

Segmentation Markovienne vectorielle non supervisée d'images radar polarimétriques

Unsupervised Vectorial Markov Segmentation of Polarimetric Radar Imaging

par Safwan EL ASSAD, Ali SAAD, Dominique BARBA

SEI (Systèmes Electroniques et Informatiques) IRESTE,
Rue Christian Pauc La Chantrerie B.P. 60601 44306 Nantes cedex 03. FRANCE.
Tél : (33)0240683036 Fax : (33)0240683066 E-mail : nelassad@ireste.

résumé et mots clés

La segmentation des images ROS (Radar à Ouverture Synthétique) consiste à produire une partition de l'image initiale en classes possédant certaines caractéristiques d'homogénéité au sens de la rétrodiffusion. Elle se pose en préalable à la plupart des schémas d'interprétation et de décision. Parmi les différentes techniques de segmentation possibles, nous avons privilégié celles basées sur une modélisation markovienne. A cet effet, un modèle statistique des images ROS polarimétriques est utilisé pour les segmenter en classes homogènes ayant chacune des caractéristiques de rétrodiffusion similaires. Cette segmentation utilise l'estimateur MAP (Maximum A Posteriori) que nous avons implémenté soit par l'algorithme déterministe ICM Modifié (Iterative Conditional Modes), soit par l'algorithme stochastique RS (Recuit Simulé). L'estimateur MAP correspond à la maximisation de la probabilité *a posteriori* de l'image d'un champ d'étiquettes de segmentation conditionnellement à la mesure faite. Pour cela un champ aléatoire de Markov représentant la distribution des étiquettes des classes est combiné avec un modèle (distribution Gaussienne ou distribution-K) représentant la distribution des données polarimétriques pour chacune des classes données. L'estimation des paramètres du modèle est non supervisée. Elle est basée sur une partition non contextuelle de l'image initiale obtenue par un groupement multidimensionnel des vecteurs caractéristiques soit par les algorithmes CMFMAP-NSO (Centres Mobiles Flous Non Supervisée et Optimale) et CMF-NS, soit par la méthode MV (Maximum de Vraisemblance) implémentée par l'algorithme SEM (Stochastic Estimation Maximisation). Les résultats de partition initiale et de segmentation obtenus sur une image mono-vue polarimétrique de la forêt des Landes, démontrent l'aptitude de l'algorithme proposé CMFMAP-NSO à fournir une partition de bonne qualité d'une part et la capacité de l'algorithme ICM à affiner cette partition et à produire une segmentation de meilleure qualité d'autre part.

Analyse d'images, segmentation vectorielle non supervisée, modèles statistiques, champs markoviens, imagerie radar à ouverture synthétique, groupement multidimensionnel, classification, estimation.

abstract and key words

A statistical image model is used for segmenting polarimetric synthetic aperture radar (SAR) data into regions of homogeneous polarimetric backscatter characteristics. A model for the conditional distribution of the polarimetric complex data is combined with a Markov random field representation for the distribution of the region labels, to obtain the *a posteriori* distribution. Optimal region labels of the data are those which maximize the *a posteriori* distribution of the region labels given the polarimetric complex data. Implementation of the MAP (Maximum A Posteriori) technique is accomplished either by the modified deterministic ICM (Iterative Conditional Modes) algorithm or by the stochastic SA (Simulated Annealing) algorithm. Unsupervised parameter estimation procedures are obtained either by the CMFMAP-NSO (unsupervised fuzzy partition - optimal number of classes) and CMF-NS or by SEM (Stochastic Estimation Maximization) algorithm. The CMFMAP-NSO algorithm performs well in situations of large variability of cluster shapes, and densities. Results, using fully polarimetric SAR forest data, obtained by the CMFMAP-NSO following by the ICM algorithm with a K distribution model are quite satisfactory.

Image analysis, unsupervised vectorial segmentation, statistical models, Markov field, synthetic aperture radar, multidimensional clustering, classification, estimation.

1. introduction

L'imagerie radar à visée latérale et à ouverture synthétique (ROS) a considérablement pris de l'importance depuis qu'il est devenu possible d'envoyer des satellites (ou avions) d'observation autour de notre planète. En effet, cette technique de télédétection hyperfréquence a l'avantage, par rapport à la télédétection optique, d'être insensible à la luminosité et aux conditions d'environnement d'ordre météorologique. Par ailleurs, l'information apportée par l'imagerie radar est complémentaire de celle fournie par l'imagerie optique puisqu'elle concerne la réflectivité électromagnétique (e.m) du sol et caractérise l'interaction physique entre l'onde e.m et la surface, alors que celle fournie par les ondes optiques concerne la couleur. L'imagerie radar a trouvé de très nombreuses applications, que ce soit dans le domaine civil pour la cartographie géologique (sol et sous-sol), l'hydrologie des surfaces continentales, l'océanographie (étude des vagues et des courants), l'agriculture (suivi de la croissance des cultures), et l'urbanisme, ou dans le domaine militaire pour l'observation et la surveillance des territoires. L'image de ROS polarimétrique a permis de quantifier et de mieux comprendre les propriétés de rétrodiffusion des cibles naturelles [Ula90].

Au vu de la nature aléatoire des textures dans les images ROS que nous analysons, les approches statistiques sont bien adaptées à l'analyse de ce type d'images. En général, ce type d'approches s'appuie sur des modélisations statistiques locales (niveau région) et globales (niveau image).

Van Zyl *et al.* [Van89] et Lim *et al.* [Lim89] ont utilisé un classifieur de type maximum de vraisemblance *a priori*, n'utilisant que des modèles locaux des données d'une région. Cette approche est moins intéressante que celle de type maximum de vraisemblance (MAP). En effet, l'introduction de modèles statistiques globaux améliore fortement la segmentation des images texturées [Dre87, Dre90, Dub89, Gem86, Har79, Gra86, Man91, Mar94]. Les modèles markoviens en particulier, permettent la prise en compte des variabilités de la structure des textures naturelles analysées.

Dans le cas où le nombre de régions et les paramètres des modèles sont connus *a priori*, ou font l'objet d'un apprentissage préalable, les résultats de la segmentation supervisée sont de bonne qualité [Kon88, Her92]. Les techniques non supervisées visent à s'affranchir de l'apprentissage *a priori* et on note de nombreux efforts récents dans ce domaine [Che85, Fan86, Kho89, Gem90, Man91, Won92, Des93]. Il faut noter qu'une segmentation robuste complètement non supervisée reste actuellement hors d'atteinte. Dans [Fan86, Man91, Rig91] les paramètres de modèles statistiques stationnaires de textures sont identifiés sur des fenêtres disjointes. Les fenêtres présentant des caractéristiques voisines sont ensuite fusionnées. Cette première segmentation est ensuite raffinée par des méthodes de type relaxation.

Lakshmanan *et al.* [Lak89] ont proposé un algorithme de segmentation non supervisée s'appuyant sur des modèles markoviens

dont les paramètres sont estimés simultanément avec l'image segmentée. Une technique similaire a été présentée par Won *et al.* [Won92] pour des modèles gaussiens markoviens de texture, avec la différence que le nombre de régions ici est déterminé à l'aide d'un critère statistique d'information. Une approche voisine est développée dans [Zha88, Gat89, Xie91] pour déterminer le nombre de classes dans une image.

Dans la classification des données ROS utilisant le classifieur de Bayes, la majorité des techniques sont supervisées car elles demandent la sélection de zones d'apprentissage pour chaque classe de terrain, ou semi-supervisées parce qu'elles connaissent seulement le nombre de classes *a priori* [Kel88, Rig92].

Dans ce papier nous traitons le problème de la segmentation vectorielle complètement non supervisée des données polarimétriques en classes homogènes au sens de la rétrodiffusion. La segmentation est contextuelle et est basée sur une approche markovienne permettant de pouvoir faire de l'estimation au sens du maximum *a posteriori*. Pour cela un modèle de distribution conditionnelle représentant les données ROS dans une classe est combiné avec un modèle de Markov représentant la distribution des labels des classes. Deux distributions conditionnelles *a priori* des données mesurées sont testées et comparées. La première est la distribution de Gauss couramment employée dans la littérature, la seconde est la distribution K , mieux adaptée au type d'images que nous traitons dans ce papier [Elas93].

Deux méthodes d'optimisation pour le classifieur MAP ont été implantées, l'algorithme ICM (Iterative Conditional Modes) [Bes74, Bes86] sous une forme modifiée, la méthode du recuit simulé RS (qui est beaucoup plus longue à converger).

L'initialisation de l'algorithme ICM nécessite une image de segmentation initiale pour l'estimation des paramètres du modèle statistique adopté. Cette segmentation non contextuelle est obtenue par l'un des trois algorithmes suivants; le SEM (Stochastic Estimation Maximisation), le CMF-NS (Centres Mobiles Flous Non Supervisée) utilisant une distance euclidienne simple et le CMFMAP-NSO (Centres Mobiles Flous Non Supervisée et Optimale) utilisant une distance exponentielle au sens du maximum de vraisemblance *a posteriori*. Une analyse comparative est proposée du comportement de ces algorithmes de classification sur des données vectorielles expérimentales, en soulignant la spécificité de chacun.

Contrairement à l'algorithme SEM qui fournit une partition dépendante de la configuration initiale, l'algorithme de classification automatique des données CMFMAP-NSO que nous proposons a l'avantage de contourner cet inconvénient et d'être adaptable par classe. Par contre les trois algorithmes développés ne nécessitent pas la connaissance du nombre de classes.

Le papier est organisé de la façon suivante. Dans la partie 2, la segmentation non contextuelle de l'image est posée, vue comme un problème de mélange de densités. Dans la section 2.1, nous rappelons la méthode du maximum de vraisemblance et son implémentation par l'algorithme SEM. Dans la section 2.2, nous décrivons l'algorithme de classification CMFMAP-NSO

avec toutes ses étapes y compris les procédures de recherche automatique des centres initiaux des classes. La section 2.3 décrit les données expérimentales et le vecteur caractéristique utilisé. Dans la section 2.4, nous commentons et discutons les résultats de classification obtenus par les trois algorithmes cités auparavant.

Dans la partie 3, nous introduisons la segmentation contextuelle au sens du critère MAP. Dans la section 3.1, les modèles statistiques du signal d'image ROS sont posés. Puis la section 3.2 décrit l'optimisation du critère MAP par les algorithmes ICM modifié et le RS.

Les résultats expérimentaux sur les images ROS polarimétriques sont fournis et discutés. Une comparaison entre les deux méthodes d'optimisation est également donnée dans la partie 4 avant de conclure.

2. segmentation non contextuelle de l'image, vue comme un problème de mélange de densités

Nous savons que la méthode d'optimisation ICM nécessite une distribution initiale des étiquettes. L'image des étiquettes est obtenue par une segmentation non contextuelle des données radar. Cette opération peut être vue comme un problème de reconnaissance de mélange de densités de probabilités paramétrées de même structure. L'objectif est de retrouver la composante dont est issue chaque point du mélange. Cela revient à déterminer les paramètres du mélange à partir de l'échantillonnage d'observation E . Ce problème peut être vu sous deux aspects [Fir94] :

- l'aspect estimation par l'approche du maximum de vraisemblance,
- l'aspect classification par la Méthode des Centres Mobiles Flous-Non Supervisée et Optimale (méthode CMFMAP-NSO) [Cel89], [Gat89].

Dans la suite nous exposons ces deux aspects.

2.1. approche du maximum de vraisemblance [Ren84]

On admet que les données traitées constituent un échantillon $E = \{X_1, \dots, X_n, \dots, X_N\}$ de N observations indépendantes d'une variable aléatoire X à valeurs dans R^d avec $X_n = (X_{n1}, \dots, X_{nd})^T$, d étant la dimension de l'espace des données. On suppose que les observations appartiennent à K classes et

proviennent d'un mélange de K densités de probabilités de lois identiques. Soit X_n un point de l'échantillon, sa loi est alors :

$$P(X = X_n) = \sum_{k=1}^K P(X \in C_k) P(X = X_n / X \in C_k) \quad (1)$$

avec

$$\sum_{k=1}^K P(X \in C_k) = 1, \quad 0 < P(X \in C_k) \leq 1 \quad \forall k = 1, \dots, K \quad (2)$$

Le théorème de Bayes permet d'écrire :

$$P(X \in C_k / X = X_n) = \frac{P(X \in C_k) P(X = X_n / X \in C_k)}{P(X = X_n)} \quad (3)$$

Par la suite, pour simplifier les notations, on note :

$P(C_k / X_n) = P(X \in C_k / X = X_n)$ la probabilité conditionnelle *a posteriori* de la classe C_k sachant X_n , $P(X_n / C_k) = P(X = X_n / X \in C_k) = \psi(X_n, \Theta_k)$ est la densité de probabilité conditionnelle *a priori*, dépendante du vecteur des paramètres inconnus Θ_k de la classe C_k , $P_k = P(X \in C_k)$ est la probabilité *a priori* de la classe C_k ; $\Theta_k = (\mu_k, \Sigma_k)$ le vecteur des paramètres du modèle dans le cas d'une loi gaussienne.

On estime les paramètres du mélange P_k et Θ_k par les estimateurs du maximum de vraisemblance \hat{P}_k et $\hat{\Theta}_k$ qui sont obtenus en maximisant le logarithme de la vraisemblance de l'échantillon E , donné (sous la contrainte de l'équation (2)) par :

$$\text{Log}(L) = \text{Log} \left[\prod_{n=1}^N P(X_n) \right] = \sum_{n=1}^N \text{Log} [P(X_n)] \quad (4)$$

En utilisant la méthode du multiplicateur de Lagrange, nous obtenons une nouvelle fonction $L1$:

$$\text{Log}(L1) = \sum_{n=1}^N \text{Log} [P(X_n)] - \lambda \left[\sum_{k=1}^K P_k - 1 \right] \quad (5)$$

L'estimation \hat{P}_k est obtenue par annulation des dérivées partielles par rapport à P_k de l'expression (5), puis multiplication par \hat{P}_k et utilisation de l'équation (3), soit (voir annexe) :

$$\hat{P}_k = \frac{1}{\lambda} \sum_{n=1}^N \hat{P}(C_k / X_n) \quad (6)$$

La valeur de $\lambda = N$ est trouvée en sommant l'équation (6) par rapport à k et en utilisant les équations (1)-(3).

Par ailleurs l'annulation du gradient partiel par rapport au vecteur de paramètres Θ_k de l'expression (5) et l'utilisation de l'équation (3), permet d'obtenir (voir annexe) :

$$\sum_{n=1}^N P(C_k / X_n) \frac{\partial}{\partial \Theta_k} \text{Log} [\psi(X_n, \Theta_k)] = 0 \quad (7)$$

Si X est une variable aléatoire complexe dont les parties réelle et imaginaire sont considérées comme des variables aléatoires

gaussiennes centrées et de même variance, alors on montre [Duv91] que X suit une loi gaussienne circulaire.

Une densité de probabilité conditionnelle $\psi(X_n, \Theta_k)$ gaussienne circulaire, est complètement déterminée par sa moyenne $\mu_k = E[X_n]_k$ et sa matrice de covariance $\Sigma_k = E[(X_n - \mu_k)(X_n - \mu_k)^T]_k$:

$$\begin{aligned} \psi(X_n, \Theta_k) &= g_k(X_n, \mu_k, \Sigma_k) \\ &= \pi^{-d} |\Sigma_k|^{-1} \exp \left\{ -(X_n - \mu_k)^T \Sigma_k^{-1} (X_n - \mu_k) \right\} \end{aligned} \quad (8)$$

Le calcul des dérivées partielles respectivement par rapport à μ_k et Σ_k du Log de $g_k(X_n, \mu_k, \Sigma_k)$. Puis leur remplacement dans (7), permettent d'estimer $\hat{\mu}_k$ et $\hat{\Sigma}_k$ soit :

$$\hat{\mu}_k = \frac{\sum_{n=1}^N \hat{P}(C_k/X_n) X_n}{\sum_{n=1}^N \hat{P}(C_k/X_n)} \quad (9)$$

$$\hat{\Sigma}_k = \frac{\sum_{n=1}^N \hat{P}(C_k/X_n) [(X_n - \hat{\mu}_k)(X_n - \hat{\mu}_k)^T]}{\sum_{n=1}^N \hat{P}(C_k/X_n)} \quad (10)$$

Les formules (3), (6), (9), (10) seront calculées de façon itérative par l'algorithme SEM que nous rappelons ci-après. L'application de la méthode du maximum de vraisemblance nécessite donc l'estimation de tous les paramètres du mélange de densités. Pour chaque composante du mélange, les paramètres inconnus à calculer sont :

- la probabilité *a priori*,
- le vecteur moyen à d inconnues,
- la matrice de covariance qui est symétrique définie positive et a donc au maximum $d(d+1)/2$ inconnues.

Ceci implique $K(d+1)(d+2)/2$ paramètres inconnus à déterminer pour l'ensemble des classes.

L'algorithme SEM [Cel86], [Mas93].

Cet algorithme fonctionne ainsi.

Au départ, on fixe le paramètre K majorant (nombre maximal supposé) du nombre de composantes du mélange et un seuil $c(N, d)$ compris entre 0 et 1.

Initialisation :

En chaque point X_n , $n = 1, \dots, N$, on choisit au hasard les probabilités initiales d'appartenance à l'une des composantes :

Soient

$$P^1(C_k/X_n) k = 1, \dots, K$$

avec

$$0 < P^1(C_k/X_n) < 1 \quad \text{et} \quad \sum_{k=1}^K P^1(C_k/X_n) = 1 \quad (11)$$

Itération i

Étape stochastique :

On tire en chaque point X_n la variable aléatoire multinomiale

$$e^i(X_n) = \{e^i(C_k/X_n); k = 1, \dots, K\} \quad (12)$$

d'ordre un et de paramètre $\{P^i(C_k/X_n), k = 1, \dots, K\}$.

Les réalisations $e^i(X_n)$ définissent une partition

$\mathcal{P}^i = \{\mathcal{P}_1^i, \dots, \mathcal{P}_K^i\}$ de l'échantillon avec :

$$\mathcal{P}_k^i = \{X_n / e^i(C_k/X_n) = 1\}$$

Si pour un certain k , $\text{card}(\mathcal{P}_k^i)$ est plus petit que $N \cdot c(N, d)$, l'algorithme est ré-initialisé avec le nombre de classes majorant diminué de un, $K = K - 1$. Sinon, on passe à l'étape de maximisation.

Étape maximisation :

On calcule les estimateurs du maximum de vraisemblance $[\hat{P}_k^{i+1}, \hat{\Theta}_k^{i+1}]$ des paramètres du mélange sur la base des sous-échantillons $\mathcal{P}_k^i, k = 1, \dots, K$.

L'estimation de \hat{P}_k^{i+1} est donnée par l'équation (6) en remplaçant $\hat{P}(C_k/X_n)$ par $e^i(C_k/X_n)$. L'estimation des $\hat{\Theta}_k^{i+1}$ dépend bien sûr de la famille paramétrée, posée *a priori*, des composantes du mélange. Dans le cas où l'espérance μ_k , ou la matrice de covariance Σ_k sont les constituants des paramètres (cas de mélanges gaussiens, de Poisson, d'exponentielles, loi en K, \dots), les estimations à l'itération i sont données respectivement par les relations (9) et (10) en remplaçant dans celles-ci les $\hat{P}(C_k/X_n)$ par $e^i(C_k/X_n)$.

Étape estimation :

À partir des $\hat{P}_k^{i+1}, \Theta_k^{i+1}$, on calcule pour $k = 1, \dots, K; n = 1, \dots, N$, $\hat{P}^{i+1}(C_k/X_n)$ donné par la relation (3).

Cette procédure est répétée jusqu'à la stabilité.

2.2. méthode des Centres Mobiles Flous Non Supervisée et Optimale (CMFMAP-NSO)

Contrairement à la méthode du maximum de vraisemblance qui est une méthode d'estimation, la méthode utilisée ici est une méthode de classification ou partitionnement, dérivée de la méthode générale des nuées dynamiques, dans laquelle chaque classe est représentée par son centre de gravité [Cel89], [Bez81]. Il s'agit de trouver le meilleur couple partition représentation optimisant un critère donné.

Dans la méthode classique des centres mobiles, le critère à optimiser est l'inertie intraclasse de la partition, la méthode a tendance à chercher des classes d'égal volume hypersphérique, dans un espace multidimensionnel, et de faible inertie. D'autre part, comme pour la plupart des méthodes de partitionnement, la solution obtenue dépend de la configuration initiale et nécessite la connaissance du nombre de classes.

La méthode proposée ici répond d'une part aux problèmes de formes et de densités quelconques des classes et d'autre part aux problèmes de l'initialisation et de la connaissance préalable du nombre des classes.

L'algorithme se déroule selon la stratégie générale suivante :

A) Classification selon la méthode de centre mobile classique CM utilisant la procédure de recherche automatique des centres des classes (section 2.2.3).

B) Classification selon l'algorithme CM flou utilisant une distance exponentielle adaptative par classe, estimée au sens de maximum de vraisemblance (raffinement de l'étape A) (section 2.2.1).

C) Calcul des critères de performances (section 2.2.2) de la classification obtenue et test sur le nombre k de classes atteint. Si $k < K$ alors sauvegarde des résultats des critères de performances $f(k)$ et retour à l'étape A), sinon arrêt et $\text{argmin}_k f(k)$ donne le nombre exact de classes. L'algorithme est alors lancé avec le nombre exact de classes.

2.2.1. l'algorithme des centres mobiles flous par maximum de vraisemblance

L'algorithme des centres mobiles flous CMF est basé sur la minimisation de façon itérative de la fonction $E_q(\mathcal{P}, \mu)$ dépendante du couple $\{\mathcal{P}, \mu\}$, où :

\mathcal{P} est l'ensemble des partitions $(\mathcal{P}_1, \dots, \mathcal{P}_K)$ à K classes et μ est l'ensemble des centres (μ_1, \dots, μ_k) des différentes classes.

$$E_q(\mathcal{P}, \mu) = \sum_{n=1}^N \sum_{k=1}^K [P(C_k/X_n)]^q \cdot d_A^2(X_n, \mu_k) \quad (13)$$

avec

$$d_A^2(X_n, \mu_k) = (X_n - \mu_k)^T A (X_n - \mu_k) \quad (14)$$

$P(C_k/X_n)$ est la probabilité *a posteriori* pour que le vecteur caractéristique X_n appartienne à la classe C_k , $d_A(X_n, \mu_k)$ est une distance métrique entre X_n et μ_k , A est une matrice hermitienne définie positive de dimension $d \times d$, où d est la dimension du vecteur caractéristique, N est le nombre des vecteurs caractéristiques, K est le nombre de classes et $q > 1$ est un paramètre qui contrôle le degré de flou de la classification obtenue.

L'enchaînement des opérations pour minimiser (13) se déroule comme suit :

1) Etant donnés les centres initiaux μ_k .

2) Calcul de $P(C_k/X_n)$ pour $k = 1, \dots$, et $n = 1, \dots, N$

$$P(C_k/X_n) = \frac{[1/d_A^2(X_n, \mu_k)]^{1/(q-1)}}{\sum_{k=1}^K [1/d_A^2(X_n, \mu_k)]^{1/(q-1)}} \quad (15)$$

et de $E_q(\mathcal{P}, \mu)$ donnée par (13), avec :

$$d_A^2(X_n, \mu_k) = d_I^2(X_n, \mu_k) = |X_n - \mu_k|^2 \quad (16)$$

la distance euclidienne ($A = I$: matrice identité).

3) Calcul des nouveaux centres $\hat{\mu}_k$, matrices de covariance $\hat{\Sigma}_k$ et probabilités *a priori* \hat{P}_k pour $k = 1, \dots, K$.

$$\hat{\mu}_k = \frac{\sum_{n=1}^N [P(C_k/X_n)]^q \cdot X_n}{\sum_{n=1}^N [P(C_k/X_n)]^q} \quad (17)$$

$$\hat{\Sigma}_k = \frac{\sum_{n=1}^N [P(C_k/X_n)]^q [(X_n - \hat{\mu}_k)(X_n - \hat{\mu}_k)^T]}{\sum_{n=1}^N [P(C_k/X_n)]^q} \quad (18)$$

$$\hat{P}_k = \frac{1}{N} \sum_{n=1}^N [P(C_k/X_n)]^q \quad (19)$$

4) Mise à jour de $P(C_k/X_n)$ à $\hat{P}(C_k/X_n)$ selon (15) et de $E_q(\mathcal{P}, \mu)$ à $\hat{E}_q(\mathcal{P}, \hat{\mu})$ selon (13), avec :

$$d_A^2(X_n, \hat{\mu}_k) = d_e^2(X_n, \hat{\mu}_k) = \frac{|\hat{\Sigma}_k|^{1/2}}{\hat{P}_k} \exp \left\{ \frac{1}{2} (X_n - \hat{\mu}_k)^T \hat{\Sigma}_k^{-1} (X_n - \hat{\mu}_k) \right\} \quad (20)$$

une distance exponentielle [Day69], permettant de tenir compte des formes ellipsoïdales des classes ainsi que des densités variables (le nombre de points dans une classe n'est pas proportionnel au volume de la classe considérée).

5) Comparaison de $\hat{E}_q(\mathcal{P}, \hat{\mu})$ avec celle de l'itération précédente. Si la différence est inférieure à ε alors arrêt, sinon retour à l'étape 3. L'algorithme CMF-NS (utilisé par la suite) se caractérise, par rapport à l'algorithme CMFMAP-NSO, par l'utilisation d'une distance euclidienne au niveau de l'étape 4).

2.2.2. calcul des critères de performances : détermination du nombre de classes

La détermination du nombre optimal de classes peut être liée au résultat attendu d'une classification, au sens de la maximisation du pourcentage de bonne classification des pixels.

Une classification optimale doit satisfaire certaines conditions [Gat89], [Xie91] :

– une bonne séparation entre les classes générées,

- une compacité maximale des classes (concentration maximale du nombre des vecteurs caractéristiques dans une classe donnée autour de son centre),
- un volume minimal de l'ensemble des classes.

A cet effet deux critères de performance sont proposés :

1) la fonction de compacité et de séparation CS définie comme étant le rapport entre la compacité des classes et le carré de la distance minimale entre les centres des classes. Elle est donnée par :

$$CS = \frac{E_2(\mathcal{P}, \mu)}{N \cdot d_{\min}^2} \quad (21)$$

On remarque que la compacité est la fonction de variation spatiale totale divisée par N . Elle dépend de la forme de la distribution des données dans les classes, mais est indépendante du nombre total N de ces données.

Une valeur de plus en plus faible de la compacité indique que les classes sont de plus en plus compactes. Donc plus la fonction de compacité et de séparation des classes CS est faible, meilleure est la partition obtenue, dans le sens où CS indique qu'on a atteint le meilleur rapport entre compacité des classes et une bonne séparation entre celles-ci. Notons que la relation (21) reste valable pour $q \neq 2$.

2) la fonction d'hypervolume V des classes définie par :

$$V = \sum_{k=1}^K |\Sigma_k|^{1/2} \quad (22)$$

L'hypervolume considéré ici n'est autre que le produit mixte des vecteurs formant la matrice de covariance Σ_k donnée par la relation (18). La meilleure partition est obtenue pour un hypervolume minimal.

2.2.3. procédures de recherche automatique des centres initiaux des classes

Les algorithmes de classification de type nuées dynamiques (ex : algorithme des centres mobiles) fournissent une solution de partitionnement dépendante des centres des classes choisis à l'initialisation. A l'issue de Q passages de ces algorithmes sur un même ensemble d'objets, on obtient Q partitions généralement non identiques. Pour trouver et accélérer la convergence vers la partition optimale, un algorithme d'initialisation des positions initiales optimales des centres de classes est vital. Trois types d'initialisation développés dans la littérature ont été testés, ainsi que la généralisation d'une de ces méthodes qui s'avère très efficace après comparaison selon un critère de performances précisé ultérieurement. Une description brève de ces méthodes est donnée ci-dessous.

Méthode aléatoire

Choix aléatoire des centres initiaux. Ceci est utilisé en général dans l'algorithme classique des centres mobiles.

Méthode du centre le plus éloigné

Le premier centre est le centroïde de l'ensemble des vecteurs à classer dans les K centres. Le deuxième centre choisi est celui le plus éloigné du premier au sens de la distance euclidienne. Puis à chacun des deux centres on associe les vecteurs les plus proches, ainsi on obtient deux classes dont on calcule les centroïdes correspondants. Pour choisir le troisième centre, on procède comme suit. Dans chaque classe, on sélectionne l'élément le plus mal classé dans sa classe, toujours au sens de la distance euclidienne. Puis parmi les plus mal classés, le nouveau centre est celui qui a la plus grande distance par rapport à son centre. On affecte ensuite les éléments aux centres les plus proches et on calcule les nouveaux centroïdes. Ainsi à chaque ajout d'un nouveau centre, on répète les mêmes opérations faites lors du choix du troisième centre et cela jusqu'à l'épuisement des centres.

Méthode de Linde et Linde modifiée

Le principe de cette méthode [Lin80], utilisée dans la quantification vectorielle, est décrit par les étapes suivantes :

- le premier centre est le centroïde de l'ensemble des vecteurs à classer dans les K classes,
- chaque centre est divisé en deux centres voisins $C_{i+\varepsilon}$ et $C_{i-\varepsilon}$, où ε est une petite perturbation, puis groupement des éléments aux centres les plus proches et calcul des nouveaux centroïdes,
- si le nombre des centres obtenu est égal au nombre K souhaité, on arrête; sinon, on recommence l'étape 2.

L'inconvénient de cette méthode est que le nombre total des centres est une puissance de deux. Or dans notre application le nombre de centres est un entier positif. Alors la modification de la méthode de Linde consiste, à chaque itération, à ne diviser qu'un seul centre. Le centre à diviser parmi tous les centres existants est celui de variance maximale. L'expression de la variance est donnée par :

$$\text{Var}_k = \frac{1}{N_k} \sum_{X_n \in C_k} d_I(X_n, \mu_k)$$

où N_k et μ_k sont respectivement le nombre de vecteurs et le centroïde de la classe C_k .

Comparaison des performances entre les méthodes

Nous avons comparé les méthodes entre elles selon les critères de performance suivants :

- erreur quadratique moyenne entre les centres trouvés par une méthode et les « vrais » centres obtenus après convergence de l'algorithme des centres mobiles,
- temps d'exécution d'une méthode.

Les résultats obtenus (voir tableau 1 ci-après) montrent que la méthode de Linde modifiée est la meilleure. C'est cette méthode que nous utilisons dans l'algorithme CMFMAP-NSO.

Tableau 1. – Performances des différentes méthodes.

Méthode utilisée	Nombre d'itérations	Temps de calcul(s) (Rs6000 IBM)	Erreur quadratique moyenne
Aléatoire	58	1	70.47
Centre de plus éloigné	46	1	0.805
Linde Généralisée	35	1	0.17

2.3. données expérimentales et choix du vecteur caractéristiques X_n utilisé dans les algorithmes SEM, CMF-NS et CMFMAP-NSO

Pour comparer les performances de ces trois algorithmes, nous avons utilisé les données en bande L (1.225 GHz) issues de la campagne MAESTRO-1 faite par la NASA/JPL utilisant le système polarimétrique AIRSAR avec un angle d'incidence compris entre 40° et 50°. Les données forment une image 1-vue et la résolution d'un pixel de l'image est de 6,6 m verticalement et de 3 m horizontalement. L'image traitée est une vue de la forêt des Landes. La forêt contient 5 classes d'arbres âgés de 1 à 46 ans, une classe contenant de l'herbe et quelques structures routières.

Nous avons découpé l'image globale en 5 sous-images pour tester la robustesse des algorithmes sur plusieurs images réelles. Chaque sous-image est formée de (512*250/9) fenêtres, la taille de chaque fenêtre est de (3*3) pixels. Chaque point vectoriel de dimension quatre, est calculé à partir des trois composantes polarimétriques HH , HV et W sur la fenêtre choisie.

La matrice de covariance utilisée dans les modèles représentant les données polarimétriques à l'intérieur d'une classe C_k (Gaussien, Distribution-K), caractérise complètement les statistiques de rétrodiffusion au second ordre de la classe considérée. Il est naturel alors de choisir comme composantes du vecteur caractéristique X_n , les éléments de la matrice de covariance, soit :

$$X_n = \begin{bmatrix} \sigma_{HH} = 10 \log_{10}(E[|HH|^2]) \\ \sigma_{HV} = 10 \log_{10}(E[|HV|^2]) \\ \sigma_{VV} = 10 \log_{10}(E[|VV|^2]) \\ 10 \log_{10}(|E[HH VV^*]|) \end{bmatrix} \quad (24)$$

Les trois premières composantes représentent les sections efficaces radar en db, mesurées à différentes polarisations de la fenêtre considérée.

La quatrième composante représente l'amplitude de la corrélation entre HH et VV . L'intérêt de considérer le logarithme des éléments de la matrice de covariance est double :

1) le speckle qui est un bruit multiplicatif sera transformé en un bruit additif, ce qui rend les classes moins sensibles à la présence du speckle. En effet lorsque le bruit est multiplicatif, la variance de

chaque composante du vecteur mesuré par le radar polarimétrique est proportionnelle au carré de sa valeur moyenne [Lee81]. Or la plupart des algorithmes de classification sont basés entre autres sur la minimisation de la variation spatiale totale intraclasse.

2) Avant passage au log la composante HV est presque toujours très faible par rapport aux composantes HH et VV , ce qui rend le résultat de la classification vectorielle principalement dépendant des termes de co-polarisation, à moins de donner des poids arbitraires aux différentes composantes, variant avec la nature des cibles et les conditions environnantes. Dans le domaine logarithmique cette difficulté est en partie contournée.

2.4. résultats expérimentaux et discussion

Nous avons repéré sur les sous-images à partir de la vérité terrain (carte réelle de la forêt Les-Landes) 6 types de parcelles de tranches d'âges différents concernant les arbres et une zone d'herbe.

C1 : classe de terrain contenant de l'herbe,

C2 : classe d'arbre d'âges 1-5 ans,

C3 : classe d'arbre d'âges 6-10 ans,

C4 : classe d'arbre d'âges 11-14 ans,

C5 : classe d'arbre d'âges 15-18 ans,

C6 : classe d'arbre d'âges 19-30 ans,

C7 : classe d'arbre d'âges 33-44 ans.

Par ailleurs, chaque sous-image est globalement formée de deux régions de réflectivité distinctes au sens de la rétrodiffusion : une région sombre R1 et une région claire R2.

Les résultats obtenus par les différents algorithmes de classification de données vectorielles sur les cinq sous-images sont comparés selon un premier critère (subjectif) qui juge la qualité visuelle de la classification et un second critère (objectif) qui quantifie le pourcentage de bonne classification des vecteurs.

Problèmes liés à l'application de l'algorithme SEM

Dans l'application de l'algorithme SEM il y a deux problèmes fondamentaux.

Le premier est le choix du seuil pour pouvoir éliminer une classe. Ce seuil doit être fixé à l'aide d'une connaissance *a priori* sur la taille minimale des objets dans l'image à traiter, ce qui rend cette méthode non complètement automatique. La précision de la classification des objets dépend de leur taille. En effet, si la taille des objets est statistiquement insuffisante pour faire une estimation au sens de maximum de vraisemblance des paramètres du modèle, cette méthode ne pourra les classer correctement.

Le deuxième point important est le critère d'arrêt de l'algorithme. Il dépend aussi de l'image à traiter. Si on augmente le nombre d'itérations de l'algorithme pour assurer la convergence, cela revient à augmenter le temps de calcul de la méthode. Par ailleurs le nombre d'itérations et le seuil à fixer sont liés, ce qui rend leur choix un peu plus difficile.

Commentaires des résultats

Dans le tableau 2 suivant, nous présentons quelques résultats concernant le temps de calcul et le nombre de classes trouvé par l'algorithme SEM pour les différentes sous-images. Ceci pour un nombre de classes majorant de départ $K = 10$, un seuil de 250 et un nombre d'itérations égal à 36 choisi de sorte que le résultat de classification ne change pas, une fois que l'algorithme trouve le nombre de classe. Le seuil est choisi en fonction de la taille des objets dans les images traitées. Si on diminue le seuil, le nombre d'itérations augmente et par conséquent le temps de calcul aussi. C'est ce qu'on a remarqué en traitant à nouveau la sous-image 2 avec un seuil fixé à 200. Le nombre de classes trouvé est alors égal à six et le nombre d'itérations nécessaire pour avoir la stabilité est de 62. Par ailleurs, lorsque le seuil est faible l'algorithme a tendance à prendre comme nombre de classes le nombre majorant K initial.

Tableau 2. - Temps de calcul de l'algorithme SEM.

	temps de calcul (s)	nombre de classes trouvés
sous-image 1	817	3
sous-image 2	849	4
sous-image 3	780	5
sous-image 4	863	7
sous-image 5	812	5

Pour déterminer le nombre optimal des classes pour l'algorithme CMFMAP-NSO, nous avons calculé les deux critères définis précédemment, à savoir respectivement la fonction de compacité et de séparation CS et l'hypervolume V sur l'ensemble des sous-images. Les courbes des figures 1 et 2 montrent les résultats obtenus. Nous remarquons que le critère V fournit le nombre deux comme premier choix optimal sur l'ensemble des sous-images. Ceci est prévisible puisque chaque sous-image est globalement divisée en deux régions R1 et R2. Par ailleurs, le critère de

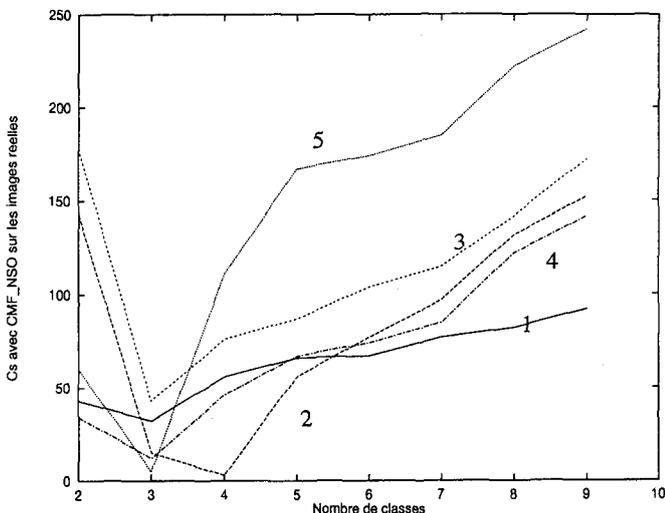


Figure 1. - Courbes de performances des différentes sous-images utilisant le critère CS .

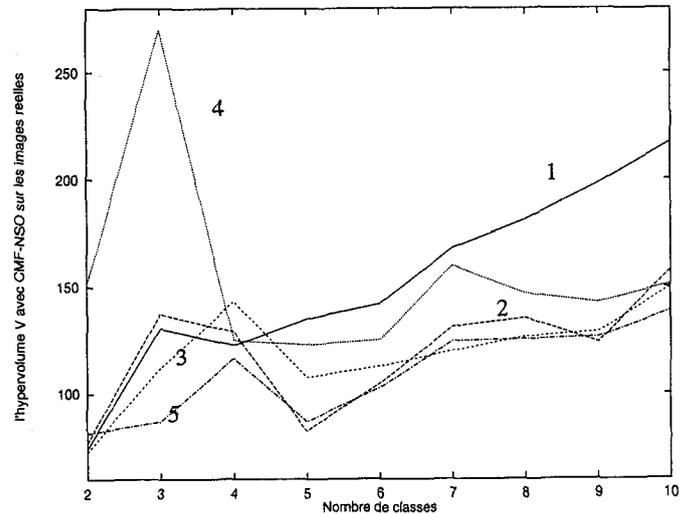


Figure 2. - Courbes de performances des différentes sous-images utilisant le critère V .

l'hypervolume V dans lequel la matrice de covariance est calculée à l'aide des probabilités *a posteriori* utilisant une distance au sens du MAP est bien adapté.

Pour la sous-image 5 et selon le critère CS , le premier choix optimal est trois. Ce mauvais choix est lié à la nature même du critère CS mieux adapté pour l'algorithme CMF-NS.

Les résultats de classification obtenus par l'algorithme CMFMAP-NSO utilisant le critère V sur les différentes sous-images, montrent que celles-ci sont formées essentiellement de deux régions homogènes R1 et R2 ayant un contraste net (voir figure 3). En fait, chaque région est constituée de plusieurs classes de très faible contraste entre elles et le nombre de classes obtenu, qui correspond par ailleurs au nombre de classes d'après la vérité terrain, est celui donné par le critère de performances en deuxième choix optimal (second minimum). Au vu de ces résultats, nous pouvons dire que l'algorithme CMFMAP-NSO a tendance à grouper les classes très proches au sens de la rétrodiffusion en une seule classe. En effet l'affectation d'un vecteur à une classe est faite d'une façon déterministe, par le choix de la classe la plus proche au vecteur X , c'est-à-dire la classe la plus probable au sens du maximum de vraisemblance. Or dans une région constituée de classes de très faible contrastes entre elles, les probabilités *a posteriori* de ces classes sont conditionnées effectivement par les probabilités *a priori* P_k des classes puisque les probabilités conditionnelles $P(X/C_k)$ ont des valeurs très proches. Par conséquent la classe la plus grande fusionne les vecteurs des autres classes au fur et à mesure des itérations dans sa propre classe en formant ainsi une région homogène.

Par contre, l'algorithme SEM affecte un vecteur à une classe par un tirage aléatoire entre les différentes probabilités *a posteriori* d'appartenance du vecteur aux différentes classes. Le tirage aléatoire multinomial choisi, favorise les classes de grandes probabilités. Mais comme la région est formée de zones de très faible contraste, les probabilités *a posteriori* calculées sont proches entre elles et le tirage multinomial se comporte comme un tirage selon une loi uniforme, ce qui donne l'effet d'une classification aléatoire dans la région en question.

Tableau 3. – Résultats de bonne classification des pixels par les 3 méthodes.

	SEM région R1	SEM région R2	CMF-NS région R1	CMF-NS région R2	CMFMAP NSO région R1	CMFMAP NSO région R2
sous-image 1	95%	67%	71%	45%	89,8%	95%
sous-image 2	82%	33%	69%	54%	90,7%	79%
sous-image 3	78%	24%	83%	43%	70%	90,6%
sous-image 4	88%	43%	57%	46,4%	80%	90,6%
sous-image 5	80%	22%	85%	33%	78%	90,2%

L'algorithme CMF-NS se distingue dans son comportement de classifieur par rapport aux algorithmes CMFMAP-NSO et SEM. L'impression aléatoire de la classification obtenue par cet algorithme dans une région composée de plusieurs zones de différence de contrastes faibles est trompeuse. En effet, comme les vecteurs sont groupés dans les différentes classes en fonction de leurs caractéristiques de rétrodiffusion, indépendamment de leurs dispositions spatiales dans l'image, nous pouvons avoir des vecteurs qui sont proches spatialement, mais ayant des caractéristiques de rétrodiffusion assez séparées donc appartenant à des classes différentes. D'où la forme de la classification obtenue. Pour une taille de fenêtre plus grande, l'algorithme CMF-NS aura une tendance d'homogénéisation spatiale des classes proches au sens de la rétrodiffusion. On peut penser que ce raisonnement est aussi valable pour l'algorithme SEM, or il n'en est rien, puisque en plus de l'aspect aléatoire, le nombre des vecteurs diminue et par conséquent, les procédures statistiques d'estimation des paramètres sont moins fiables. Le temps de calcul de l'algorithme CMFMAP-NSO est de l'ordre de 38 mn.

Dans le tableau 3, nous présentons les pourcentages de bonne classification des pixels dans deux régions R1 (sombre) et R2 (claire) de réflectivités distinctes (voir figure 3), obtenus par les trois algorithmes SEM, CMF-NS et CMFMAP-NSO.

Ces résultats montrent que globalement, c'est l'algorithme CMFMAP-NSO, suivi de l'algorithme SEM, qui fournit le meilleur pourcentage de bonne classification des pixels.

3. segmentation contextuelle au sens du MAP

Un des grands intérêts des méthodes markoviennes réside dans les possibilités qu'elles offrent dans l'estimation des paramètres du modèle et dans la segmentation au sens du MAP (Maximum de vraisemblance *A Posteriori*) ou au sens du MPM (Maximum de vraisemblance *a Posteriori* de la densité Marginale).

On note Y le champ des observations, et L le champ des étiquettes. Les étiquettes prennent leurs valeurs dans $\Omega = \{\omega_i\}$. Pour procéder à une analyse statistique de l'image y (une réalisation du champ Y , mesures observées), il faut spécifier la distribution *a posteriori* $P(L = l/Y = y)$, qui exprime la probabilité

d'avoir l'étiquette sous-jacente l conditionnellement à la mesure observée y .

Le critère MAP est le plus communément utilisé. Le champ des étiquettes estimé est tel que :

$$\hat{l} = \max_l P(L = l/Y = y) \quad (25)$$

Par utilisation du théorème de Bayes :

$$P(L = l/Y = y) = \frac{P(Y = y/L = l) \cdot P(L = l)}{P(Y = y)} \quad (26)$$

Comme $P(Y = y)$ est indépendant de l , ceci implique que :

$$\hat{l} = \max_l (P(Y = y/L = l) \cdot P(L = l)) \quad (27)$$

Nous nous intéressons à la segmentation d'image ROS, aussi le champ des étiquettes l sera pour nous la segmentation obtenue. Donc, pour l'obtention de cette segmentation nous avons besoin de réaliser deux étapes distinctes.

La première étape est la modélisation du problème, c'est-à-dire trouver le modèle statistique de $P(Y = y/L = l)$ et le modèle de $P(L = l)$.

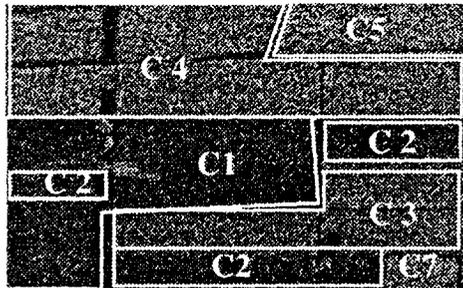
Pour la modélisation de $P(L = l)$, un modèle markovien permettra de prendre en compte les dépendances statistiques spatiales entre sites adjacents; modèle que nous introduisons dans le paragraphe 3.1.1.

Pour la modélisation des observations deux modèles sont possible pour la segmentation des données ROS polarimétriques : le modèle gaussien très utilisé dans la littérature, le modèle de la distribution K mieux adapté aux cibles naturelles (paragraphe 3.1.2 et 3.1.3).

La deuxième étape est l'optimisation. Deux approches seront développées : une approche de type déterministe implémentée par l'algorithme ICM modifié, une seconde, de type stochastique, implémentée par recuit simulé RS. Ces deux approches sont décrites dans la section 3.2.

3.1. modèles statistiques du signal d'image Ros

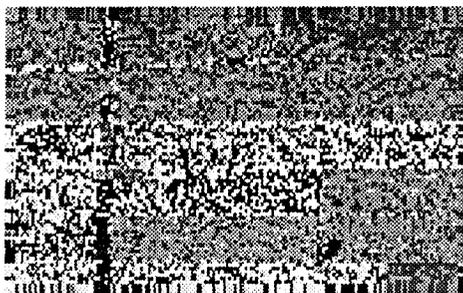
Dans cette section nous présenterons les modèles utilisés pour l'image de segmentation et pour les observations.



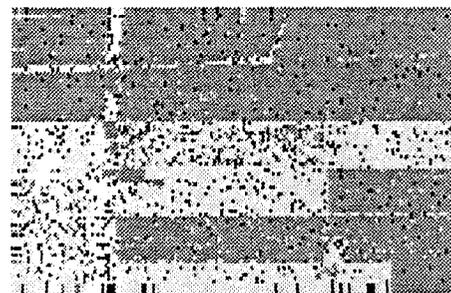
sous-image 2



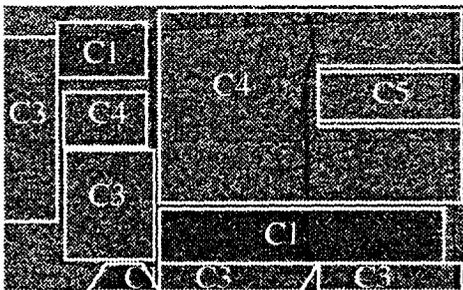
Classification par SEM.



Classification par CMF-NS



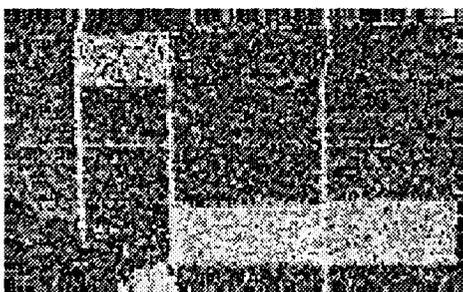
Classification par CMFMAP-NSO



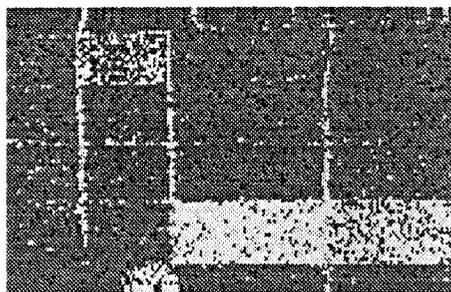
sous-image 5



Classification par SEM.



Classification par CMF-NS



Classification par CMFMAP-NSO

Figure 3. – Résultats de classification des trois algorithmes sur les sous-images deux et cinq.

3.1.1. modèle markovien du champ des classes

L'image du champ des classes est considérée comme une réalisation d'un champ aléatoire de Markov. Ce dernier tient compte de la structure géométrique et de l'interaction locale entre pixels voisins. Ainsi, la probabilité conditionnelle pour qu'un pixel appartienne à une classe donnée connaissant toutes les autres étiquettes de tous les pixels dans l'image, dépend uniquement des étiquettes du voisinage (propriété des champs de Markov). Ceci s'exprime par :

$$P(L_s = l_s / L_{s'} = l_{s'}, s' \in S - \{s\}) \\ = p(L_s = l_s / L_{s'} = l_{s'}, s' \in V_s) \quad (28)$$

où V_s est le voisinage du site s défini par [Gem86].

Avec la condition de positivité suivante :

$$\forall l \in \Omega^S, \quad P(L = l) > 0 \quad (29)$$

alors le champs aléatoire L est un champ de Markov.

Le théorème d'Hammersley-Clifford [Bes74] énonce qu'un champ aléatoire de Markov L suit une loi de probabilité de Gibbs donnée par :

$$P(L = l) = \frac{1}{Z} \exp \{-U_2(l)\} \quad (30)$$

où Z est une constante de normalisation définie par :

$$Z = \sum_{l \in \Omega} \exp \{-U_2(l)\} \quad (31)$$

Ω est l'espace de toutes les configurations possibles. $U_2(l)$ est appelée fonction d'énergie, elle représente une somme de fonctions locales $V_c(l)$, dites fonctions de potentiel :

$$U_2 = \sum_{s \in S} \sum_{c \in C} V_c(l_s)$$

C est un ensemble de cliques engendrées par le voisinage choisi.

A l'échelle d'un site, la fonction d'énergie associée à ce site devient :

$$U_2^s = \sum_{c \in C} V_c(l_s) \quad (33)$$

Pour notre application, nous avons utilisé un système de huit-voisinage engendrant trois types de cliques, parmi lesquels nous n'avons retenu que les cliques d'ordre deux (deux pixels voisins). La fonction de potentiel V_c considérée ici est du type « tout ou rien ». Le modèle dans ces conditions est connu sous le nom du modèle MLL [Gem84] (MultiLevel Logistic), et la fonction potentiel U_2^s prend la forme suivante

$$U_2^s(l_s = l_k / L_V) = -\beta \cdot \left[\sum_{v \in V_s} \delta(l_k - l_v) \right] \quad (34)$$

Dans cette expression, l_k est l'étiquette de la classe k , L_V est l'ensemble des étiquettes voisines, δ représente le symbole de Kronecker et β est une constante positive et supérieure à 1 dont le rôle est de favoriser l'homogénéité spatiale des pixels voisins. En toute rigueur, il faudrait estimer la valeur du paramètre β simultanément au procédé de segmentation. Cependant, dans [Rign91], un test sur les données ROS montre que pour $\beta \in [1, 0; 1, 6]$ la précision de la segmentation est quasi-indépendante de l'image traitée. La valeur $\beta = 1,4$ est adoptée après l'étude de l'influence de ce paramètre sur la segmentation des données ROS polarimétriques que nous possédons.

3.1.2. modèle Gaussien circulaire de données polarimétriques dans une classe

Un modèle de la distribution conditionnelle des données pour une classe (région) connue est présenté. L'étiquette de la classe C_k , $k \in 1, \dots, K$ du site/pixel s est noté l_s , K étant le nombre de classes dans l'image.

Soit X_s le vecteur complexe mesuré par le radar polarimétrique, contenant toutes les informations polarimétriques (les quatre termes de la matrice de rétrodiffusion) relatives au site/pixel s . Pour une image une-vue, X_s s'écrit :

$$X_s = \begin{bmatrix} HH \\ HV \\ VV \end{bmatrix} = \begin{bmatrix} HH_p + jHH_q \\ HV_p + jHV_q \\ VV_p + jVV_q \end{bmatrix} \quad (35)$$

avec : HH l'onde émise et l'onde reçue sont polarisées horizontalement (co-polarisation).

VV : l'onde émise et l'onde reçue sont polarisées verticalement.

HV : l'onde émise est polarisée horizontalement, l'onde reçue est polarisée verticalement (cross-polarisation).

p : désigne les ondes en phase.

q : désigne les ondes en quadrature de phase.

La composante VH ne figure pas dans la relation (35), cela revient à supposer que la cible mesurée obéit au principe de la réciprocité : $VH = HV$. C'est le cas des cibles naturelles [Yuc89].

Conditionnellement à une classe donnée (région homogène), la distribution du vecteur X_s supposé Gaussien circulaire, est donnée par [Kon87] :

$$P(X_s / l_s = l_k) = \frac{1}{\pi^3 |\Sigma_k|} \exp \left(-X_s^\dagger \Sigma_k^{-1} X_s \right) \quad (36)$$

où l_k est l'étiquette de la classe C_k , $\Sigma_k = E[X_s X_s^\dagger]_k$ est la matrice de covariance 3×3 des données polarimétriques de la classe C_k , $|\Sigma_k|$ est son déterminant et \dagger signifie le transposé complexe conjugué.

Ce modèle suppose que les données à l'intérieur d'une classe présentent des statistiques de rétrodiffusion stationnaires, c'est-à-dire, que la matrice de covariance est invariante par translation. L'évaluation de (36) implique un calcul important. Pour réduire ce dernier, on fait généralement l'hypothèse de cibles symétriques en

azimut [Bor87] qui est vraie pour la majorité des cibles naturelles. Cela suppose que l'amplitude complexe HV est non corrélée avec les amplitudes complexes HH et VV . Dans ce cas la matrice de covariance Σ_k se réduit à :

$$\Sigma_k = \sigma_k \begin{bmatrix} 1 & 0 & \rho_k \sqrt{\gamma_k} \\ 0 & \varepsilon_k & 0 \\ \rho_k^* \sqrt{\gamma_k} & 0 & \gamma_k \end{bmatrix} \quad (37)$$

Avec :

$$\begin{aligned} \sigma_k &= E[|HH|^2]_k & \varepsilon_k &= \frac{E[|HV|^2]_k}{E[|HH|^2]_k} & \gamma_k &= \frac{E[|VV|^2]_k}{E[|HH|^2]_k} \\ \rho_k &= \frac{E[HH VV^*]_k}{\sqrt{E[|HH|^2]_k E[|VV|^2]_k}} \end{aligned} \quad (38)$$

* : signifie complexe conjugué.

En remplaçant les éléments par leurs valeurs dans l'expression (36), et après développement cette dernière se met sous la forme suivante [Rig92] :

$$\begin{aligned} P(X_s/l_s = l_k) &= \frac{1}{\pi^3} \exp \left\{ -\frac{|HH|^2}{\sigma_k(1-|\rho_k|^2)} \right. \\ &\quad - \frac{|HV|^2}{\sigma_k \varepsilon_k} - \frac{|VV|^2}{\sigma_k \gamma_k (1-|\rho_k|^2)} + 2 \frac{\Re(HH \cdot VV^* \cdot \rho_k)}{\sigma_k \sqrt{\gamma_k} (1-|\rho_k|^2)} \\ &\quad \left. - \text{Log} [\sigma_k^3 \varepsilon_k \gamma_k (1-|\rho_k|^2)] \right\} \end{aligned} \quad (39)$$

La prise en compte de l'information supplémentaire contenue dans le voisinage d'un site/pixel, conduit naturellement à l'amélioration du procédé d'affectation d'une étiquette à ce dernier. La probabilité conditionnelle du site/pixel donnée par (39) est alors remplacée par la probabilité conjointe du pixel en question et des pixels voisins.

Dans les images ROS des cibles naturelles, les pixels voisins sont peu corrélés. Typiquement la longueur de corrélation est de l'ordre de deux pixels et l'amplitude du coefficient de corrélation est de 0.2 [Wat87], [Ula86]. Dans ce papier, nous considérerons que les pixels de l'image sont non corrélés, donc indépendants du fait qu'ils sont gaussiens. Par ailleurs en considérant l'expression exacte de la probabilité conditionnelle conjointe, il est montré dans [Rig91] que l'amélioration de la segmentation est pratiquement insignifiante en regard de la complexité et du temps de calcul supplémentaire.

A chaque site/pixel s pris comme centre on associe une fenêtre homogène de taille M pixels.

La probabilité conditionnelle du vecteur $Xf_s = [X_1, \dots, X_M]^T$ de chaque fenêtre de l'image est alors donnée par :

$$P(Xf_s/l_s = l_k) = \prod_{i \in M} P(X_i/l_i = l_k) \quad (40)$$

soit encore

$$P(Xf_s/l_s = l_k) = \frac{1}{\pi^{3M}} \exp \{-MU_1^s(X_s/l_s = l_k)\} \quad (41)$$

où U_1^s est la fonction d'énergie définie par :

$$\begin{aligned} U_1^s(Xf_s/l_s = l_k) &= \frac{|HH^M|^2}{\sigma_k(1-|\rho_k|^2)} + \frac{|HV^M|^2}{\sigma_k \varepsilon_k} + \frac{|VV^M|^2}{\sigma_k \gamma_k (1-|\rho_k|^2)} \\ &\quad - 2 \frac{\Re(HH^M \cdot VV^{*M} \cdot \rho_k)}{\sigma_k \sqrt{\gamma_k} (1-|\rho_k|^2)} + \text{Log} [\sigma_k^3 \varepsilon_k \gamma_k (1-|\rho_k|^2)] \end{aligned} \quad (42)$$

avec :

$$\begin{aligned} HH^M VV^{*M} &= \frac{1}{M} \cdot \sum_{i=1}^M HH_i VV_i^* \\ |HH^M|^2 &= \frac{1}{M} \cdot \sum_{i=1}^M |HH_i|^2; & |HV^M|^2 &= \frac{1}{M} \cdot \sum_{i=1}^M |HV_i|^2; \\ |VV^M|^2 &= \frac{1}{M} \cdot \sum_{i=1}^M |VV_i|^2 \end{aligned} \quad (43)$$

L'opération décrite par l'expression (43) est le calcul de la moyenne de l'intensité de M pixels indépendants. La conséquence est la réduction du speckle, mais au prix d'une dégradation de la résolution. Or le filtrage de speckle est généralement une étape préliminaire à la segmentation, destinée à améliorer cette dernière [Saa96]. Cela justifie d'une autre façon l'utilisation de l'expression (41) à la place de l'expression (39).

3.1.3. modèle de distribution-K de données polarimétriques dans une classe

Dans la section 3.1.2 nous avons supposé que la distribution des données polarimétriques dans une classe (région homogène) est gaussienne circulaire. La seule source de la variabilité spatiale du signal image polarimétrique est donc le bruit aléatoire multiplicatif connu sous le nom de speckle. Or plusieurs travaux [Jek76], [Jek80], [Jao84] ont admis qu'il est plus réaliste de considérer que les cibles naturelles (ex : terrain, mer) sont spatialement non homogènes. Pour tenir compte de cette réalité, Novak [Nov89] a proposé un modèle dans lequel le vecteur complexe mesuré par le radar polarimétrique est supposé être le produit d'un vecteur de loi gaussienne circulaire qui représente le speckle et d'une variable aléatoire réelle \sqrt{g} source de la variabilité spatiale :

$$Y_s = X_s \sqrt{g} \quad (44)$$

La distribution de X_s a été donnée dans l'équation (36). Si la distribution de g est du type gamma (d'autres distributions comme celles de Weibul, log-normale sont souvent utilisées), alors la distribution du vecteur mesuré Y_s est la fonction de Bessel modifiée de deuxième espèce d'ordre $3 - \alpha$ appelée aussi la distribution-K généralisée, donnée par :

$$P(Y_s/l_s = l_k) = \frac{2(\bar{g}Y_s^T \Sigma_k^{-1} Y_s)^{(\alpha-3)/2}}{\pi |\Sigma_k| \bar{g}^\alpha \Gamma(\alpha)} K_{3-\alpha} \left(2\sqrt{\frac{Y_s^T \Sigma_k^{-1} Y_s}{\bar{g}}} \right) \quad (45)$$

$$\text{avec} \quad E[g] = \bar{g}\alpha \quad (46)$$

Cette distribution est caractérisée par les deux paramètres \bar{g} et α . Il a été montré [Jao84] que, pour un modèle de terrain composé, le paramètre α seul est suffisant pour caractériser la distribution-K de l'amplitude normalisée de chacune des composantes du vecteur mesuré.

Ce résultat a été étendu [Yuc89] dans le cas multidimensionnel. On obtient la même expression donnée par l'équation (45) en posant dans cette dernière $\bar{g} = 1/\alpha$ soit :

$$P(Y_s/l_s = l_k) = \frac{2\alpha^{(3+\alpha)/2} (Y_s^T \Sigma_k^{-1} Y_s)^{(\alpha-3)/2}}{\pi^3 |\Sigma_k| \Gamma(\alpha)} K_{3-\alpha} \left(2\alpha^{1/2} \sqrt{Y_s^T \Sigma_k^{-1} Y_s} \right) \quad (47)$$

Γ est la fonction gamma, α est un paramètre de forme caractéristique de la distribution-K, permettant d'identifier les différents types de régions (classes) en respectant leur homogénéité. Les régions spatialement homogènes sont caractérisées par un paramètre α élevé, tandis que les régions spatialement non homogènes sont caractérisées par un paramètre α faible. Dans [Elas93], α est exploité comme paramètre d'identification de l'âge des arbres des différentes zones de la forêt Les-Landes.

Remarquons que lorsque $\alpha \rightarrow \infty$ la distribution-K tend vers la distribution gaussienne circulaire donnée par l'équation (36).

Le paramètre α est estimé à partir des moments des intensités normalisées théorique et pratique respectivement données par [Jek80][Yuc89] :

$$I^{(m)} = \frac{m! \Gamma(\alpha + m)}{\alpha^m \Gamma(\alpha)} \quad (48)$$

$$I_{HH}^{(m)} = \frac{E(|HH|^{2m})}{[E(|HH|^2)]^m} \quad I_{HV}^{(m)} = \frac{E(|HV|^{2m})}{[E(|HV|^2)]^m}$$

$$I_{VV}^{(m)} = \frac{E(|VV|^{2m})}{[E(|VV|^2)]^m} \quad (49)$$

Où m est l'ordre du moment. En particulier le moment d'ordre 2 est lié à α par

$$I^{(2)} = 2(1 + 1/\alpha) \quad (50)$$

3.2. optimisation du MAP par ICM et RS

Comme on l'a vu dans la section 3, la segmentation au sens du MAP consiste à prendre comme estimée la valeur \hat{l} qui

maximise la probabilité *a posteriori*. A cet effet, nous avons mis en œuvre deux méthodes (ICM et RS) qui sont assez éloignées dans leur complexité opératoire. Rappelons aussi que l'ICM dépend fortement de l'initialisation à la différence (théoriquement) de RS.

3.2.1. optimisation par algorithme ICM modifié

L'algorithme de relaxation déterministe ICM, calcule de façon itérative l'expression suivante

$$P(Y_s/l_s = l_k) \times p(L_s = l_s / L_{s'} = l_{s'}, s' \in V_s) \quad (51)$$

pour chaque pixel et pour chaque classe de l'image. Il affecte au pixel courant (traité) l'étiquette qui maximise l'expression (51). Cette opération est répétée jusqu'à la stabilité de la segmentation. Cette méthode simple et rapide est convergente, puisque à chaque itération on diminue toujours l'énergie, mais tout comme les méthodes de gradient, elle converge presque sûrement vers un optimum local.

Afin d'accélérer la convergence de l'ICM, nous avons utilisé une pile d'instabilité qui contient tous les pixels instables. Alors nous ne traitons que les pixels instables, c'est-à-dire ceux qui ont une probabilité de Gibbs inférieure à un seuil μ (nous avons pris 0.97). En fait c'est un test d'homogénéité permettant de savoir si le pixel courant est homogène avec ses voisins ou non. Tous les pixels instables donc hétérogènes sont traités ainsi que leurs voisins immédiats (8-voisinage). Par ailleurs l'estimation des paramètres se fait à chaque itération permettant ainsi une meilleure segmentation.

Description de l'algorithme ICM modifié :

Etant données les données polarimétriques (les trois images HH , HV , VV) ainsi qu'une segmentation initiale (pré-segmentation) de l'image, l'algorithme ICM modifié est alors décrit par les étapes suivantes :

On itère jusqu'à la stabilité de la segmentation les opérations suivantes :

1) Pour chaque pixel, dans une fenêtre de taille fixe (3*3), on calcule la probabilité de Gibbs pour toutes les classes de l'image, et on cherche la probabilité de Gibbs maximale P_{Gibbs}

Si $P_{Gibbs} > \mu$, alors le site courant est stable on passe au site suivant,

Sinon, on réalise les étapes suivantes :

a) Pour chaque classe :

– Estimation des paramètres.

– Calcul de l'expression (51).

b) Recherche de la classe qui maximise (51).

c) Affectation de l'étiquette de la classe trouvée dans b au site courant.

2) Test de stabilité de la segmentation obtenue, si la stabilité est assurée (le nombre de sites qui changent de classe est inférieur à 1%) alors, l'algorithme s'arrête, sinon on revient à l'étape 1.

3.2.2. optimisation par recuit simulé (RS)

L'objectif de cette méthode est de garantir l'obtention d'un minimum global, sans risquer de rester bloqué dans un minimum local rencontré éventuellement.

Le point clé du recuit simulé est qu'il autorise (à température élevée) des transitions vers des configurations d'énergie plus forte, pour pouvoir se dégager des minima locaux.

Le cheminement permettant d'atteindre le minimum global [Gem84] est le même que celui de l'ICM, à la différence qu'ici, nous introduisons la température et le choix de décroissance de celle-ci, ainsi que le choix stochastique permettant l'autorisation des transitions montantes augmentant l'énergie.

Dans la simulation réalisée, le schéma de réduction de la température choisi est : $T(i) = \tau^i \cdot T_0$ avec $\tau \in [0,98, 0,9999]$, où τ est le facteur de réduction de la température et $T_0 = 293^\circ K$.

4. résultats expérimentaux obtenus par ICM et RS

Dans cette partie nous présentons les résultats de segmentation obtenus par l'algorithme ICM modifié (pour les deux modèles adoptés, distribution Gaussienne : ICM_Gauss, distribution_K : ICM_K) et le recuit simulé sur les sous-images deux et cinq. L'initialisation de l'ICM est faite à partir des images classifiées par l'algorithmes CMFMAP-NSO. Les résultats en pourcentage de bonne classification des pixels sont donnés dans les tableaux 4, 5 et 6. Nous remarquons que l'algorithme ICM_K fournit un meilleur résultat que l'algorithme ICM_Gauss. Cela était prévisible puisque plusieurs travaux ont déjà montré que, pour ce type d'images, les données mesurées conditionnellement à une classe suivent bien la distribution_K. Le RS augmente de 8 à 10% le taux de bonne classification des pixels, pour un temps de calcul très supérieur à celui de l'ICM. Nous pouvons constater aussi que globalement l'algorithme ICM améliore les résultats du CMFMAP-NSO. Cette amélioration se traduit par un gain approximatif de 10% à 20% en termes de bonne classification des pixels et par l'obtention d'une qualité visuelle meilleure des images segmentées (voir figure 4).

La construction des régions se base sur une propriété purement radiométrique (dans le sens de la rétrodiffusion). Pour une région donnée (R1 par exemple) les pixels qui sont mal classés se retrouvent pour une partie dans l'autre région (R2) et pour la partie restante éparpillés aléatoirement dans l'image de segmentation (voir figure 4).

Dans le tableau 7 nous avons indiqué le temps de calcul de segmentation des différents algorithmes.

Tableau 4. – Pourcentage de bonne classification des pixels obtenu par l'algorithme ICM_Gauss.

ICM_Gauss	région R1	région R2
sous-image 2	86%	81%
sous-image 5	80%	94,2%

Tableau 5. – Pourcentage de bonne classification des pixels obtenu par l'algorithme ICM_K.

ICM_K	région R1	région R2
sous-image 2	100%	88%
sous-image 5	98%	91%

Tableau 6. – Pourcentage de bonne classification des pixels obtenu par le RS.

RS_Gauss	région R1	région R2
sous-image 2	100%	90%
sous-image 5	99%	97%

Tableau 7. – Temps de calcul en seconde des différents algorithmes sur RS6000.

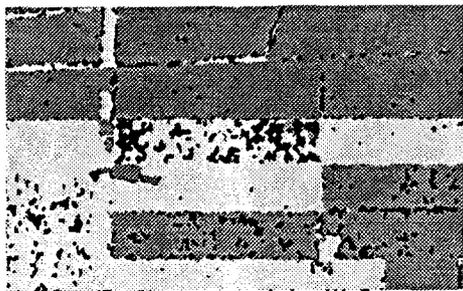
Temps de calcul (s)	ICM_Gauss	ICM_K	RS_Gauss
sous-image 2	150	850	4200
sous-image 5	146,37	722	400

5. conclusion

Dans ce papier nous avons traité le problème du groupement multidimensionnel au sens de la rétrodiffusion et de la segmentation vectorielle markovienne des données radar à visée latérale et à ouverture synthétique. Le groupement multidimensionnel non supervisé a été abordé sous deux aspects :

- l'estimation par la méthode du maximum de vraisemblance *a posteriori*, implémentée par l'algorithme SEM,
- la classification par affectation des données aux classes les plus proches au sens de la distance, implémentée soit par l'algorithme CMF-NS utilisant une distance euclidienne, soit par l'algorithme CMFMAP-NSO utilisant une distance exponentielle au sens du maximum de vraisemblance *a posteriori*.

Globalement, c'est l'algorithme CMFMAP-NSO qui donne le meilleur pourcentage de bonne classification des pixels, la meilleure qualité visuelle des images segmentées et la meilleure adaptativité aux formes elliptiques des classes. Cependant le temps de calcul mis en œuvre est plus important, puisque à chaque vecteur est associée une probabilité floue et le calcul des paramètres



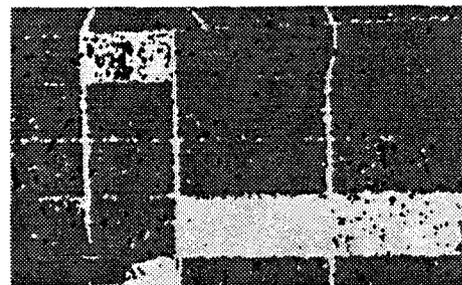
Segmentation par ICM_Gauss



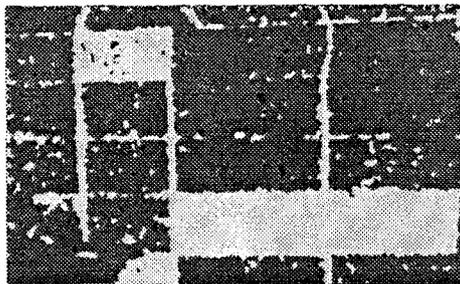
Segmentation par ICM_K



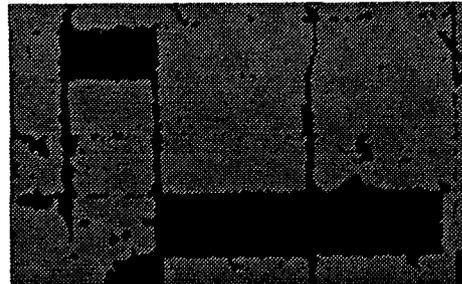
Segmentation par RS_Gauss



Segmentation par ICM_Gauss



Segmentation par ICM_K



Segmentation par RS_Gauss

Figure 4. – Résultats de segmentation des différents algorithmes sur les sous-images deux et cinq.

de chaque classe tient compte de tous les vecteurs de l'échantillon. Cet algorithme a la propriété de grouper les classes ayant des différences de contrastes très faibles en une seule classe homogène. L'algorithme SEM par contre, a la propriété d'affecter les vecteurs de façon aléatoire pour ces types de classes. Par ailleurs, dans l'algorithme SEM, le choix fixé *a priori* du seuil et du nombre d'itérations qui dépendent de l'image à traiter, rend l'algorithme non complètement automatique.

Nous avons constaté l'efficacité d'une approche markovienne dans la segmentation vectorielle utilisant l'estimateur MAP implémenté par l'ICM et le RS. Ces derniers affinent et améliorent la segmentation déjà obtenue par un des algorithmes de groupement multidimensionnel. Les résultats obtenus montrent aussi l'intérêt de prendre comme distribution conditionnelle *a priori* la distribution_K, qui décrit convenablement les amplitudes des

échos radar des scènes naturelles (Forêt Les-Landes) par rapport à la distribution gaussienne souvent utilisée dans la littérature.

ANNEXE

Démonstration de la relation (6) :

Pour estimer \hat{P}_k , nous dérivons l'expression (5) par rapport à P_k

$$\frac{\partial}{\partial P_k} \left[\sum_{n=1}^N \text{Log} [P(X_n)] - \lambda \left(\sum_{k=1}^K P_k - 1 \right) \right] = 0$$

$$P(X_n) = \sum_{k=1}^K P_k \psi(X_n, \Theta_k)$$

Remplaçons $P(X_n)$ par sa valeur, nous obtenons :

$$\frac{\partial}{\partial P_k} \left[\sum_{n=1}^N \text{Log} \left[\sum_{k=1}^K P_k \psi(X_n, \Theta_k) \right] - \lambda \left(\sum_{k=1}^K P_k - 1 \right) \right] = 0$$

Soit :

$$\sum_{n=1}^N \frac{\psi(X_n, \Theta_k)}{\sum_{k=1}^K P_k \psi(X_n, \Theta_k)} - \lambda = 0$$

Multiplions par \hat{P}_k et utilisons la relation (3), nous obtenons la relation (6)

Démonstration de la relation (7) :

Annulons le gradient partiel par rapport à Θ_k de l'expression (5), nous obtenons :

$$\frac{\partial}{\partial \Theta_k} \left[\sum_{n=1}^N \text{Log} \left(\sum_{k=1}^K P_k \psi(X_n, \Theta_k) \right) \right] = 0$$

Soit encore :

$$\sum_{n=1}^N \frac{P_k \frac{\partial}{\partial \Theta_k} \psi(X_n, \Theta_k)}{\sum_{k=1}^K P_k \psi(X_n, \Theta_k)} = 0$$

Multiplions le numérateur et le dénominateur par $\psi(X_n, \Theta_k)$ et utilisons la relation (3), nous obtenons :

$$\sum_{n=1}^N P(C_k/X_n) \frac{\frac{\partial}{\partial \Theta_k} \psi(X_n, \Theta_k)}{\psi(X_n, \Theta_k)} = 0$$

Cette dernière relation est bien identique à la relation (7).

BIBLIOGRAPHIE

[Bes74] Besag J., «Spatial interaction and the statistical analysis of lattice system», *J. Royal Stat. Soc. Serie B*, Vol. 36, pp. 192-236, 1974.
 [Bes86] Besag J., «Analysis of dirty pictures», *JRSS B*, 1986.
 [Bez81] Bezdek J. C., «Pattern Recognition with Fuzzy Objective Function Algorithms», *Plenum*, New York, 1981.
 [Bor87] Borgeaud M., Shin R. T., and Kong J. A., «Theoretical models for polarimetric radar clutter», *J. Electron. Waves Appl.*, Vol. 1, pp. 67-86, 1987.
 [Cel86] CELEUX G., DIEBOLT J., «L'algorithme SEM : un algorithme d'apprentissage probabiliste pour la reconnaissance de mélange de densités», *Revue de statistiques appliquées* Vol. 34, n°2, pp 35-52, 1986.
 [Cel89] Celeux G., Diday E., Govaert G., Lechevallier Y. et Ralamondrainy H., «Classification automatique des données», *Dunod* 1989.

[Che85] Chellappa and Chatterjee, «Classification de texture using Gaussian Markov Random fields», *EEE Trans Acoust. Speech& Signal Processing*, pp. 959-963, Aug. 85.
 [Day69] Day N. E., «Estimating the components of mixture of normal distributions», *Biometrika*, Vol. 56, pp. 463-474, 1969.
 [Des93] Descombes X. «Champs markoviens en analyse d'images» thèse de l'ENST, Paris 1993.
 [Dre87] Derin H. and Elliot H., «Modeling and segmenting of noisy and textured images using Gibbs random field»*EEE Trans. PAMI*, Vol. 9, N°1, pp. 39-55, Jan. 1987.
 [Dre90] Derin H., Kelly, Vezina & Labit C., «Modeling and segmentation of speckled images using complex data» *IEEE Trans Geosci. Remote sens.* Vol. 28, janv 90.
 [Dub89] Dubes R. C. and Jain A. K., «Random field models in images analysis» *J. Applied Statistics*, Vol. 6, N°2, pp. 131-163, 1989.
 [Duv91] Duvaut P., «Traitement du signal, concepts et applications», pp. 582-586, Hermes 1991.
 [Elas93] El Assad S., Belhadj Z. and Saillard J., «Statistical Analysis of forest clutter Measured with polarimetric SAR», *Rapport interne*, Janvier 1993.
 [Fan86] Fan Z. and Cohen F. S., «unsupervised texture image segmentation» In proc. Int. Congress Intelligent Autonomous Systeme, Amsterdam, Holland, Dec. 1986.
 [Fir94] Firmin C. and Hamad D., «Gaussian based neural networks applied to pattern classification and multivariate probability density estimation. *IEEE Int. Conf. On Networks. IEEE WCCI*, Vol. V., June 1994, pp 2985-2990.
 [Gath89] Gath I. and Geva A. B., «Unsupervised optimal fuzzy clustering», *IEEE Trans. Anal. Machine Intell.*, Vol. 11, n°7, pp. 773-781, July 1989.
 [Gem84] Geman and Geman, «Stochastic relaxation, Gibbs distribution and the bayesian restoration of images», *IEEE Trans. Pattern Anal. and Machine Intell.* Vol. PAMI.6, pp. 721-741, Nov 84.
 [Gem86] Geman D. and Graffigne C., «Markov random fields image models and their application to computer vision», in proe. int. Congress of Mathematicians, pp. 609-628, 1986.
 [Gem90] Geman D., Geman S., Graffigne C., and Pong D., «Boundary detection by contrained optimisation»*IEEE Trans. PAMI*, Vol. 12, N°7, pp. 609-628, July 1990.
 [Gra86] Graffigne C., «Experiments in texture analysis and segmentation», Phd thesis, Brown University, 1986.
 [Har79] Haralick R. M. «Statistical and structural approaches to textures», *Porc. EEE*, Vol. 67, No. 5, pp. 786-804, May 1979.
 [Her92] Hermann and Kahazenic, «Classification of multispectral Remote sensing data using a Back Propagation Neural Networks», *EEE Trans Geosci. Remote sens.* vol. 30 N°1, Junary 92.
 [Jao84] Jao J. K., «Amplitude distribution of composite terrain radar clutter and the Kdistribution», *IEEE Trans. Antennas Propagat.* Vol. AP-32, n°10, pp. 1049-1062, 1984.
 [Jek76] Jekeman E., and Pusey P. N., «A model for non-rayleigh sea echo», *IEEE Trans. Antennas Propagat.*, Vol. AP-24, n°6, 1976 pp. 806-814.
 [Jek80] Jekeman E., «On the statistic of K-distribution noise», *J. Phys. A : Math. Gen.*, Vol. 13, pp. 31-48, 1980.
 [Kel88] Kelly, Drein and Harrit, «Adaptative segmentation of speckled images using a hierarchical random field model», *IEEE Trans Acoust. Speech & Signal Processing* vol 36, pp. 1628-1641, oct 88.
 [Kho89] Khotanzad A. and Chen J.Y., «Unsupervised segmentation of textured images by edges detection in multidimensional features», *IEEE Trans. PAMI*, Vol. 1, N°4, pp. 414-421, 1989.

[Kon87] Kong J. A., Swartz A. A., Yuch H. A., Novak L. M. and Shin R.T., «Identification of terrain cover using the optimum polarimetric classifier», *J. Electron. Waves Appl.*, Vol. 2., pp. 171-194, Feb. 1987.

[Kon88] Kong J. A., Swartz A. A., Yuch H. A., Novak L. M. and R.T. Shin, «Identification of terrain cover using the optimum polarimetric classifier», *J. Electron. Waves appl.*, vol.2, pp. 171-194, Feb. 88.

[Lak89] Lakshmanan S. and Drein H., «Simultaneous parametre estimation and segmentation of Gibbs random fields using simulated annealing», *EEE Trans. PAMI* Vol. 11, N°8, pp. 799-813, 1989.

[Lee81] Lee J. S., «Speckle analysis and smoothing of synthetic aperture radar images», *Computer Graphics and Image Processing* 17, pp. 24-32, 1981.

[Lim89] Lim, Swartz, Yeuh, Kong, Shin and Van Zyl, «Classification of Earth terrain using polarimetric SAR images», *J. Geophy. Res.*, vol. 94, pp. 7049-7057, June 89.

[Lin80] Linde Y., Buzo A. and Gray R.M., «An Algorithrn for Vector Quantizer Design», *IEEE Trans. Commun.* COM-28, pp. 84-95, Jan. 1980.

[Man91] Manjunath and Chellappa R., «Unsupervised texture segmentation using markov random fields» *EEE Trans. PAMI*, Vol. 13, pp. 478-482, 1991.

[Mar94] Marhic A. N., «Modélisation stochastique et classification des images satellite par méthodes statistiques», Thèse Univ. Rennes I

[Mas93] Masson. P. and Pieczynski. W., «SEM Algorithm and Unsupervised Statistical Segmentation of Satellite Images » *IEEETEE Trans. on Geo. and Rem. Sen* Vol. 31, N°3, pp.618-633, May 1993.

[Nov89] Novak L. M., Sechtin M. B., and Cardullo M. J., «Studies on target Detection Algorithms Which Use Polarimetric Radar Data», *IEEE Trans. Aerospace Electronic Systems*, Vol. AES-25, n°2, pp. 150-169, 1989.

[Ren84] Render A. R. and Walker H. A., «Mixture densities, maximum likelihood and the EM algorithm», *SIAM Review*, Vol. 26, n°2, April 1984.

[Rig91] Rignot E. and Chellappa R., «Segmentation of synthetic aperture radar complex data», *J. Opt. Soc. Amer.*, A, pp. 1499-1509, Sept. 1991.

[Rign91] Rignot E., «Segmentation of multifrequency and polarimetric synthetic aperture radar data», Ph. D. dissertation, *Dep. of EE-systems, Univ of Southern California, Los Angeles, CA*, May 1991.

[Rig92] Rignot E. and Chellappa R., «Segmentation of polarimetric synthetic aperture radar data», *IEEE Trans. On Image Processing*, Vol. 1, n°3, pp. 281-296, July 1992.

[Saa96] Saad A., El Assad S. et Barba D., «Speckle Filtering in SAR Images by contrast modification, comparaison with a large class of filters», *Annales des Télécommunications*, Vol. 51, pp. 233-244, n°5-6, Juin 1996.

[Ula86] Ulaby F. T., Moore R. K. and Fung A. K., «Microwave Remote Sensing : Active and Passive», *From Theorie to Applicatton*. Vol. 10, Dedham, MA : Aetech House, 1986, pp. 1907-1935.

[Ula90] Ulaby F. T. and Elachi C., «Radar polarimetry for Geoscience applications», *Norwood, MA : Artech House*, 1990.

[Van89] Van Zyl J. J. et al., «Unsupervised classification of scattering behavior using radar polarimetry data», *IEEE Trans. Geosci. Rentote Sens.*, Vol. 27, N°1, pp. 3644, Janvier 1989. Wat87]

[Wat 87] Watts S. and Ward K. D., «Spatial Correlation in K-Distribution Sea Clutter», *IEE Proceedings*, Vol. 134, n°6, pp. 526-532, October 1987.

[Won92] Won C. S. and Drein H., «Unsupervised segmentation of noisy and textured images using markov random fields», *CVGIP*, Vol. 54, N°4, pp. 308-328, July 1992.

[Xie91] Xie X. L. and Beni G., «A Validity Measure for Fuzzy Clustering», *IEEE trans. PAMI*, Vol. 13, n°8, pp. 841-847, August 1991.

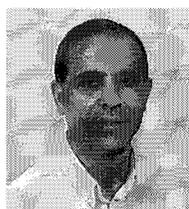
[Yuc89] Yuch H. A., Kong J. A., Jao J. K., Shin R.T. and Novak L. M., «K-distribution and polarimetric terrain radar clutter», *J. Electron. Waves Appl.*, Vol. 1, pp. 747-768, Aug. 1989.

[Zha88] Zhang J. and Modestino J.W., «Unsupervised images segmentation using a gaussian model», In proc. CISS 1988, pages 7-10, Princeton, NJ, 1988.

Manuscrit reçu le

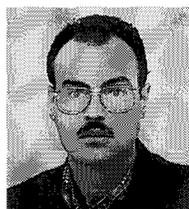
LES AUTEURS

Safwan EL ASSAD



Safwan El Assad est né au Liban en 1957. Il est Docteur de l'Université de Lille I en 1987 dans le domaine de la CEM. Depuis 1988, il est Maître de Conférences à l'IRESTE (Institut de Recherche et d'Enseignement Supérieur aux Techniques de l'Electronique - Ecole d'Ingénieurs de l'Université de Nantes). De 1988 à 1997 il anime un thème transversal sur «l'Imagerie Radar» au sein du laboratoire SEI. Ses activités de recherche portent sur les sujets suivants : Télédétection micro-ondes, Méthodes de reconstruction de l'image SAR et l'image interférométrique, Modélisation statistiques des échos radar, Filtrages scalaire et vectoriel des images SAR, Classification et segmentation markovienne des images SAR polarimétriques, Compression des signaux radar en vue de la transmission.

Ali SAAD



Ali Saad a obtenu sa Thèse de Doctorat en Septembre 1996 à l'IRESTE, Université de Nantes, sur le «Filtrage et segmentation d'images radar polarimétriques». Depuis Octobre 1996, il est en Post-Doctoral au «Baylor College of Medecine» Houston U.S.A où il effectue des recherches sur le filtrage et classification des images médicales 3-D.

Dominique BARBA

Dominique Barba est né en France en 1944. Il est titulaire d'un doctorat de 3ème cycle en télécommunications de l'Université de Rennes I obtenu en 1969 et d'un doctorat es Sciences Mathématiques (mention informatique) de l'Université de Paris VI sur un sujet lié au traitement numérique des images avec critère psychovisuel de qualité. Il est actuellement Professeur à l'IRESTE (Université de Nantes) et responsable d'une équipe de recherche en Analyse des images et Vidéocommunications au sein du laboratoire Systèmes Electroniques et Informatiques (SEI). Ses centres d'intérêt concement le traitement numérique des images fixes et animées, spécialement en vue de l'analyse ou du codage des images.