats montrent la classification des zones urbaines, les pixels assignés aux autres classes étant regroupés en une classe globale dite de « non-urbain ». On peut donc voir trois classes sur les images de résultats : les deux classes urbaines et la classe non-urbaine.

La figure 8 montre trois résultats de classification de la même image. Le premier est celui de la classification par maximum de vraisemblance sur un modèle gaussien sur les 10 premières

Classification de textures hyperspectrales fondée sur un modèle markovien

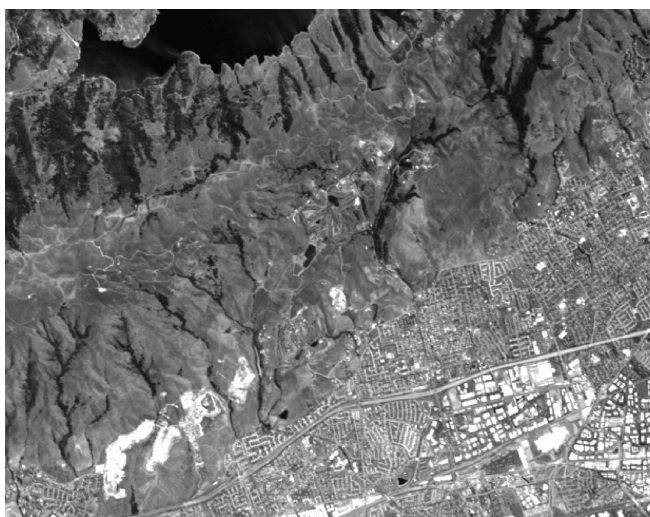


Figure 7. – Affichage des canaux {30,20,10} d'une scène AVIRIS prise au-dessus de Moffett Field en Californie.



(a)



(b)



(c)



(d)

Figure 8. – a) Une bande de l'image initiale, b) résultat de classification sur les 10 premières composantes principales, par maximum de vraisemblance sur modèle gaussien pixélique, c) résultat de classification sur 10 bandes extraites par poursuite de projection sur modèle gaussien pixélique, d) résultat de classification sur 10 bandes extraites par poursuite de projection sur le modèle markovien gaussien.

composantes principales, le deuxième est celui donné par la méthode de Jimenez et Landgrebe [JL99] par poursuite de projection sur modèle gaussien pixélique en gardant également 10 bandes, et le dernier résultat est celui donné par la méthode proposée dans cet article en gardant 10 bandes. Le choix de 10 bandes est justifié par le fait qu'il offre un bon compromis entre le temps de calcul et le taux de bonne classification, car ce dernier augmente faiblement au-delà de ce seuil de 10 bandes.

En ce qui concerne la différenciation entre urbain et non urbain, il semble que les trois méthodes aient à peu de chose près la même efficacité, bien que les deux premières produisent des fausses détections au niveau des eaux peu profondes de l'étang. En dehors de ces fausses alarmes et au niveau de l'analyse intra-

urbaine, on note une efficacité équivalente des deux méthodes de poursuite de projection comparées à l'ACP qui a tendance à détecter de gros bâtiments dans les zones résidentielles.

La figure 9 montre les résultats que donnent les méthodes citées ci-dessus, en gardant 20 bandes au lieu de 10. Dans ce cas, on constate que les deux premières méthodes produisent un grand nombre de fausses détections de grands bâtiments, dont beaucoup sont détectés dans les zones résidentielles.

D'autre part, comme on pouvait s'y attendre, le résultat de la méthode utilisant l'ACP à 20 composantes diffère très peu de celui à 10 composantes, puisque les bandes ajoutées sont de moins en moins informatives. La première méthode produit une meilleure discrimination entre les deux types de bâti.

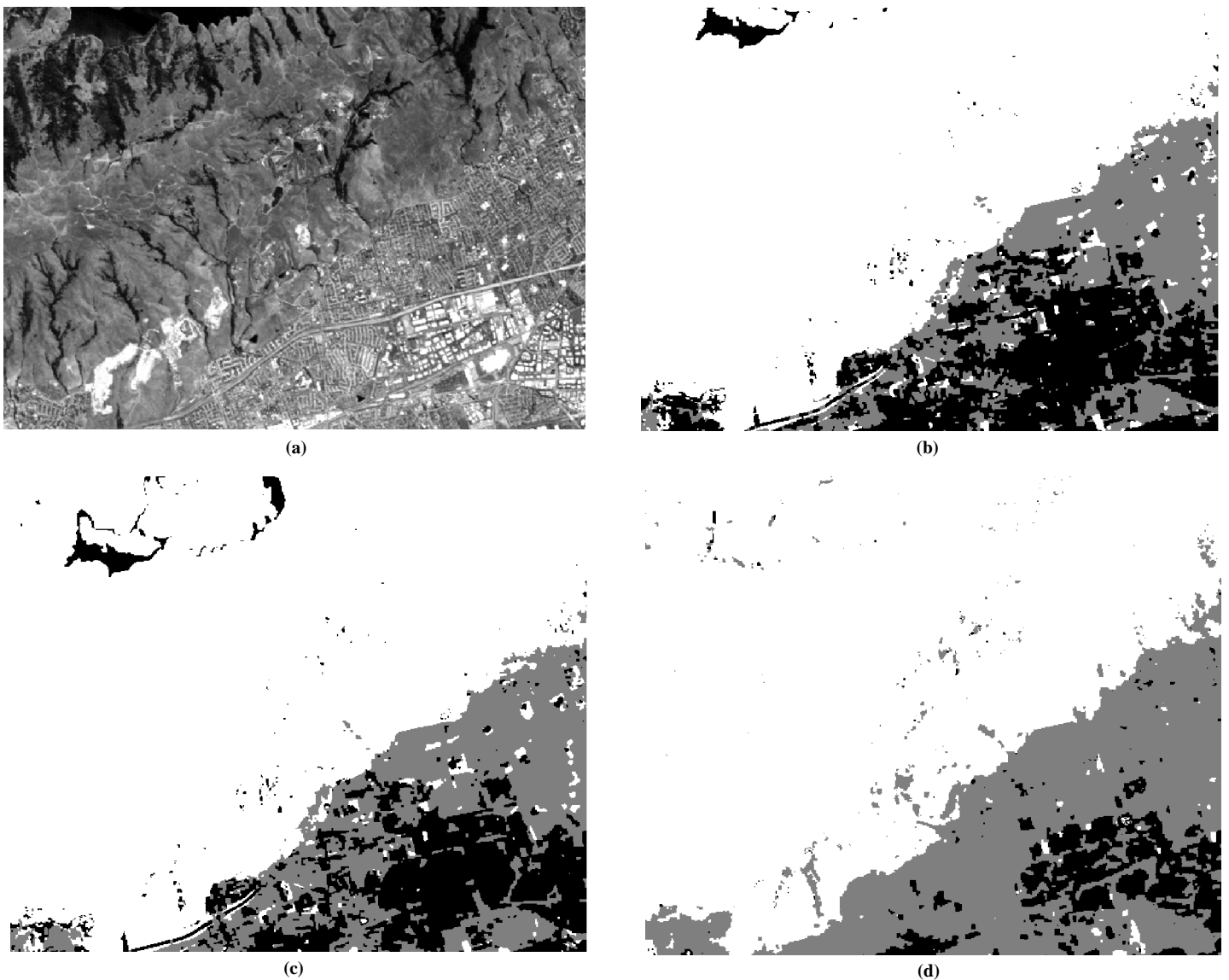


Figure 9. – a) Une bande de l'image initiale, b) résultat de classification sur les 20 premières composantes principales, par maximum de vraisemblance sur modèle gaussien pixélique, c) résultat de classification sur 20 bandes extraites par poursuite de projection sur modèle gaussien pixélique, d) résultat de classification sur 20 bandes extraites par poursuite de projection sur le modèle markovien gaussien.

5.3. temps de calcul

Le temps de calcul de la matrice de projection pour les différentes méthodes est très variable. Tout d'abord, notons que pour l'ACP ce temps dépend de la dimension des espaces de départ et d'arrivée, mais aussi du nombre de pixels de l'image. En ce qui concerne la poursuite de projection, ce temps dépend des dimensions de l'espace initial, de celles de l'espace de projection et du nombre de pixels d'apprentissage.

Dans le cas de la scène AVIRIS traitée au paragraphe précédent, de dimension spatiale 614×512 , et ayant 224 bandes spectrales, les temps de calcul des matrices de projection correspondant aux différentes méthodes sont donnés dans le tableau 1.

Pour l'ACP, le temps de calcul est alourdi par le fait que l'image a des dimensions trop grandes pour être chargée en mémoire directement, et qu'il faut donc faire des lectures fréquentes sur le disque. La différence entre les deux méthodes de poursuite de projection s'explique par le fait que l'indice de projection est plus complexe dans le cas contextuel, puisqu'il s'agit d'une somme (ou d'un minimum) de l'indice sur les pixels d'apprentissage. Par ailleurs, le nombre d'itérations sur l'ensemble des bandes ainsi que le critère d'arrêt ϵ utilisé lors de la poursuite de projection influencent les temps de calcul des algorithmes de poursuite de projection.

Tableau 1. – Temps de calcul de la matrice de projection pour les 3 méthodes dans le cas de la scène AVIRIS de Moffett Field (sur une station Sun Ultra10).

Type de réduction d'espace	Temps de calcul de la projection
ACP	1h25mn
PP-G	30mn
PP-MG	1h40mn

6. conclusion

Cet article présente une méthode de classification supervisée pour l'imagerie hyperspectrale, utilisant l'information de texture présente dans les images.

Cette méthode est fondée sur une modélisation globale par champ de Markov gaussien multivarié et sur l'utilisation d'un algorithme de poursuite de projection pour réduire la dimension de l'espace de représentation de l'observation.

On a, dans un premier temps, exposé les raisons du passage de l'espace original à un espace de dimension réduite pour éviter les problèmes dus au phénomène de Hughes.

Dans une deuxième partie, on a présenté les méthodes les plus utilisées pour effectuer cette réduction de dimension, ainsi qu'un algorithme moins connu, la poursuite de projection, qui a la propriété d'être beaucoup moins sensible aux problèmes de dimensionalité, puisque les calculs y sont effectués dans le sous-espace projeté. On a proposé une variante de l'algorithme approché de Jimenez et Landgrebe [JL99], qui consiste, étant donné un indice de projection caractérisant la discrimination entre les classes, à découper le spectre en groupes de bandes ayant des contributions homogènes à l'indice de projection, et à extraire de chacun de ces groupes une composante, en maximisant l'indice de projection correspondant à l'espace généré par l'ensemble des composantes extraites. Mais cet algorithme, du fait de l'ensemble réduit des espaces de projection qu'il peut atteindre, n'assure pas d'obtenir l'espace de projection pour lequel l'indice de projection est maximal.

Enfin, la troisième partie de cet article a été consacrée à la modélisation des textures par champs de Markov gaussiens, qui permet de prendre en compte les textures des images et qui n'est pas très restrictive puisque dans le cadre d'une image non-texturée, donc sans dépendance spatiale inter-pixel, on retrouve le cas gaussien pixélique. On a proposé une méthode d'estimation des paramètres de ce modèle dans le cadre d'un modèle simplifié du fait d'hypothèses dues aux caractéristiques particulières des images hyperspectrales.

Cette modélisation est utilisée pour l'adaptation de l'algorithme de poursuite de projection. Dans cette adaptation, l'indice de projection à maximiser est lié à la distance de Bhattacharyya entre les classes, qui est calculée en se fondant sur cette modélisation.

Les résultats présentés dans la dernière partie font ressortir plusieurs points intéressants :

- Premièrement, la distinction entre classes urbaines et rurales semble être facilement réalisée, quelle que soit la méthode utilisée.
- Deuxièmement, on voit que dans le cas d'images dont les classes sont bien séparées radiométriquement, les résultats sont équivalents lorsque l'on utilise des méthodes classiques comme l'ACP, des méthodes moins classiques comme la poursuite de projection utilisant un modèle pixélique, ou utilisant le modèle contextuel proposé.
- Troisièmement, dans le cas de données réelles, on constate que les résultats paraissent meilleurs si l'on utilise le modèle markovien, même si ceci peut difficilement être validé en l'absence d'une vérité de terrain faite par un expert. Toutefois, il semble que le modèle proposé soit plus capable de saisir toute la variabilité des tissus urbains.

En ce qui concerne les travaux futurs, il est possible d'adapter l'algorithme au cas non-supervisé. La comparaison avec d'autres méthodes d'analyse de texture est également envisagée, par exemple celle utilisant les matrices de cooccurrence.

remerciements

Le premier auteur souhaite remercier la région Provence-Alpes-Côte d'Azur et Alcatel Space Industries pour le soutien financier apporté à sa thèse. Les auteurs remercient le JPL pour la mise à disposition de l'image AVIRIS de Moffett Field en Californie, le Dr. Robin Morris du Centre de Recherche Ames de la Nasa pour la fourniture des cartes, ainsi que les relecteurs de la revue *Traitement du Signal* pour leurs critiques constructives.

BIBLIOGRAPHIE

- [Bel61] R. Bellman, *Adaptive control processes : a guided tour*. Princeton Univ. Press., 1961.
- [Bes74] J. Besag, Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Stat. Soc. B*, 36:192-236, 1974.
- [BK99] J. Bennett et A. Khotanzad, Maximum likelihood estimation methods for multispectral random field image models. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 21(6):537-543, 1999.
- [CCB85] R. Chellappa, S. Chatterjee, et R. Bagdazian, Texture synthesis and compression using Gauss-Markov random fields. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(2):298-303, Mars-Avril 1985.
- [DE87] H. Derin et H. Elliott, Modeling and segmentation of noisy and textured images using Gibbs random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 9:721-741, 1987.
- [Des93] X. Descombes, *Champs markoviens en analyse d'images*. Thèse de Doctorat, ENST Paris, 1993.
- [Des97] X. Descombes, A dense class of Markov Random Fields and associated parameter estimation. *Journal of Visual Communication and Image Representation*, 8(3):299-316, Septembre 1997.
- [DMZ97] X. Descombes, R. Morris, et J. Zerubia, Quelques améliorations à la segmentation d'image bayésienne. *Traitement du Signal*, 14(4):373-382, 1997.
- [DSP99] X. Descombes, M. Sigelle, et F. Prêteux, Estimating Gaussian Markov random field parameters in a nonstationary framework : application to remote sensing imaging. *IEEE Transactions on Image Processing (IP)*, 8(4):490-503, 1999.
- [Dun89] G. H. Dunteman, *Principal Components Analysis*. Sage Inc., 1989.
- [Fuk90] K. Fukunaga, *Introduction to statistical pattern recognition*. Academic Press, 2^e édition, 1990.
- [GBSC88] A.A. Green, M. Berman, P. Switzer, et M.D. Craig, A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing (GRS)*, 26(1):65-74, 1988.
- [GG84] S. Geman et D. Geman, Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 6(6):721-741, Novembre 1984.
- [Gim99] G.L. Gimel'farb, *Image texture and Gibbs random fields*. Kluwer, 1999.
- [Har79] R. M. Haralick, Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786-804, 1979.
- [Haz00] G. G. Hazel, Multivariate Gaussian MRF for multispectral scene segmentation and anomaly detection. *IEEE Transactions on Geoscience and Remote Sensing (GRS)*, 38(3):1199-1211, 2000.
- [HL93] P. Hall et K. Li, On almost linearity of low dimensionality projections from high dimensional data. *The Annals of Statistics*, 21(2):867-889, 1993.
- [Hub85] P. J. Huber, Projection pursuit. *The Annals of Statistics*, 13(2):435-475, 1985.
- [Hug68] G.F. Hughes, On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14(1):55-63, 1968.
- [Jai89] A. K. Jain, *Fundamentals of image processing*, chapter 2 : Two-dimensional systems and mathematical preliminaries. Prentice Hall, 1989.
- [JF91] A. K. Jain et F. Farrokhnia, Unsupervised texture segmentation using Gabor filters. *Pattern Recognition*, 24(12):1167-1186, 1991.
- [JL98] L.O. Jimenez et D.A. Landgrebe, Supervised classification in high-dimensional space : geometrical, statistical and asymptotical properties of multivariate data. *IEEE Transactions on Systems, Man and Cybernetics-Part C : Applications and Reviews*, 28(1):39-54, Février 1998.
- [JL99] L. O. Jimenez et D. A. Landgrebe, Hyperspectral data analysis and supervised feature reduction via projection pursuit. *IEEE Transactions on Geoscience and Remote Sensing (GRS)*, 37(6):2653-2667, 1999.
- [JR99] X. Jia et J. A. Richards, Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification. *IEEE Transactions on Geoscience and Remote Sensing (GRS)*, 37(1):538-542, 1999.
- [KH95] C. Kervrann et F. Heitz, A Markov random field model-based approach to unsupervised texture segmentation using local and global spatial statistics. *IEEE Transactions on Image Processing (IP)*, 4(6):856-862, Juin 1995.
- [LC00] C. Lee et E. Choi, Bayes error evaluation of the Gaussian ML classifier. *IEEE Transactions on Geoscience and Remote Sensing (GRS)*, 38(3):1471-1475, 2000.
- [LDZ00] A. Lorette, X. Descombes, et J. Zerubia, Urban areas extraction based on texture analysis through a Markovian modelling. *International Journal of Computer Vision (IJCV)*, 36(3):219-234, 2000.
- [PH95] D. K. Panjwani et G. Healey, Markov random field models for unsupervised segmentation of textured color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 17(10):939-954, 1995.
- [SM00] S. M. Schweizer et J. M. F. Moura, Hyperspectral imagery : Clutter adaptation in anomaly detection. *IEEE Transactions on Information Theory*, 46(5):1855-1871, 2000.
- [Win95] G. Winkler, *Image analysis, random fields and dynamic Monte Carlo methods : a mathematical introduction*. Springer, 1995.
- [Wu82] F.Y. Wu, The Potts model. *Reviews of modern physics*, 54(1):235-267, 1982.

Manuscrit reçu le 18 janvier 2002

LES AUTEURS

Guillaume RELIER



Guillaume RELIER est diplômé de l'École Nationale Supérieure des Télécommunication et du DEA de Traitement d'Image de l'Université de Rennes I depuis 1999. Entre 1999 et 2002, il était doctorant en Analyse d'Image à l'INRIA Sophia Antipolis. Il a obtenu le grade de docteur fin 2002. Ses recherches concernent le traitement d'images et plus particulièrement l'analyse d'images hyperspectrales.

Frédéric FALZON

Frédéric FALZON est responsable du Service de Recherche sur les Systèmes d'Observation dans le Département Recherche d'Alcatel Space à Cannes. Il travaille sur le traitement des images et des signaux bords et sols ainsi que sur l'étude de nouvelles structures d'instruments dédiés à l'observation du ciel et de la terre.

Xavier DESCOMBES



Xavier DESCOMBES a obtenu les diplômes d'ingénieur de l'ENST en 1989, de DEA de mathématiques pures de l'université Pierre et Marie Curie en 1990 et de doctorat de l'ENST en 1993. Il est ensuite successivement à l'université catholique de Leuven (Belgique), à l'INRIA (projet PASTIS) et à l'institut Max Planck de Leipzig (Allemagne). Il est actuellement chargé de

recherche à l'INRIA depuis 1997 et chargé de cours à Sup'Aero depuis 1999. Ses travaux, concernant les champs de Markov et les modélisations par processus objet, ont été appliqués à l'imagerie satellitale et médicale en vue de la segmentation, la restauration d'image ou l'extraction d'objets.

Josiane ZERUBIA



Josiane ZERUBIA est chercheur permanent à l'INRIA depuis 1989. Elle est directrice de recherche depuis 1995, a été responsable du projet PASTIS (INRIA Sophia Antipolis) de 1995 à 1997 et est responsable du projet ARIANA (projet commun CNRS/INRIA/UNSA) depuis 1998. Elle est également professeur à Sup'Aero depuis 1999. Auparavant, elle était à l'institut de traitement

des signaux et des images de l'université de Californie du Sud à Los Angeles en tant que Post doctorante. Elle a également travaillé comme chercheur au Lassy (Université de Nice et CNRS) entre 1984 et 1988, au laboratoire de recherche de Hewlett Packard en France et à Palo-Alto entre 1982 et 1984. Elle a obtenu un diplôme d'ingénieur de l'ENSIEG, Grenoble, en 1981, un doctorat d'ingénieur en 1986, un doctorat d'Université en 1988 et une habilitation à diriger des recherches en 1994. Elle est « fellow » de l'IEEE, membre du comité technique IEEE IMDSP, membre du « board of governors » de la société IEEE SP, adjointe à l'éditrice en chef de IEEE Trans. IP et membre du comité éditorial du bulletin de la SFTP. Son domaine d'intérêt est le traitement des images, en particulier la restauration d'image, la segmentation, classification, le groupement perceptuel, la super-résolution en utilisant des modèles probabilistes. Elle travaille également sur l'estimation de paramètres et les techniques d'optimisation.