

THESE

présentée à

L'UNIVERSITE DES SCIENCES ET TECHNOLOGIES DE LILLE

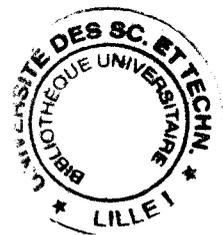
pour obtenir le titre de

DOCTEUR DE L'UNIVERSITE

Spécialité: Electronique

par

Abdelwaheb MARZOUKI



SEGMENTATION STATISTIQUE D'IMAGES RADAR

Soutenu le vendredi 8 Novembre 1996 devant la commission d'examen composée de:

Président : M Jean Paul DUBUS, Professeur à l'Université de Lille I

Rapporteurs : M Alain HILLION, Professeur à l'ENSTB

M Jacques LEMAIRE, Professeur à l'Université de Nice

Examineurs: M^{me} Agnès BEGUE, Ingénieur de recherche au CIRAD

M Armand LOPES, Maître de Conférences à l'Université Paul Sabatier de Toulouse

M Wojciech PIECZYNSKI, Professeur à l'Institut National des Télécommunications

M Yves DELIGNON, Maître de Conférences à l'Ecole Nouvelle d'Ingénieurs en Communication



*A la mémoire de mon père,
à ma mère,
à mes frères et mes soeurs*

La vraie logique de ce monde se trouve dans le calcul des probabilités

James Clerk Maxwell

REMERCIEMENTS

Cette thèse n'aurait pu aboutir sans le parfait encadrement et la constante présence de Monsieur Yves Delignon, Maître de Conférences à l'ENIC, à qui j'adresse mes vifs remerciements et le témoignage de ma reconnaissance.

Mes remerciements s'adressent, également à mon directeur de thèse, Monsieur Wojciech Pieczynski Professeur à l'INT pour la confiance qu'il m'a accordée et pour ses riches conseils.

Je remercie Monsieur Jean Paul Dubus, Professeur à l'Université de Lille I de m'avoir fait l'honneur de présider le Jury de la thèse et de m'avoir enseigné le traitement du signal au cours des années qui ont précédé cette thèse.

Je témoigne ma reconnaissance à Monsieur Alain Hillion, Professeur à l'ENSTB pour avoir accepté d'être rapporteur de ma thèse et pour l'aide qu'il m'a apportée pour mener à bien la rédaction de ce rapport.

Je remercie vivement Monsieur Jacques Lemaire, Professeur à l'Université de Nice de bien vouloir être rapporteur de ma thèse.

Je remercie Madame Agnès Begue Ingénieur au CIRAD de m'avoir fourni les données qui m'ont permis de valider ce travail et d'avoir accepté de faire partie du jury de cette thèse.

Je remercie Monsieur Armand Lopes, Maître de conférences à l'Université Paul Sabatier de Toulouse, pour avoir accepté de faire partie du jury de ma thèse.

Je remercie Monsieur Jean Claude Damien, directeur de l'ENIC de la confiance qu'il m'a accordée en m'acceptant comme chercheur à l'ENIC.

Je tiens à remercier Monsieur Faouzi Ghorbel pour la confiance qu'il m'a accordée en m'acceptant comme membre du GRIF, pour ses fructueuses remarques et pour ses riches conseils.

Je tiens à remercier, également Monsieur Philippe Clabaut pour l'aide qu'il m'a apporté pour l'interprétation des résultats obtenus au cours de cette thèse.

Je remercie Monsieur Patrick Kennis et Monsieur Yves Druelle, Chefs du département Electronique de l'ENIC pour leur gentillesse et pour leurs riches conseils.

J'adresse un grand Merci à Ahmed Mokadem, David Low, Lamia Ben Youssef, Mourad Zribi, Néjib Haouech et Sabine Toustain de m'avoir aidé à finir ce rapport et de leurs vives remarques.

J'adresse un salut amical à tout le personnel de l'ENIC et particulièrement à tous les membres du GRIF.

Un immense merci à la famille en Tunisie pour la pensée qu'elle m'a accordé au cours de mon séjour en France.

Table des matières

INTRODUCTION GENERALE	3
I. MODELISATION STATISTIQUE DES IMAGES RADAR	7
I.1 LES RADARS IMAGEURS.....	8
<i>I.1.a Principe du radar à visée latérale.....</i>	<i>9</i>
<i>I.1.b Caractéristiques des radars à visée latérale.....</i>	<i>11</i>
I.2 MODELE HIERARCHIQUE GENERALISE	13
I.3 MODELISATION DU CHAMP DE CLASSES	14
<i>I.3.a Modélisation locale du champ de classes</i>	<i>15</i>
<i>I.3.b Modélisation globale du champ de classes</i>	<i>15</i>
I.4 MODELISATION DU PROCESSUS DE MESURE	22
<i>I.4.a Système de distributions de Pearson</i>	<i>23</i>
<i>I.4.b Estimation des paramètres du système de Pearson.....</i>	<i>30</i>
I.5 SEGMENTATION BAYESIENNE	38
I.6 CONCLUSION.....	40
II. SEGMENTATION LOCALE DES IMAGES MONOSPECTRALES	43
II.1 ESTIMATION DES PARAMETRES DANS LE MODELE LOCAL STATIONNAIRE.....	44
<i>II.1.a Méthode du maximum de la vraisemblance</i>	<i>45</i>
<i>II.1.b Méthode de l'espérance conditionnelle.....</i>	<i>51</i>
<i>II.1.c Méthode d'estimation par prélèvement aléatoire et contrôlé.....</i>	<i>53</i>
II.2 ESTIMATION DES PARAMETRES DANS LE MODELE NON STATIONNAIRE	57
<i>II.2.a Contexte général</i>	<i>57</i>
<i>II.2.b Procédures d'estimation itératives.....</i>	<i>58</i>

IV.4 CONCLUSION	112
------------------------------	------------

V. SEGMENTATION GLOBALE DES IMAGES MULTISPECTRALES..... 115

V.1 ESTIMATION DES PARAMETRES.....	115
---	------------

<i>V.1.a Algorithme EM gibbsien.....</i>	<i>116</i>
--	------------

<i>V.1.b Algorithme ICE gibbsien</i>	<i>119</i>
--	------------

V.2 ANALYSE DES RESULTATS.....	120
---------------------------------------	------------

<i>V.2.a Segmentation de l'image de synthèse</i>	<i>120</i>
--	------------

<i>V.2.b Segmentation de l'image SIR-C/X-SAR</i>	<i>122</i>
--	------------

<i>V.2.c Segmentation de la zone forestière</i>	<i>123</i>
---	------------

V.3 CONCLUSION.....	125
----------------------------	------------

CONCLUSION GENERALE..... 129

ANNEXES

ANNEXE 1: EQUATIONS DE VRAISEMBLANCE DANS LE CAS MONOSPECTRAL.....	133
---	------------

ANNEXE 2: INITIALISATION DES PARAMETRES DANS LES METHODES LOCALES	135
--	------------

ANNEXE 3: EQUATIONS DE VRAISEMBLANCE DANS LE CAS MULTISPECTRAL	137
---	------------

REFERENCES BIBLIOGRAPHIQUES..... 141

NOTATIONS, TABLEAUX ET FIGURES

LISTE DES NOTATIONS.....	149
---------------------------------	------------

LISTE DES TABLEAUX	151
---------------------------------	------------

LISTE DES FIGURES.....	153
-------------------------------	------------

INTRODUCTIION GENERALE

INTRODUCTION GENERALE

La segmentation des images radar se présente comme une étape indispensable pour un certain nombre d'applications telles que la discrimination des cultures, le suivi du processus de déforestation ou la cartographie de la végétation dans les zones alluviales.

Ce traitement a pris une forte ampleur de nos jours, en raison de l'opportunité qu'offrent les satellites d'observation de la terre, qui utilisent des radars à ouverture synthétique. L'intérêt des images acquises par ces capteurs est double. D'une part le radar est un système actif, dans la mesure où il fournit sa propre source d'illumination. Il opère, de ce fait, de jour comme de nuit et indépendamment des conditions météorologiques. D'autre part, la technique de synthèse d'ouverture utilisée permet d'obtenir une résolution spatiale d'environ 20 m qui est comparable à celle obtenue par des capteurs optiques. Cependant, et contrairement aux images issues de capteurs optiques, la scène satellitaire est difficilement interprétable vu le fort niveau du speckle, qui donne un aspect granulaire aux différentes textures homogènes observées.

Dans le cas des images RADAR, plusieurs méthodes de segmentation peuvent être employées. Parmi ces méthodes, nous distinguons les méthodes géométriques qui utilisent des notions de distances et qui ne font pas intervenir des hypothèses probabilistes et les méthodes statistiques qui sont basées sur une modélisation stochastique de l'image. L'avantage des méthodes statistiques est qu'elles permettent de tenir compte de la nature stochastique du bruit par une modélisation adéquate du processus de rétrodiffusion. Le caractère non supervisé des méthodes statistiques permet en plus d'ignorer le nombre de classes et les paramètres associés.

Soulignons que l'étude que nous proposons s'intègre dans une suite de travaux réalisés par les chercheurs de l'INT et de l'ENST de Bretagne. Nous proposons une amélioration aux modèles adoptés habituellement pour décrire le processus de rétrodiffusion. Les recherches effectuées dans ce domaine ayant conclu à une forme variable de la distribution qui modélise ce processus, nous proposons de tenir compte de cette variabilité de forme en modélisant la rétrodiffusion relative à chaque texture par les distributions unimodales du système de Pearson.

CHAPITRE I

MODELISATION STATISQUE DES IMAGES RADAR

I. Modélisation statistique des images radar

La segmentation statistique des images a pris une forte ampleur depuis une dizaine d'année et particulièrement depuis l'usage des champs markoviens pour la modélisation des scènes naturelles homogènes. Le but d'une telle opération est de procéder à la répartition de l'image en plusieurs régions homogènes (ou classes) en minimisant une probabilité d'erreur de classification. Les méthodes de segmentation que nous employons sont dites non supervisées car elles ne demandent, ni le nombre de classes, ni des informations ou des données complémentaires sur la nature du terrain. Ces méthodes sont issues d'une modélisation de l'image qui tient compte de la richesse des scènes naturelles imagées par un radar, d'une part, et de la répartition des différentes textures et variabilité de chacune d'elles, d'autre part. Le modèle retenu pour une telle description de l'image est le modèle hiérarchique généralisé [HH87].

Par ailleurs, il est bien établi que les images issues de la rétrodiffusion présentent des distributions de formes variables avec la nature de la texture [DGH90, DKV90, Del93, LLN90, LNT90, LNT93].

L'étude de la distribution du niveau de gris au sein d'une même région homogène a montré le caractère unimodal de cette distribution et la variabilité de sa forme en fonction de la texture. Cette étude a d'autre part montré qu'un ajustement de l'histogramme relatif aux textures homogènes d'une image radar est mieux réalisé par les distributions du système de distributions de Pearson, que par des distributions dont la forme est fixée telle que la loi normale ou la loi gamma [Del93, JK70, MDP93, MDQ93, Quel93]. L'originalité de notre étude réside dans l'adoption des distributions de ce système pour modéliser la rétrodiffusion au sein des scènes naturelles homogènes et d'utiliser cette modélisation pour réaliser une segmentation de l'image.

L'organisation de ce chapitre est la suivante:

Dans un premier paragraphe, nous donnons un aperçu des différents radars utilisés pour la collection des images traitées au cours de notre étude et un bref sommaire sur les caractéristiques techniques des radar imageurs.

Dans le paragraphe 2 nous rappelons les bases du modèle hiérarchique généralisé.

brièvement les principes des radars à antenne latérale afin de définir les caractéristiques des cellules de résolution et la quantité mesurée par le niveau de gris.

I.1.a Principe du radar à visée latérale

Le radar embarqué se déplaçant à une vitesse V dans la direction azimutale émet une impulsion vers la terre. Lorsque l'onde électromagnétique atteint le sol, les cibles contenues dans l'empreinte du radar réfléchissent vers l'antenne une partie de l'onde émise. Le temps de propagation d'une onde réfléchi par une cible étant proportionnel à la distance radar-cible, le radar localise la cible dans la direction radiale (Fig I.1). Le signal réfléchi par l'ensemble des cibles contenues dans l'empreinte permet d'obtenir après échantillonnage une ligne de l'image. Entre deux impulsions, le radar s'est déplacé et à la seconde impulsion, on obtient par le même procédé une seconde ligne de l'image. En itérant ce processus pour toutes les impulsions émises, on obtient après échantillonnage et quantification une image numérique de la surface observée.

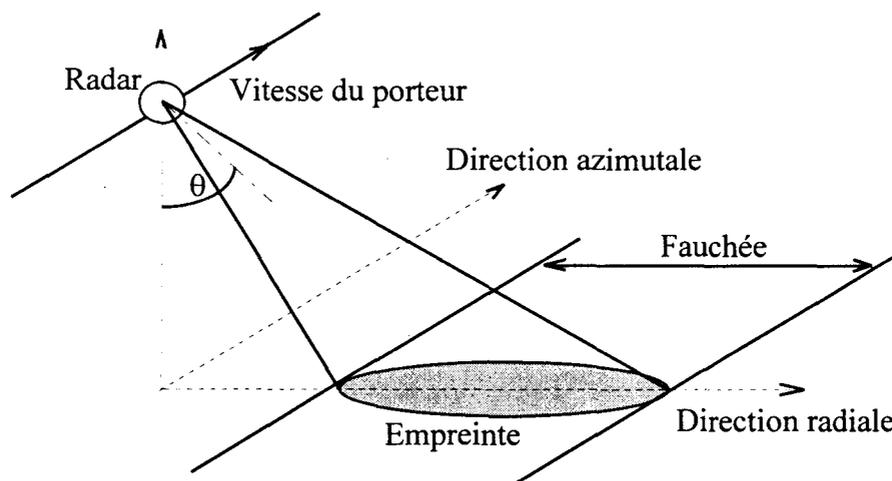


Fig I.1: Principe d'un radar imageur à visée latérale.

Une image d'une zone observée est alors constituée de pixels qui représentent une surface au sol dont les dimensions sont définies dans la direction azimutale et radiale (fig I.2).

Le niveau de gris des images SEASAT et ERS1 est calculé à partir du module de l'onde rétrodiffusée d'une cellule de résolution. Afin d'améliorer la lisibilité des images radar, l'aspect granulaire est diminué par une technique de moyennage multivue. Cette technique consiste à moyennner l'intensité (carré du module) des n_1 ondes rétrodiffusées

- ouverture synthétique. Le principe du radar à ouverture synthétique à visée latérale consiste à utiliser la variation linéaire de la fréquence Doppler de l'onde réfléchie d'une cible fixe balayée par le faisceau radar. Par analyse des fréquences Doppler du signal, on synthétise un radar de plus faible ouverture et donc de meilleure résolution azimutale.

I.1.b Caractéristiques des radars à visée latérale

SEASAT

SEASAT [All84, Way91] fut le premier satellite dont la mission était l'observation des océans. Lancé le 27 juin 1978 à une altitude de 800 km et animé d'une vitesse de $6,7 \text{ km.s}^{-1}$, le satellite effectuait plus de 14 tours de la terre chaque jour fournissant ainsi pendant les trois mois de son existence des images représentant plus de 100 millions de km^2 de la surface terrestre. Le radar à antenne synthétique à visée latérale fonctionnait en bande L (1.25 GHz, 23.5 cm) et avait un angle d'incidence θ de 20° permettant d'observer la terre sur une fauchée de 100 km. Les images produites par le SAR ont une résolution radiale de 25 mètres et une résolution azimutale de 25 m pour le mode 4 vues.

ERS1

Le satellite ERS1 [Way91] est le second satellite à vocation océanographique à embarquer un radar à ouverture synthétique. Développé par l'European Space Agency, ERS1 a été lancé le 16 Juillet 1991 et a une mission de deux ans orientée principalement vers l'observation des glaces et de l'océan [ARR81].

Le radar à antenne synthétique est aussi à visée latérale, il fonctionne en bande C (5.6 GHz, 5.3 cm) en polarisation verticale et a un angle d'incidence de 23° . Les images sont obtenues sur une fauchée de 100 km, leurs résolutions radiale et azimutale sont de l'ordre de 20 mètres pour 3 vues.

JERS1

Le satellite JERS1 est également développé par la NASDA japonaise. Il opère dans la bande L-band (1,25 Ghz, 24cm) en polarisation HH et a un angle d'incidence de 32° .

I.2 Modèle hiérarchique généralisé

Désignons par M le nombre de canaux de l'image multispectrale et par S l'ensemble des sites de l'image. L'intensité de l'onde retrodiffusée par le site s et collectée par le canal m peut être décrite par une variable aléatoire $Y_{m,s}$.

L'ensemble de ces variables aléatoires forme un champ aléatoire $Y = (Y_{m,s})_{s \in S; m=1, \dots, M}$

ou tout simplement: $Y = (Y_s)_{s \in S}$

$$\text{avec } Y_s = \begin{pmatrix} Y_{1,s} \\ \vdots \\ Y_{M,s} \end{pmatrix}$$

Le modèle hiérarchique permet de distinguer l'information texture qu'on cherche à extraire de l'information bruit que l'on observe effectivement sur une image. Il définit pour cela une image comme étant la réalisation particulière de deux processus aléatoires, un processus de classes et un processus de mesure.

Le processus des classes est décrit par un champ aléatoire:

$$X = (X_s)_{s \in S} \quad (I.2)$$

Où chaque X_s est une variable aléatoire qui prend ses valeurs dans un ensemble fini d'étiquettes ou ensemble de classes:

$$\Omega = \{\omega_1, \dots, \omega_K\} \quad (I.3)$$

Sur chaque site s et pour chaque canal m , le processus de mesure est décrit par la variable aléatoire $Y_{m,s}$ conditionnelle à chacune des K réalisations $X_s = \omega_1, \dots, X_s = \omega_K$.

Nous noterons désormais par $\xi_{m,s}^k$ chacune de ces variables.

La superposition du processus de classes et du processus de mesure donne lieu au champ des intensités réellement observable Y .

La figure I.5 montre le principe de la superposition du processus de classes et du processus de mesure dans le cas général multispectral.

I.3.a Modélisation locale du champ de classes

Le modèle local admet l'indépendance des composantes X_s du champ de classes.

Notons $\Pi_{k,s}$ la probabilité d'avoir sur un site s la classe ω_k ($X_s = \omega_k$). La loi du champ X est donc entièrement déterminée par les probabilités $\Pi_{k,s}$.

Selon que ces quantités dépendent ou non de la position du pixel s , nous distinguons les deux modélisations locales suivantes:

I.3.a.i Modélisation locale stationnaire

La modélisation locale stationnaire admet que les probabilités a priori $\Pi_{k,s}$ sont indépendantes de la position du site s . Les proportions des différentes classes sont donc indépendantes de l'endroit où l'on se trouve dans l'image. Bien que cette hypothèse ne soit pas valable dans le cas général, elle reste d'une utilité pratique. En effet, seules K probabilités Π_k sont nécessaires pour déterminer la loi du champ de classes.

I.3.a.ii Modélisation locale non stationnaire

Dans le cas local non stationnaire, les probabilités $\Pi_{k,s}$ dépendent de la position des sites. Si N désigne la taille de l'échantillon image, la loi du champ de classes est décrite par $N \times K$ paramètres. Cette modélisation permet de tenir compte de la dépendance des proportions des classes avec la position du pixel.

I.3.b Modélisation globale du champ de classes

La modélisation globale tient en compte de la dépendance entre les composantes X_s du champ X . Dans la mesure où le champ de classes est un vecteur aléatoire dont le nombre d'éléments est égal à N , sa loi est définie sur l'ensemble Ω^N dont le nombre d'éléments est égal à K^N . En pratique, il est impossible de calculer une telle loi. D'où le recours à un modèle simplifié qui prend en compte la dépendance statistique entre les composantes de ce champ d'une part et qui permet une plus simple écriture de la loi de ce champ d'autre part. Le modèle markovien permet une telle simplification. Son

leurs sont associées sont représentés respectivement sur les Fig.I.6 et les Fig.I.7.

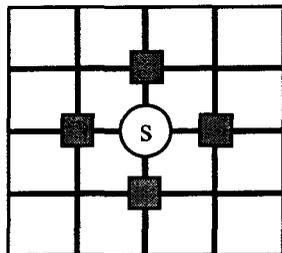


Fig I.6.a: Voisinage d'ordre 1

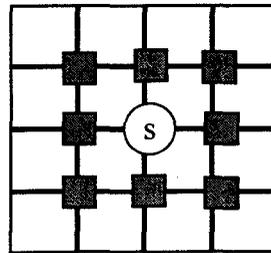


Fig I.6.b: Voisinage d'ordre 2

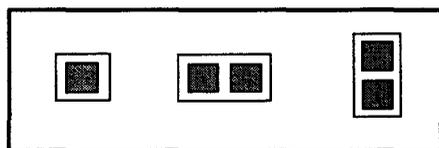


Fig I.7.a: Cliques associées au voisinage d'ordre 1

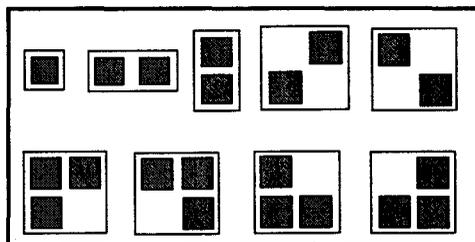


Fig I.7.b: Cliques associées au voisinage d'ordre 2

■: voisin du pixel s.

Définition III: Champ de Markov

Soit un champ aléatoire $X = \{X_s, s \in S\}$, où chaque coordonnée X_s est à valeur dans un ensemble fini $\Omega = \{\omega_1, \dots, \omega_K\}$.

X est un champ de Markov relativement au système de voisinage V si:

- α) La probabilité de toutes les réalisations de X est non nulle: $P(x) > 0$
- β) Pour tout $s \in S$: $P(X_s = x_s / X_t = x_t, t \neq s) = P(X_s = x_s / X_t = x_t, t \in V_s)$ où V_s est un voisinage de s dont la forme géométrique est indépendante de s .

Les probabilités conditionnelles $P(X_s = x_s / X_t = x_t, t \in V_s)$ s'appellent les caractéristiques locales du champ X . Nous notons désormais par $P_{X_s}^{V_s}(x_s)$ ces probabilités.

Proposition

Soit x une réalisation particulière d'un champ de Markov relativement au système de voisinage V . L'énergie U de la mesure de Gibbs s'écrit:

$$U(x) = \sum_i \left(\sum_{C \in \zeta_i} \varphi_i(x_C) \right) \quad (I.5)$$

où l'indice i représente les types de cliques, ζ_i l'ensemble des cliques de type i dans l'ensemble S et φ_i est une application qui associe à chaque configuration de (X_s) ,

$s \in C$ un nombre réel.

Le calcul de la distribution du champ markovien X nécessite donc la connaissance des fonctions φ_i . La constante de normalisation Z est une somme sur l'ensemble de toutes les configurations possibles du champ X . Dans le cas d'une image à deux classes et de taille 256×256 , le calcul de cette constante nécessite une somme sur 2^{65536} configurations. Une telle opération n'est pas réalisable directement. Il est par conséquent impossible de calculer pratiquement la distribution du champ X . Il reste cependant possible de simuler une réalisation de ce champ à partir de ses caractéristiques locales. L'échantillonneur de Gibbs permet cette simulation lorsque les probabilités conditionnelles $P_{X_s}^{V_s}$ sont connues.

Si le système de voisinage V est petit, le nombre de cliques qui lui sont associées est relativement faible. Les probabilités conditionnelles sont obtenues par:

$$P_{X_s}^{V_s}(\omega_k) = \frac{e^{-U(\omega_k, V_s)}}{\sum_{j=1}^{j=K} e^{-U(\omega_j, V_s)}} \quad (I.6)$$

avec:

$$U(\omega_k, V_s) = \sum_i \left(\sum_{C \in \zeta_i} \varphi_i(x_C) \right) \quad (I.7)$$

Markov qui converge vers l'une des réalisations du champ aléatoire X .

Cette suite est construite de la manière suivante:

Itération initiale:

On se donne une configuration initiale $X=x^0$.

Itération $n+1$:

Sur chaque site s , nous calculons la loi de X_s conditionnelle aux réalisations x_t situées dans le voisinage V_s à partir de (I.12).

On effectue un tirage selon cette loi et on remplace la classe figurant sur le pixel s par la classe obtenue par le tirage (qui peut être éventuellement la même). On passe au pixel suivant et on effectue la même opération et ainsi de suite. Après N tirages on aura effectué un balayage complet de l'image.

On obtient ainsi une suite $x^0, x^1, \dots, x^n, \dots$ de réalisations aléatoires du champ X , dont la loi converge vers (I.4).

La figure I.7 est un exemple de réalisations d'un champ de Markov par l'échantillonneur de Gibbs obtenu au cours de 50 itérations. Ce champ qui comporte trois classes a une énergie de liaison entre deux pixels voisins égale à 3.

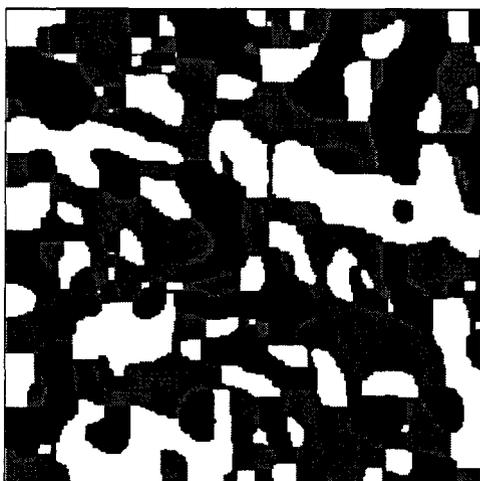


Fig I.8: Exemple de champs markoviens simulé par l'échantillonneur de Gibbs

I.4.a Système de distributions de Pearson

Les distributions du système de Pearson sont solutions de l'équation différentielle suivantes:

$$\frac{df(x)}{dx} = -\frac{(x+a)}{c_0 + c_1x + c_2x^2} f(x) \quad (\text{I.14})$$

Notons par a_1 et a_2 les racines réelles ou complexes de l'équation: $c_0 + c_1x + c_2x^2 = 0$ et posons respectivement:

$$p = \frac{a + a_1}{c_2(a_2 - a_1)} \quad (\text{I.15})$$

$$q = -\frac{a + a_2}{c_2(a_2 - a_1)} \quad (\text{I.16})$$

L'équation (I.14) s'écrit alors sous la forme:

$$\frac{df(x)}{dx} = \left(\frac{p}{x - a_2} + \frac{q}{x - a_1} \right) f(x) \quad (\text{I.17})$$

La résolution de cette équation permet de distinguer huit distributions unimodales différentes. L'ensemble de ces distributions est représenté sur le graphe de Pearson, lieu des points (β_1, β_2) .

Si m_2 , m_3 et m_4 désignent respectivement les moments centrés d'ordre 2, 3 et 4, le skewness de la distribution est donné par:

$$\beta_1 = \frac{(m_3)^2}{(m_2)^3} \quad (\text{I.18})$$

Le kurtosis de la distribution est:

$$\beta_2 = \frac{m_4}{(m_2)^2} \quad (\text{I.19})$$

Elle est aussi donnée par:

$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx \quad (I.22)$$

La fonction gamma à une variable p est donnée de la même manière par:

$$\Gamma(p) = \int_0^{+\infty} x^{p-1} e^{-x} dx \quad (I.23)$$

La forme de la distribution varie avec les paramètres p et q . Nous distinguons en particulier les 3 formes suivantes:

La forme en U: si $p \leq 0$ et $q \leq 0$.

La distribution est alors infinie pour $x = a_1$ et $x = a_2$ et elle admet un antimode sur l'intervalle $[a_1, a_2]$.

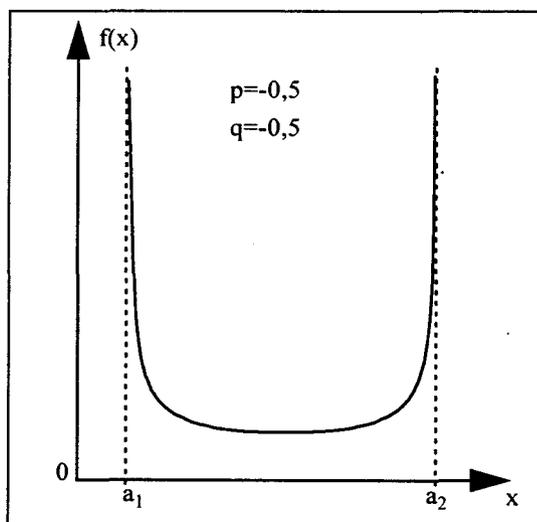


Fig I.10: Distribution bêta de première espèce en U

La forme en J: si $(p < 0$ et $q > 0)$ ou $(p > 0$ et $q < 0)$.

Cette forme de distributions n'admet pas de modes sur l'intervalle $[a_1, a_2]$.

$$\begin{cases} f(x) = \frac{\lambda^{q+1}(a_2 - x)^q e^{-\lambda(a_2 - x)}}{\Gamma(q+1)} & \text{pour } x \in]-\infty, a_2] \\ f(x) = 0 & \text{sinon} \end{cases} \quad (\text{I.24})$$

- De même si $p \rightarrow +\infty$, $a_2 \rightarrow +\infty$ et $\frac{q}{a_2} = \lambda < \infty$, la solution de l'équation (I.17) est la **distribution gamma droite** de paramètres a_1 , λ et p et définie par:

$$\begin{cases} f(x) = \frac{\lambda^{p+1}(x - a_1)^p e^{-\lambda(x - a_1)}}{\Gamma(p+1)} & \text{pour } x \in [a_1, +\infty[\\ f(x) = 0 & \text{sinon} \end{cases} \quad (\text{I.25})$$

La distribution gamma a été choisie de manière empirique par H. Laur [Laur89] pour décrire la texture d'images de surfaces boisées.

- La forme symétrique de la loi bêta de première espèce est souvent appelée **distribution de type II**. Le lieu des points de cette distribution dans le graphe de Pearson est un segment de droite. Elle est définie sur l'intervalle $[a_1, a_2]$ par:

$$f(x) = \frac{(-x^2 + (a_1 + a_2)x - a_1 a_2)^p}{(a_2 - a_1)^{2p+1} B(p+1, p+1)} \quad (\text{I.26})$$

Dans le cas particulier où $p \rightarrow +\infty$, $a_1 \rightarrow -\infty$ et $a_2 \rightarrow +\infty$, la limite de la distribution de type II est la **loi normale** de variance $\sigma^2 = \frac{a_1^2}{2p}$ et de moyenne: $\mu = \frac{a_1 + a_2}{2}$

- Si $p = q = 1$, la distribution bêta de seconde espèce est la **distribution uniforme** sur $[a_1, a_2]$.

I.4.a.ii Famille bêta de seconde espèce

Si les racines a_1 et a_2 sont toutes les deux réelles et si produit $(a_1 - \mu_1)(a_2 - \mu_2)$ est positif, la distribution correspondante dans le graphe de Pearson est la distribution bêta de seconde espèce. Cette distribution porte aussi le nom de distribution de type 6.

s'écrit:

$$\begin{cases} f(x) = \frac{(x-b)^\beta e^{\left(\frac{\lambda}{x-b}\right)}}{(-\lambda)^{\beta+1} \Gamma(-\beta-1)} & \text{si } x \in]b, +\infty[\\ f(x) = 0 & \text{sinon} \end{cases} \quad (\text{I.29})$$

- Si $q \rightarrow -\infty$, $a_2 \rightarrow +\infty$ et $\frac{-q}{a_2} = \lambda < \infty$, la limite de la loi bêta de seconde espèce gauche est la **loi gamma** de paramètres a_1 , p et λ .
- Si $p \rightarrow +\infty$, $q \rightarrow -\infty$, $(a_2 - a_1)q = \lambda < \infty$ et $p + q = \beta < \infty$, la limite de la loi bêta de seconde espèce droite est la **loi gamma inverse** de paramètres a , β et λ et dont la distribution s'écrit:

$$\begin{cases} f(x) = \frac{(a-x)^\beta e^{\left(\frac{\lambda}{x-a}\right)}}{(-\lambda)^{\beta+1} \Gamma(-\beta-1)} & \text{si } x \in]-\infty, a[\\ f(x) = 0 & \text{sinon} \end{cases} \quad (\text{I.30})$$

I.4.a.iii Famille de type 4

Les paramètres a_1 , a_2 , p et q sont dans ce cas complexes.

Une solution réelle de l'équation (I.17) existe, à condition que les paramètres a_1 et a_2 ainsi que les paramètres p et q soient complexes conjugués l'un de l'autre, soit

$$\begin{cases} a = a_1 = a_2 \\ p = q^* \end{cases} \quad (\text{I.31})$$

Nous notons par: $z_r = \text{Re}(z)$ (resp $z_i = \text{Im}(z)$) la partie réelle (resp imaginaire) d'un nombre complexe z quelconque.

La solution de l'équation différentielle est de la forme:

$$f(x) = \frac{(-2p_r - 1)B(-p, -p^*)(x^2 - 2a_r x + a_r^2 + a_i^2)^{2p_r} e^{-2p_i \text{Arctg}\left(\frac{x-a_r}{a_i}\right)}}{2\Pi(2a_i)^{2p_r+1}} \quad (\text{I.32})$$

$$\mu_n = \int_{-\infty}^{+\infty} x^n f_X(x) dx \quad (I.34)$$

Le moment μ_n de la distribution peut être approché en utilisant l'estimateur:

$$\hat{\mu}_n = \frac{1}{N} \sum_{i=1}^{i=N} x_i^n \quad (I.35)$$

Où N désigne la taille de l'échantillon.

Par ailleurs, la réécriture de l'équation différentielle définissant le système de distributions de Pearson permet d'obtenir l'équation récurrente:

$$-n\mu_{n-1}c_0 - (n+1)\mu_n c_1 - (n+2)\mu_{n+1}c_2 + \mu_n a + \mu_{n+1} = 0 \quad (I.36)$$

En centrant la variable aléatoire X et en identifiant les moments centrés d'ordre 2, 3, et 4 avec leurs analogues empiriques, nous obtenons quatre équations fonctions des estimateurs \hat{c}_0 , \hat{c}_1 , \hat{c}_2 et \hat{a} . La résolution de ces équations en \hat{c}_0 , \hat{c}_1 , \hat{c}_2 et \hat{a} est la suivante:

$$\hat{c}_0 = \frac{\hat{m}_2(4\hat{\beta}_2 - 3\hat{\beta}_1)}{10\hat{\beta}_2 - 12\hat{\beta}_1 - 18} \quad (I.37.a)$$

$$\hat{c}_1 = \frac{\sqrt{\hat{\beta}_1 \hat{m}_2} (\hat{\beta}_2 + 3)}{10\hat{\beta}_2 - 12\hat{\beta}_1 - 18} \quad (I.37.b)$$

$$\hat{c}_2 = \frac{2\hat{\beta}_2 - 3\hat{\beta}_1 - 6}{10\hat{\beta}_2 - 12\hat{\beta}_1 - 18} \quad (I.37.c)$$

$$\hat{a} = \hat{c}_1 \quad (I.37.d)$$

Le skewness et le kurtosis empiriques sont estimés par:

$$\hat{\beta}_1 = \frac{(\hat{m}_3)^2}{(\hat{m}_2)^3} \quad (I.38)$$

La log vraisemblance est une expression analytique plus simple en général que celle de la vraisemblance. Sa maximisation est donc souvent préférée à cette dernière.

Dans le cas du système de Pearson, l'estimation par la méthode du maximum de vraisemblance pose deux difficultés:

1: La résolution des équations ne peut être effectuée de manière analytique. On a alors recours à des méthodes numériques.

2: On ne peut séparer l'estimation du type de la loi de l'estimation des paramètres comme dans le cas de la méthode des moments.

Pour contourner cette difficulté, nous proposons de sélectionner la loi dans le graphe de Pearson en estimant β_1 et β_2 par la méthode des moments puis les paramètres de la loi sélectionnée par la méthode du MV. Cette méthode nous conduit à distinguer deux cas suivant les valeurs de (β_1, β_2) .

1^{er} cas:

La position de la distribution obtenue dans le graphe de Pearson après application de la méthode du MV est confinée dans une surface et est éloignée de toute autre distribution (Fig I.15). Dans ce cas, la loi de la distribution estimée par la méthode du maximum de vraisemblance est celle obtenue par la méthode des moments. Seuls les paramètres obtenus par les deux méthodes d'estimation sont différents. Cette hypothèse est justifiée par la continuité de la forme de la distribution lorsqu'on parcourt le graphe de Pearson.

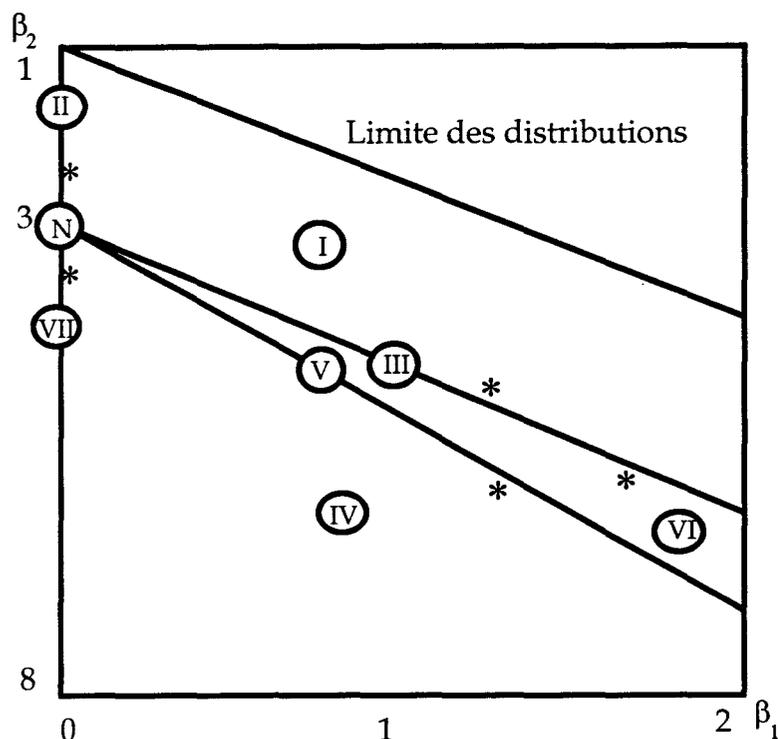


Fig 1.16: Positions de la distribution au voisinage d'une droite ou d'un point

I.4.b.iii Comparaison des deux méthodes d'estimation

L'étude comparative entre deux méthodes d'estimation est généralement effectuée par le calcul analytique du biais et de la variance, que présentent les estimateurs fournis par les deux méthodes. Le calcul de ces quantités pour les estimateurs de c_0 , c_1 , c_2 et a étant impossible à effectuer analytiquement, nous avons eu recours à une méthode numérique pour approcher le biais et la variance.

Nous proposons d'illustrer l'étude comparative des estimateurs des moments et du MV sur deux exemples.

Les deux distributions choisies admettent des formes différentes. Leurs paramètres sont présentés dans le tableau I.1.

Type de la distribution	μ_1	m_2	β_1	β_2	c_0	c_1	c_2
Bêta de 1 ^{er} espèce	50	500	0,3	3,25	111	10,7	-0,037
Bêta de 2 ^{eme} espèce	50	500	1,16	5,25	300	11	0,07

Tab I.1: Paramètres théoriques des distributions univariées utilisées

La méthode consiste, alors, à simuler un ensemble de réalisations des différents estimateurs à partir d'un ensemble d'échantillons représentant le processus des

Paramètres	Taille	β_1	β_2	c_0	c_1	c_2
Méthode des moments	16	0,7	1,5	-105	-3,64	0,09
	32	0,1	0,7	-44	-1,6	0,04
	64	0,13	0,66	-27	-0,95	0,023
	128	0,12	0,57	-19	-0,67	0,01
	256	0,12	0,57	-18	-0,63	0,015
Méthode du MV	16	0,55	1,4	-45	-0,21	0,05
	32	0,25	0,6	-15	0,0022	0,01
	64	0,08	0,2	-3,72	0,022	0,004
	128	0,04	0,11	-1,38	0,023	0,001
	256	0,03	0,07	-0,22	0,03	0,0008

Tab I.3: Valeurs du biais pour la distribution bêta de seconde espèce

Estimation de la variance

Le paramètre d'importance fondamentale dans tout problème d'estimation est la variance des estimateurs. Ce paramètre est comme le biais, fonction à la fois de la taille de l'échantillon et des paramètres théoriques de la distribution. Pour chacune des deux distributions choisies, nous avons résumé dans les tableaux I.4 et I.5 les valeurs de la variance obtenues pour chacun des estimateurs. Mise à part sa variation en fonction de la taille de l'échantillon, la variance des estimateurs varie nettement en fonction de la valeur théorique de chaque paramètre. Ceci montre la nette influence des paramètres théoriques sur l'efficacité des estimateurs. Pour une taille d'échantillon petite, les deux méthodes d'estimation offrent une variance relativement importante pour les estimateurs des deux paramètres de forme β_1 et β_2 . La variance obtenue par la méthode du MV est toutefois plus faible pour une même taille d'échantillon que celle obtenue par la méthode des moments.

sur la règle Bayes minimisant la probabilité de mauvaise classification.

Soit X (resp Y) le champ de classes (resp le champ des intensités) et soit A un ensemble quelconque tel que $A \subseteq S$. Notons par X_A la restriction de X à A . La règle de classification bayésienne induite par la fonction de perte à valeurs dans $\{0,1\}$ consiste à attribuer à chaque ensemble A la configuration x_A telle que la probabilité a posteriori $P_{X_A}^{Y=y}(x_a)$ soit maximale. Selon le choix que l'on fait pour A , on distingue deux méthodes de segmentation bayésienne:

La méthode MPM: $A = \{s\}$

L'attribution d'une classes à la composante X_s du champ de classes est effectuée à partir de la réalisation de la composante Y_s du champ des intensités. La méthode MPM est la méthode de classification naturelle si les composantes du champ de classes sont indépendantes. La règle de classification bayésienne consiste dans ce cas à attribuer au site s la classe ω_k vérifiant:

$$\omega_k = \arg (\max_{j=1, \dots, K} P_{X_s}^{Y_s=y_s}(\omega_j)) \quad (I.45)$$

La méthode reste en revanche approximative si les composantes du champ de classes ne sont pas indépendantes. Si le champ de classes est markovien, les méthodes d'estimation qui seront employées dans la suite permettront d'approcher les probabilités a posteriori les fréquences en simulant un ensemble de réalisations du champ de classes par l'échantillonneur de Gibbs.

La méthode MAP: $A = S$

Cette méthode tient en compte toute l'information $Y=y$ disponible. Si le champ de classes n'est pas indépendant, la méthode MAP représente la solution théorique adéquate. Vu qu'il est, en général, impossible de calculer analytiquement la loi du champ de classes, la règle de classification ne peut être donnée de manière explicite. Pour un champ de classes markovien, la solution peut être approchée par le recuit simulé de Geman [GeG 84]. Le maximum a posteriori est alors recherché en simulant une suite de réalisations du champ de classes selon une distribution de Gibbs dont l'énergie est donnée par:

CHAPITRE II

SEGMENTATION LOCALE DES IMAGES MONOSPECTRALES

II. Segmentation locale des images monospectrales

La segmentation locale se réfère à un modèle local pour le champ de classes. Cette méthode de classification consiste à attribuer une classe à chacun des pixels de l'image indépendamment de ses voisins. Etant basée sur le critère de classification de Bayes, la segmentation locale requiert la connaissance des probabilités a priori relatives au processus de classes et des distributions marginales relatives au processus de mesure. Or, le champ de classes n'est pas observable. Le problème de la segmentation est donc celui de l'estimation, à partir d'une réalisation particulière du champ des intensités, des paramètres modélisant la loi du champ de classes et ceux modélisant les distributions du processus de mesure. Deux méthodes sont couramment employées pour l'estimation de ces paramètres. La méthode du maximum de vraisemblance qui consiste à maximiser la vraisemblance d'un mélange de distributions et la méthode de l'espérance conditionnelle qui consiste à approcher chacun des paramètres par l'espérance conditionnelle de son estimateur. La résolution analytique des équations de vraisemblance étant impossible, la solution est approchée par des méthodes numériques, dont l'algorithme EM (Estimation Maximisation) [DH73, MP91, SDR91] et sa variante stochastique, l'algorithme SEM (Stochastic Estimation Maximisation) [CD86, MP92]. L'espérance conditionnelle est estimée itérativement par l'algorithme ICE (Iterative Conditional Estimation) [MP91, Pie92a, Pie94b]. Ces méthodes ayant été développées pour un mélange de distributions de même type, nous proposons dans ce chapitre de généraliser les algorithmes EM, SEM et ICE à un mélange de distributions de lois paramétriques appartenant au système de Pearson. Nous traitons séparément le cas où le processus de classes est stationnaire et le cas où ce processus est non stationnaire.

Ce chapitre est constitué de la manière suivante:

Le paragraphe 1 est consacré à l'estimation des paramètres dans le cadre d'une modélisation locale et stationnaire du champ de classes. Au cours de ce paragraphe, nous donnons une version des trois algorithmes itératifs EM, SEM, et ICE pour approcher numériquement l'ensemble des paramètres. Nous présentons ensuite une

$$f_{Y_s} = \sum_{k=1}^K \Pi_k f_{Y_s}^{X_s=\omega_k} \quad (\text{II.1})$$

Les probabilités a priori Π_k et les distributions $f_{Y_s}^{X_s=\omega_k}$ sont indépendantes des sites s , la loi mélange donnée par (II.1) est par conséquent indépendante de ces sites. Elle est entièrement déterminée par l'ensemble des paramètres:

$$\Theta = \{\Pi_k, \theta_k; k = 1, \dots, K\} \quad (\text{II.2})$$

Par ailleurs, la loi des couples (X_s, Y_s) est déterminée par la loi des v.a X_s ainsi que par la loi des v.a Y_s conditionnelles à X_s . La donnée d'une réalisation du champ de classes et d'une réalisation du champ des intensités permet par conséquent d'estimer l'ensemble des paramètres Θ .

Or, le champ X est inobservable, seule une réalisation particulière du champ Y est accessible. Nous avons donc à estimer l'ensemble de ces paramètres à partir de l'unique réalisation $y = (y_s)_{s \in S}$ du champ Y .

Pour estimer l'ensemble des paramètres relatifs au mélange (II.1), nous proposons l'emploi de l'une des trois approches suivantes:

- Une approche qui consiste à rechercher le maximum de la vraisemblance du mélange.
- Une approche basée sur le calcul de l'espérance conditionnelle relative aux estimateurs des différents paramètres.
- Une approche basée sur le prélèvement d'échantillons représentatifs de l'ensemble des textures de l'image.

II.1.a Méthode du maximum de la vraisemblance

Rappelons que nous disposons d'un unique échantillon représentatif du champ des intensités. L'estimation de l'ensemble des paramètres Θ par la méthode du maximum de la vraisemblance revient à maximiser la vraisemblance relative à cet échantillon par rapport à chacun des éléments de l'ensemble Θ . Pour une réalisation particulière $Y=y$ du champ des intensités, la vraisemblance s'écrit:

$$V_y(\Theta) = \prod_{s \in S} f_{Y_s}(y_s) \quad (\text{II.3})$$

algorithmes sont classiquement employés pour résoudre les équations de vraisemblance d'un mélange de distributions de même loi. Nous développons ci-dessous une version de ces algorithmes, qui est valable pour un mélange de distributions de lois inconnues et qui appartiennent au système de distributions de Pearson.

II.1.a.i Algorithme EM

L'algorithme EM (E: Estimation, M: Maximisation) est une procédure itérative qui permet d'approcher numériquement le maximum de la vraisemblance [DLR77]. L'idée de cet algorithme est d'approcher les probabilités a posteriori à l'itération courante par celles obtenues au cours de l'itération précédente. Les différents paramètres relatifs au mélange (II.1) sont alors estimés à l'itération courante par (II.6) et (II.8).

Cet algorithme est construit de la manière suivante:

Itération 0: Initialisation des paramètres

On se donne le nombre exact de classes dans l'image. Les densités marginales initiales sont choisies non nulles et bornées pour l'ensemble des niveaux de gris de l'image. L'ensemble de ces distributions, ainsi que l'ensemble des probabilités a priori peuvent être obtenus par la méthode donnée dans l'annexe 2.

Soient:

$$\Theta^{(0)} = \{\Pi_k^{(0)}, \theta_k^{(0)}; k = 1, \dots, K\} \quad (\text{II.9})$$

les paramètres initiaux.

Itération n+1: Estimation-Maximisation

Nous distinguons les deux étapes de l'algorithme EM.

Estimation

Les probabilités a posteriori sont estimées à partir des paramètres obtenus au cours de l'itération n, soit:

présente la connaissance préliminaire du nombre de classes pourrait être surmonté si l'on choisit à l'initialisation un majorant de ce nombre. On peut alors s'attendre à obtenir des probabilités a priori quasi-nulles pour les classes en surplus. Cette solution introduit, cependant, des difficultés supplémentaires, car l'ajout de classes supplémentaires augmente le nombre de maxima locaux de la vraisemblance, ce qui diminue les chances de l'algorithme de converger vers les bons paramètres. Or, une segmentation non supervisée suppose l'absence de toute information sur la statistique de l'image, telle que le nombre de classes et le nombre de maxima locaux de la vraisemblance. L'algorithme SEM que nous présentons ci-dessous permet de nous affranchir des limitations de l'algorithme EM, grâce à l'introduction d'une étape stochastique précédant la maximisation.

II.1.a.ii Algorithme SEM

Afin d'empêcher une stabilisation éventuelle de l'algorithme EM au voisinage de l'un des cols de la vraisemblance, on introduit une étape stochastique précédant l'étape de maximisation. La loi a posteriori estimée dans l'algorithme EM sert à simuler un ensemble de réalisations selon cette loi. Si cet ensemble est suffisamment grand, nous pouvons approcher les probabilités a posteriori par les fréquences. Nous utilisons ensuite cette approximation au lieu et place des probabilités a posteriori lors de l'étape de maximisation. Par conséquent, les paramètres obtenus en résolvant les équations de vraisemblance de l'algorithme EM ne se stabilisent pas en un seul point. Hormis le point maximisant la vraisemblance, aucun autre point de l'espace des paramètres ne constitue une position stable pour la procédure. Cette tendance qu'à l'algorithme SEM à approcher le maximum absolu de la vraisemblance contraint les classes en surplus à voir leurs probabilités a priori tendre vers zéro. D'où le caractère non supervisé de cet algorithme, car contrairement à l'algorithme EM, il ne demande pas le nombre exact de classes. Il suffit en effet de choisir un majorant de ce nombre. L'algorithme SEM est donc le suivant:

Itération 0: Initialisation des paramètres

Il n'est pas nécessaire de choisir le nombre exact de classes. Un majorant de ce nombre suffit, car l'algorithme élimine les classes en surplus. Le choix des

SEM ne garantit pas, contrairement à l'algorithme EM, une croissance stricte de la vraisemblance en raison de son caractère stochastique.

II.1.b Méthode de l'espérance conditionnelle

II.1.b.i Contexte général

L'estimation conditionnelle itérative ICE (Iterative Conditional Estimation) proposé par W. Pieczynski est une procédure générale d'estimation des paramètres dans le cas des données manquantes [Pie94b]. Soit un couple de champs aléatoires (X, Y) dont la loi conjointe dépend d'un paramètre β et soit $\hat{\beta}$ un estimateur de ce paramètre. Nous supposons que seul le champ Y est observable et nous envisageons de mieux approcher β à partir de la réalisation $Y=y$. Selon la procédure ICE une meilleure approximation de β peut être obtenue par l'espérance conditionnelle de $\hat{\beta}$. Or l'espérance conditionnelle dépend de β qui est par nature inconnue. La procédure itérative ICE est donc la suivante:

Notons par β^n l'estimateur du paramètre β à l'itération n et par E_n l'espérance conditionnelle calculée à partir de cet estimateur. Une approximation de β est alors obtenue à l'itération suivante par:

$$\beta^{(n+1)} = E_n(\hat{\beta} / Y = y) \quad (\text{II.17})$$

Avec I_A la fonction indicatrice de A , un estimateur des probabilités a priori Π_k est donné par les fréquences empiriques:

$$\hat{\Pi}_k = \frac{1}{N} \sum_{s \in \mathcal{S}} I_{\omega_k}(x_s) \quad (\text{II.18})$$

L'espérance conditionnelle de (II.18) est alors donnée explicitement par:

$$\Pi_k^{(n+1)} = \frac{1}{N} \sum_{s \in \mathcal{S}} P_{X_s}^{y_s^{(n)}}(\omega_k) \quad (\text{II.19})$$

Il n'est pas possible d'obtenir une telle expression pour les estimateurs $\hat{\theta}_k$ des

Cette opération est effectuée en deux étapes:

Étape 1:

Simulation de D réalisations du champ de classes selon les probabilités a posteriori $P_{X_s}^{Y_s=y_s^{(n+1)}}$. Nous obtenons ainsi D partitions différentes $(Q_1^d, Q_2^d, \dots, Q_K^d)_{d=1, \dots, D}^{(n+1)}$ de S en K classes. Les estimateurs θ_k^d relatifs à l'ensemble des réalisations obtenues sont donnés par:

$$\sum_{s \in S} \frac{\partial \text{Log} f_{k,d}(y_s)}{\partial \theta_k^d} I_{\omega_k}(x_s) = 0 \quad (\text{II.22})$$

où:

$$\theta_k^d = \begin{pmatrix} c_{0,k}^d \\ c_{1,k}^d \\ c_{2,k}^d \\ a_k^d \end{pmatrix} \quad (\text{II.23})$$

Étape 2:

Les paramètres $\theta_k^{(n+1)}$ sont donnés par:

$$\theta_k^{(n+1)} = \frac{1}{D} \sum_{d=1}^D \theta_k^d \quad (\text{II.24})$$

La loi de la distribution est alors obtenue de manière unique en fonction des paramètres estimés.

Arrêt de l'algorithme

Le critère d'arrêt est la vérification de l'inégalité (II.16)

II.1.c Méthode d'estimation par prélèvement aléatoire et contrôle

Les méthodes d'estimation développées dans (II.1.a) et (II.1.b) sont robustes, dans le sens où elles garantissent la convergence vers un maximum local de la vraisemblance.

celle de l'image et la nature des textures. Des textures fines sont mieux représentées une taille d'échantillons suffisamment petites. Pour des textures occupant des surfaces grandes et compactes, l'information est en revanche plus précise pour des échantillons de grande taille. C'est pour cette raison que l'on a choisit de varier de manière continue la taille des échantillons en fonction de l'ordre l de l'élément de la suite. La variation progressive de la taille des échantillons sélectionnés permet en effet de prendre en compte la possible présence de textures de tailles différentes. La taille des échantillons tirés varie ainsi de 4×4 à 8×8 pour une image carré présentant une taille moyenne de 256×256 . La sélection des échantillons S_k étant effectuée, nous procédons ensuite à l'estimation des paramètres décrivant le mélange.

Estimation des paramètres des lois marginales

L'estimation de l'ensemble des paramètres $\theta_k^{(l)}$ ainsi que le choix de la meilleure loi pour la distribution $f_{Y_s}^{X_s=\omega_k}$ correspondante est obtenue en appliquant la méthode du maximum de la vraisemblance sur l'échantillon S_k . La méthode des moments est également applicable, mais vu que la qualité des estimateurs de $\theta_k^{(l)}$ dépend beaucoup de la taille des échantillons, il est préférable d'utiliser d'une méthode d'estimation efficace.

Estimation des probabilités a priori Π_k

Les probabilités a priori Π_k sont obtenues en minimisant l'erreur quadratique moyenne entre l'histogramme de l'image et la densité paramétrique du mélange. Cette quantité est donnée par:

$$E = \sum_{s \in S} \left(f(y_s) - \hat{f}(y_s) \right)^2 \quad (\text{II.26})$$

Pour tenir compte de condition de normalisation de la loi a priori, nous introduisons le coefficient de Lagrange λ dans l'expression de E .

$$E = \sum_{s \in S} \left(f(y_s) - \hat{f}(y_s) \right)^2 + \lambda \left(\sum_{k=1}^{k=K} \Pi_k - 1 \right) \quad (\text{II.27})$$

L'erreur est minimum par rapport à Π_k si et seulement si:

Choix du meilleur élément de la suite

L'ensemble de paramètres modélisant la distribution mélange est choisi parmi les L éléments de la suite construite. Notons par $F^{(l)}$ la fonction de répartition obtenue à partir de l'élément $\Theta^{(l)}$ de la suite et par \hat{F} la fonction de répartition empirique (histogramme cumulé) estimée à partir de l'image. Le facteur d'ajustement de Kolmogorov-Smirnov est donné par la distance:

$$d^{(l)} = \frac{1}{\sqrt{N}} \sup_{s \in S} |F^{(l)}(y_s) - \hat{F}(y_s)| \quad (\text{II.34})$$

L'élément de la suite qui ajuste le mieux l'histogramme de l'image donc celui qui fournit la distance $d^{(l)}$ minimum.

Remarque

La solution du système (II.29) ne garantit pas le caractère positif des probabilités a priori. L'obtention d'un paramètre Π_k négatif s'explique par le fait que les K sous échantillons S_1, \dots, S_K tirés ne sont pas représentatifs des K classes de l'image. Si ce cas se produit, l'élément de la suite obtenu est rejeté. Les figures (II.1, II.2 et II.3) permettent de distinguer un ensemble de cas susceptibles de se produire lorsque l'on tire au hasard deux échantillons S_1 et S_2 dans une image qui comporte deux classes.

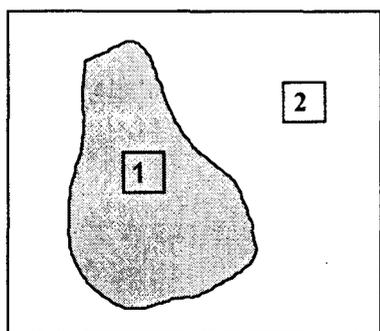


Fig II.1.a: Echantillons représentatifs des deux classes

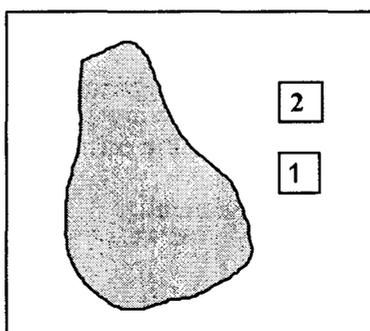


Fig II.1.b: Echantillons représentatifs d'une seule classe

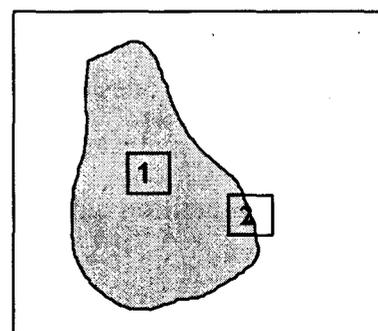


Fig II.1.c: Un seul échantillon est représentatif d'une classe

II.2 Estimation des paramètres dans le modèle non stationnaire**II.2.a Contexte général**

Le modèle local non stationnaire tient compte de la variation spatiale des probabilités a

des sites s de l'image en restreignant l'échantillon image au voisinage V_s et en appliquant ensuite l'un des algorithmes stationnaires pour l'estimation des paramètres $\Pi_{k,s}$. Comme la taille de l'échantillon V_s est relativement petite par rapport à celle de l'image, il n'est pas recommandé d'utiliser l'algorithme SEM pour estimer les probabilités a priori $\Pi_{k,s}$. Cet algorithme requiert en effet une taille d'échantillon relativement importante afin de reproduire la loi a posteriori par le tirage stochastique. La taille des échantillons V_s étant relativement petite, le tirage stochastique provoque dans ce cas une forte perturbation, ce qui cause la perte d'une quantité d'information considérable sur la loi a posteriori. Pour une taille d'échantillon V_s relativement importante (>30), l'algorithme SEM donnerait en revanche des résultats plus satisfaisant que l'algorithme EM. Comme les textures fines de l'image sont mieux sélectionnées par des petits échantillons, nous avons souvent employé des échantillons V_s de taille (<20), d'où le choix de l'algorithme EM pour l'estimation des probabilités a priori $\Pi_{k,s}$.

Les algorithmes adaptatifs que nous proposons comportent donc deux étapes:

Étape 1: Estimation des paramètres des lois marginales

Les paramètres θ_k relatifs aux distributions f_k sont estimés par l'un des algorithmes locaux et stationnaires.

Étape 2: Estimation des probabilités a priori

Les distributions f_k étant connues, l'estimation des probabilités a priori $\Pi_{k,s}$ est effectuée de la manière suivante:

- Sur chaque site s , nous choisissons un voisinage V_s , dont la taille est fixée à l'avance. Une taille de 8 lignes sur 8 colonnes est souvent utilisée dans notre étude.
- Estimation des paramètres $\Pi_{k,s}$ figurant dans le mélange (II.36) par l'algorithme EM. L'échantillon image est alors réduit au voisinage V_s entourant le pixel s .

II.2.c Comparaison de l'erreur de classification dans les deux modèles

Pour évaluer la performance des deux modélisations locales retenues au cours de notre étude, nous allons comparer les probabilités d'erreur de classification commises en adoptant l'un ou l'autre des deux modèles.

inégalité entre les probabilités d'erreurs calculées dans les deux types de modélisation.

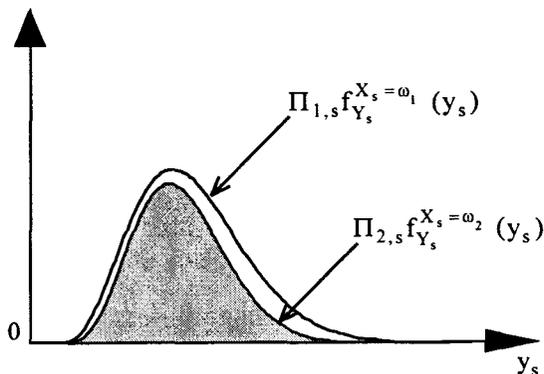


Fig II.2: Exemple de fort recouvrement avec des probabilités a priori voisines

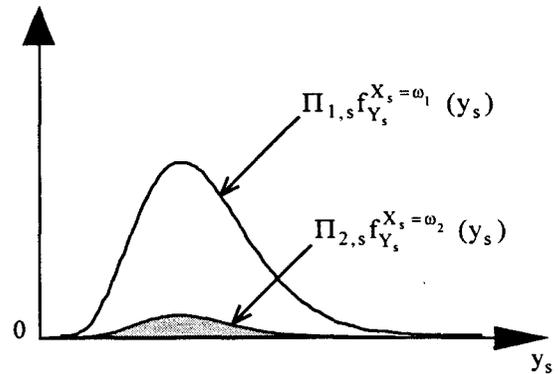


Fig II.3: Exemple de fort recouvrement avec des probabilités a priori disproportionnées

II.3 Analyse des résultats

II.3.a Segmentation locale d'une image de synthèse

Nous proposons dans cette partie l'étude de la performance de chacun des algorithmes présentés dans les paragraphes précédents à travers la segmentation d'une image de synthèse. Rappelons que dans le cas des images de synthèse, nous connaissons la réalisation du champ de classes recherchée. Ceci permet de définir le taux d'erreur de classification comme le rapport:

$$\tau = \frac{\text{Nombre de pixels mal classés}}{\text{Nombre total de pixels}} \quad (\text{II.41})$$

La donnée du champ de classes nous permet aussi de simuler une réalisation du champ de mesure. Cette opération peut être effectuée de la manière suivante:

- On se donne K distributions marginales de formes différentes et appartenant au système de distributions de Pearson. Ces distributions sont choisies de manière à ce qu'elles reproduisent le processus des intensités qu'on peut observer habituellement dans les images radar réelles.
- A partir de la réalisation du champ de classes donnée, on sélectionne pour chaque classe ω_k le sous échantillon de pixels correspondant.
- Sur chaque pixel s du sous échantillon sélectionné, nous simulons la réalisation de la distribution marginale $f_{Y_s}^{X_s=\omega_k}$ par la méthode d'inversion de la fonction de répartition.

Paramètres	Classe 1	Classe 2
π	0,67	0,33
μ	130	130
σ^2	500	500
m_3	-5914	27525
m_4	878436	3489344
β_1	0,28	0,03
β_2	3,51	3,02
Loi	6	1

Tab II.1: Paramètres théoriques du mélange

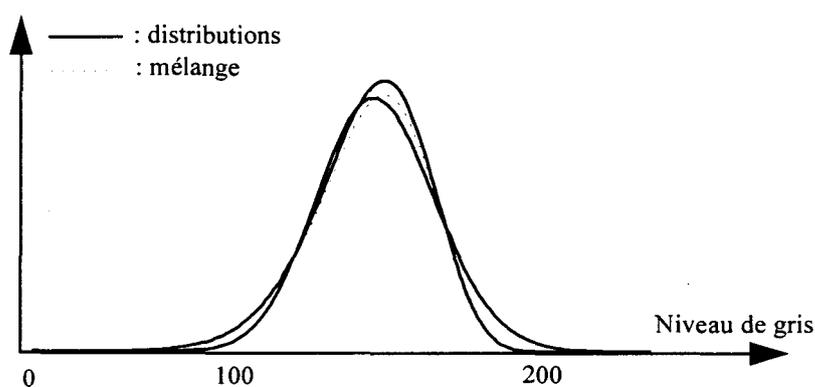


Fig II.6: Distributions marginales + mélange

II.3.a.ii Estimation des Paramètres

La comparaison entre les trois algorithmes EM, SEM et ICE locaux peut être effectuée à partir de l'erreur commise par chacun de ces algorithmes au cours de l'itération finale. Nous décrivons tout d'abord l'étape d'initialisation, avant d'évaluer la performance de l'ensemble des algorithmes.

Initialisation des algorithmes

Les densités marginales sont initialement gaussiennes. Leurs paramètres sont obtenus par la méthode de l'histogramme cumulé et les probabilités a priori initiales sont identiques.

due au caractère stochastique de ces algorithmes.

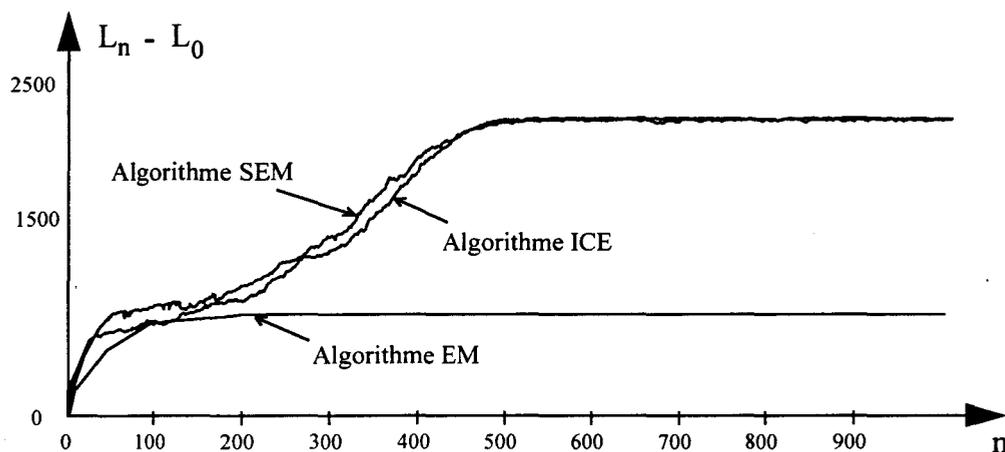


Fig II.7: Évolution de la vraisemblance de l'image de synthèse

II.3.a.iii Segmentation de l'image

En adoptant la modélisation locale et stationnaire, la probabilité d'erreur théorique calculée pour le mélange vaut approximativement 0,32. Cette quantité est représentée par la surface hachurée du graphique suivant. Notons que dans le cadre de cette modélisation, le taux d'erreur de classification est toujours supérieur à cette quantité.

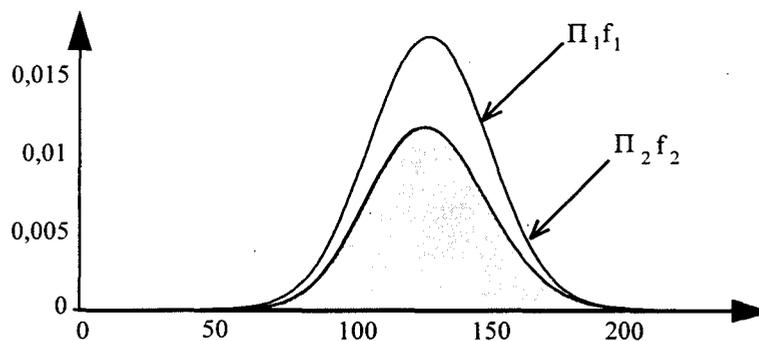


Fig II.8: Probabilité d'erreur théorique du mélange

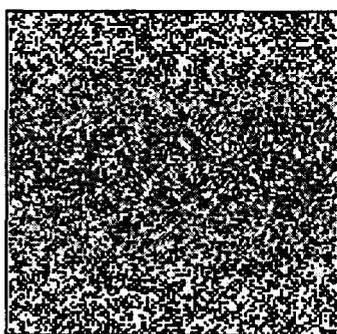


Fig II.9.a: Algorithme EM

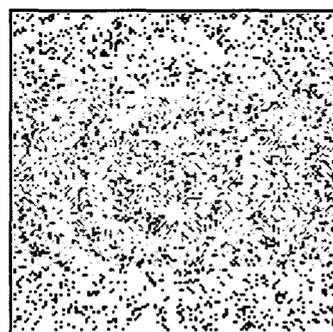


Fig II.9.b: Algorithme SEM

même, à l'intérieur de la zone maritime existe un certain nombre d'îles qui n'apparaissent pas de manière claire sur l'image satellite.

Pour étudier la performance des algorithmes sur des images réelles, nous procédons comme dans (II.3.a) à une comparaison des paramètres estimés par les différents algorithmes avant de comparer leurs performances lors de l'étape de classification.

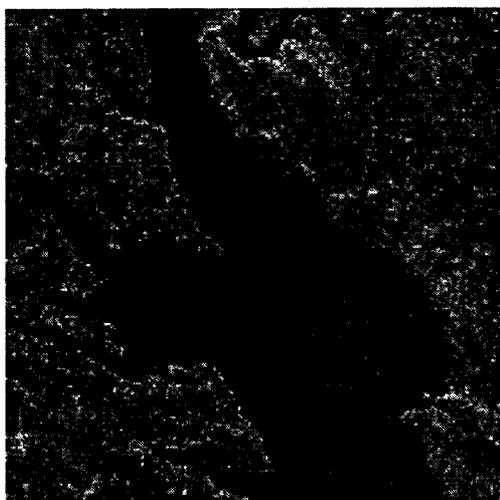


Fig II.10: Image de Brest

II.3.b.i Estimation des Paramètres

Initialisation des algorithmes

Nous avons choisi, comme dans le cas des images de synthèse, des distributions initiales gaussiennes dont les paramètres sont obtenus par la méthode de l'histogramme cumulé.

Paramètres estimés

Le tableau II.5 contient l'ensemble des résultats obtenus par les trois algorithmes d'estimation locaux et stationnaires. Les trois algorithmes convergent approximativement vers les mêmes paramètres. L'algorithme EM converge vers les mêmes paramètres que ceux obtenus par les algorithmes SEM et ICE. Ceci peut s'expliquer par le fait que l'initialisation est plus proche du maximum absolu de la vraisemblance qu'elle l'est de l'un des maximums locaux. Les résultats fournis par la méthode d'échantillonnage sont proches de ceux obtenus par les trois algorithmes itératifs précédents.



Fig II.11.a: Algorithme EM



Fig II.11.b: Algorithme SEM



Fig II.11.c: Algorithme ICE

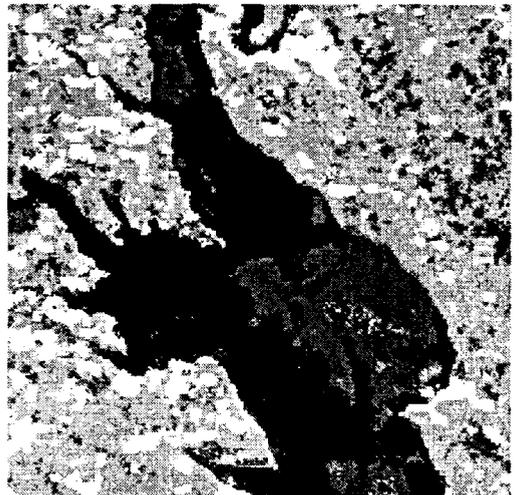


Fig II.11.d: Algorithme SEM adaptatif



Fig II.11.e: Algorithme SEM gaussien

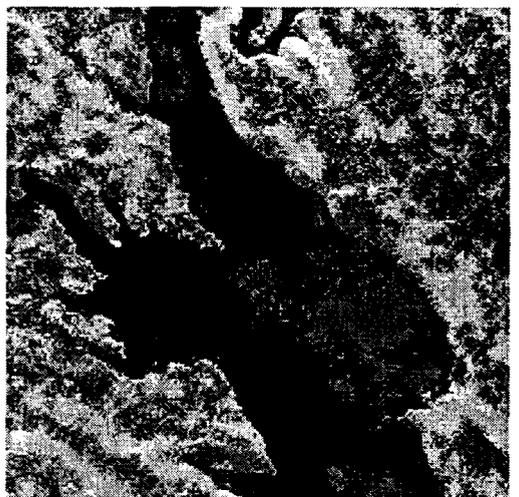


Fig II.11.f: Algorithme SEM gaussien et adaptatif

	Paramètres	Algorithme			
		EM	SEM	ICE	ECH
Classe 1	π	0,2	0,2	0,2	0,16
	μ	42	42	43	41
	σ^2	191	200	204	193
	β_1	0,15	0,18	0,2	0,2
	β_2	3,2	3,2	3,3	3,4
	Loi	1	1	6	6
Classe 2	π	0,2	0,2	0,2	0,2
	μ	73	72	72	70
	σ^2	371	357	374	337
	β_1	0,46	0,27	0,5	0,4
	β_2	3,8	3,4	3,9	3,7
	Loi	6	1	6	6
Classe 3	π	0,6	0,6	0,6	0,6
	μ	107	107	107	105
	σ^2	930	926	915	935
	β_1	0,6	0,58	0,6	0,54
	β_2	3,8	4	3,9	4,1
	Loi	6	6	6	6
Nombre d'itérations		1000	1000	1000	~

Tab II.6. Paramètres de l'image de la forêt estimés par les algorithmes locaux

II.3.c.ii Segmentation de l'image

Les figures suivantes représentent l'image de la forêt segmentée par chacun des algorithmes locaux. Les trois algorithmes stationnaires ont fourni les mêmes performances au cours de la classification, car les paramètres estimés par ces algorithmes sont très proches les uns des autres. L'algorithme SEM adaptatif présente par contre un meilleur résultat de classification meilleur. La segmentation de l'image en utilisant le modèle gaussien pour le champ de mesure a fourni des résultats sensiblement différents de ceux obtenus par notre modèle. Ce résultat n'a pas pourtant permis une séparation complète entre les pâturages et la forêt. La segmentation globale de l'image qui sera effectuée au cours du chapitre suivant permettrait une meilleure discrimination des deux classes.

II.4 Conclusion

Au cours de ce chapitre, nous avons développé quelques méthodes de segmentation statistiques basées sur une modélisation locale du champ de classes. Pour estimer les paramètres relatifs au champ de classes et ceux relatifs au champ de mesure, nous avons employé les algorithmes EM et SEM basés sur la méthode du maximum de vraisemblance et l'algorithme ICE basé sur la notion de l'espérance conditionnelle. Nous avons également introduit une méthode d'estimation basée sur le prélèvement aléatoire d'échantillons de l'image. La comparaison pratique entre les différents algorithmes construits a été réalisée grâce à l'estimation des paramètres d'une image de synthèse formée de deux classes. Nous avons remarqué une nette supériorité des algorithmes SEM et ICE vis à vis de l'algorithme EM. Les algorithmes SEM et ICE ont en effet mieux approché des distributions marginales qui modélisent le champ de mesure et les probabilités a priori qui modélisent le champ de classes. L'image étant fortement bruitée, aucun de ces trois algorithmes n'a permis de fournir des résultats de classification satisfaisants. Les résultats fournis par l'algorithme SEM adaptatif ont été en revanche probants. Appliqués à deux images réelles (SEASAT et JERS1), les résultats de la segmentation ne sont pas toujours satisfaisants, bien que l'algorithme SEM adaptatif permet d'obtenir une segmentation de meilleure qualité. Ce résultat qui concorde avec l'analyse des algorithmes sur l'image de synthèse nous encourage à proposer une modélisation des images en tenant compte de la dépendance statistique entre les composantes du champ de classes.

CHAPITRE III

SEGMENTATION GLOBALE DES IMAGES MONOSPECTRALES

III. Segmentation globale des images monospectrales

Dans le chapitre précédent, nous avons étudié le problème de segmentation sous son aspect local; celui de la classification de chaque pixel à partir de la seule donnée de son niveau de gris. Nous proposons au cours de ce chapitre l'emploi des méthodes de segmentation qui ont été utilisées couramment dans le cadre d'une modélisation markovienne du champ de classes. Les travaux réalisés dans ce domaine supposent que le processus de mesure est décrit par des lois d'un type particulier: la loi gamma ou la loi normale [DKV90, DW92, EDC84, HH87, JS 96, KDH88, LD89]. Ces études montrent, d'autre part, qu'il est possible d'estimer l'ensemble des paramètres dans le cas d'un champ de classes markovien, en utilisant une version globale de chacun des algorithmes EM et ICE [MP90, Pen92]. Nous proposons de généraliser ces deux méthodes d'estimation en modélisant le processus de mesure par les distributions du système de Pearson. Pour étudier la validité et les limitations des algorithmes construits, nous proposons leur application sur une image de synthèse et sur des images radar acquises par des capteurs différents.

Ce chapitre est constitué de la manière suivante:

Nous consacrons le premier paragraphe au développement des algorithmes d'estimation EM et ICE globaux. Nous cherchons dans un premier temps à approcher la solution des équations de vraisemblance par l'algorithme EM gibbsien. Nous développons ensuite l'algorithme ICE gibbsien, en approchant itérativement chacun des paramètres par l'espérance conditionnelle de son estimateur.

Nous analysons au paragraphe 2 l'ensemble des résultats obtenus. Les deux algorithmes construits seront testés sur une image de synthèse et des images radar SEASAT et JERS1.

III.1 Estimation des paramètres

La modélisation markovienne tient compte de la dépendance entre les variables aléatoires X_s qui composent le champ de classes X . Contrairement à la modélisation locale, la modélisation markovienne ne permet pas d'écrire la loi du couple de champs

$$P_X^{Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad (\text{III.3})$$

Où f_Y est la densité marginale du champ Y .

En posant:

$$V_y(x) = U(x) - \text{Log}f_Y^{X=x}(y) \quad (\text{III.4})$$

la loi a posteriori s'écrit sous la forme:

$$P_X^{Y=y}(x) = \frac{e^{-V_y(x)}}{f_Y(y)} \quad (\text{III.5})$$

Remarquons néanmoins que (III.5) est une distribution de Gibbs. Le champ aléatoire X conditionnel à une réalisation $Y = y$ du champ des intensités est donc markovien. La loi d'une composante X_s du champ de classes conditionnellement à une réalisation de Y_s et aux voisinages V_s est obtenue par la probabilité:

$$P_{X_s/Y_s}^{V_s}(\omega_k) = \frac{f_{Y_s}^{\omega_k}(y_s) P_{X_s}^{V_s}(\omega_k)}{\sum_{j=1}^K f_{Y_s}^{\omega_j}(y_s) P_{X_s}^{V_s}(\omega_j)} \quad (\text{III.6})$$

Il est par conséquent possible de simuler une réalisation de X selon la loi a posteriori par l'échantillonneur de Gibbs. Cette propriété est en effet essentielle tant pour la segmentation que pour l'estimation des paramètres.

III.1.a Algorithme EM gibbsien

Supposons que l'on dispose d'un ensemble $\xi = \{x^d ; d = 1, \dots, D\}$ de réalisations du champ de classes. L'estimation des paramètres relatifs à la loi du couple de champs aléatoires (X, Y) peut être effectuée dans ce cas en maximisant la vraisemblance par rapport à l'échantillon:

$$\zeta = \{(x^d, y) \mid d = 1, \dots, D\} \quad (\text{III.7})$$

Comme cette équation est impossible à résoudre (la constante de normalisation Z est incalculable analytiquement), nous proposons un estimateur du paramètre α basé sur le critère de l'erreur quadratique minimum.

Si $\hat{n}(\omega_k / V_j)$ désigne la fréquence d'apparition de la classe ω_k et de son voisinage V_j dans l'échantillon ζ , l'estimateur $\hat{\alpha}$ est celui qui minimise l'erreur quadratique moyenne:

$$\delta = \sum_{k=1}^{k=K} \sum_j \left(P_{X_s}^{V_j}(\omega_k) - \hat{n}(\omega_k / V_j) \right)^2 \quad (\text{III.14})$$

Il est par conséquent solution de l'équation:

$$\begin{cases} \sum_{k=1}^{k=K} \sum_j \frac{\partial P_{X_s}^{V_j}(\omega_k)}{\partial \alpha} \left(P_{X_s}^{V_j}(\omega_k) - \hat{n}(\omega_k / V_j) \right) = 0 \\ \frac{\partial^2 \delta}{\partial \alpha^2} > 0 \end{cases} \quad (\text{III.15})$$

Pour approcher la solution de (III.15), nous utilisons la méthode de Newton.

Dans la mesure où l'estimation des paramètres relatifs à la loi du champ de classes ne sont pas issus de l'équation (III.13), cette méthode d'estimation porte souvent le nom de la méthode de maximisation de la pseudo-vraisemblance.

Les équations (III.11) et (III.15) se prêtent à une résolution numérique immédiate, à condition de disposer de l'échantillon ζ . En pratique, on ne dispose pas d'un tel ensemble puisque la seule réalisation accessible est celle du champ Y . Rappelons à ce fait qu'on peut simuler un ensemble de réalisations ζ par l'échantillonneur de Gibbs si on connaît les probabilités conditionnelles (III.6). Mais ceci suppose aussi que l'on connaisse les probabilités conditionnelles (III.7) qui sont fonctions des paramètres à estimer. L'approche itérative de l'algorithme EM consiste donc à estimer les probabilités conditionnelles (III.7) à l'itération à courante par celles obtenues au cours de l'itération précédente.

Les paramètres $\theta_k^{(n+1)}$ relatifs aux distributions $f_k^{(n+1)}$ sont estimés par les racines de l'équation:

$$\sum_{d=1}^{d=D} \sum_{s \in S} \frac{\partial \text{Log} f_{Y_s}^{X_s = \omega_k}(y_s)}{\partial \theta_k^{(n+1)}} I_{\omega_k}(x_s^d) = 0 \quad (\text{III.18})$$

Réarrangeant la somme figurant dans (III.18) et utilisant l'expression de la loi a posteriori estimé par (III.16), nous obtenons une forme équivalente à (III.18), soit:

$$\sum_{s \in S} \frac{\partial \text{Log} f_{Y_s}^{X_s = \omega_k}(y_s)}{\partial \theta_k^{(n+1)}} (P_{X_s}^{Y_s = y_s})^{(n+1)}(x_s) = 0 \quad (\text{III.19})$$

Cette forme s'apparente à celle obtenue dans le cas local. La différence se situe dans le fait que la loi a posteriori donnée dans (III.16) n'est pas calculée analytiquement comme dans le cas local, mais approchée par les fréquences.

Arrêt de l'algorithme

Soit $\hat{n}(\omega_k)$ la fréquence d'apparition de la classe ω_k dans l'ensemble $\xi^{(n+1)}$ et soit $f^{(n+1)}$ le nouveau mélange obtenu par:

$$f^{(n+1)}(y_s) = \sum_{k=1}^K \hat{n}(\omega_k) f_k^{(n+1)}(y_s) \quad (\text{III.20})$$

La vraisemblance par rapport à $\xi^{(n+1)}$ qui est calculée sur la base des $f^{(n+1)}$ est:

$$L^{(n+1)} = \sum_{s \in S} \text{Log} f^{(n+1)}(y_s) \quad (\text{III.21})$$

Le critère d'arrêt choisit est alors la vérification de l'inégalité:

$$\frac{|L^{(n+1)} - L^{(n)}|}{|L^{(n+1)}|} < \varepsilon \quad (\text{III.22})$$

où ε est un nombre inférieur à 1.

Espérance conditionnelle

L'espérance conditionnelle des différents paramètres modélisant la loi de (X, Y) est effectuée en deux étapes:

Étape 1:

Calcul des fréquences empiriques $\hat{n}^d(\omega_k / V_s)$ à partir de chacune des D réalisations du champ de classes et estimation du paramètre α^d par (III.17).

Les estimateurs $\theta_k^{d(n)}$ sont obtenus par:

$$\sum_{s \in S} \frac{\partial \text{Log} f_{k,d}(y_s)}{\partial \theta_k^d} I_{\omega_k}(x_s^d) = 0$$

Étape 2:

Le paramètre $\theta_k^{(n+1)}$ est donnée par:

$$\theta_k^{(n+1)} = \frac{1}{D} \sum_{d=1}^D \theta_k^d \quad (\text{III.23})$$

Le paramètre $\alpha^{(n+1)}$ est donnée de manière analogue par:

$$\alpha^{(n+1)} = \frac{1}{D} \sum_{d=1}^D \alpha^d \quad (\text{III.24})$$

Arrêt de l'algorithme

Le critère d'arrêt choisit est toujours la vérification de l'inégalité (III.22).

		Algorithme		
	Paramètres	Théoriques	EM	ICE
Classe 1	μ	130	130	130
	σ^2	500	500	499
	β_1	0,28	0,29	0,27
	β_2	3,51	3,55	3,52
	Loi	6	6	6
Classe 2	μ	130	130	130
	σ^2	499	499	501
	β_1	0,01	0,03	0,01
	β_2	3,03	2,94	3,01
	Loi	1	1	1
Energie		~	-0,7	-0,9
Nombre d'itérations			300	300

Tab III.1: Paramètres de l'image de synthèse estimés par les algorithmes globaux

III.2.a.ii Segmentation de l'image

Rappelons que l'image de synthèse n'est pas une réalisation d'un champ de Markov. Ainsi, la qualité des résultats obtenus sert également à tester la capacité d'adaptation du modèle markovien aux scènes réelles non markoviennes. Les figure III.1.a et III.1.b représentent l'image de synthèse segmentée par les deux algorithmes globaux. Nous remarquons la nette supériorité de ces algorithmes par rapport aux algorithmes locaux stationnaire. Seul l'algorithme SEM adaptatif permet d'approcher la performance des algorithmes EM et ICE markoviens. Le faible taux d'erreur obtenu par ces algorithmes montre en particulier la puissance du modèle global markovien à pouvoir s'adapter aux scènes réelles non markoviennes.

Algorithme	EM gibbsien	ICE gibbsien
Taux d'erreur	0,041	0,047

Tab III.2: Taux d'erreur obtenus par les algorithmes globaux

	Paramètres	Algorithme	
		EM	ICE
Classe 1	μ	9	8
	σ^2	10	11
	β_1	0,2	0,2
	β_2	3,3	3,2
	Loi	1	1
Classe 2	μ	14	14
	σ^2	26	26
	β_1	0,5	0,4
	β_2	3,8	4
	Loi	4	4
Classe 3	μ	24	24
	σ^2	100	99
	β_1	1,3	1,2
	β_2	7	6,6
	Loi	4	4
Classe 4	μ	35	35
	σ^2	263	258
	β_1	6	6
	β_2	17	18
	Loi	6	6
Energie		-1	-0,9
Nombre d'itérations		20	20

Tab III.3: Paramètres de l'image de Brest estimés par les algorithmes globaux

III.2.b.ii Segmentation de l'image

Les figures (III.3.a et III.3.b) représentent l'image de la côte brestoise segmentée par les deux algorithmes EM et ICE globaux. Les deux images segmentées sont pratiquement les mêmes. En comparaison avec les algorithmes locaux (stationnaire ou non), nous remarquons une meilleure discrimination des classes ainsi qu'un meilleur respect des côtes.

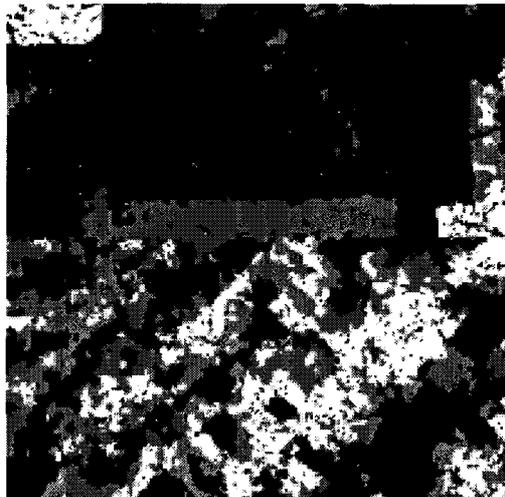


Fig III.3.a: Algorithme EM gibbsien

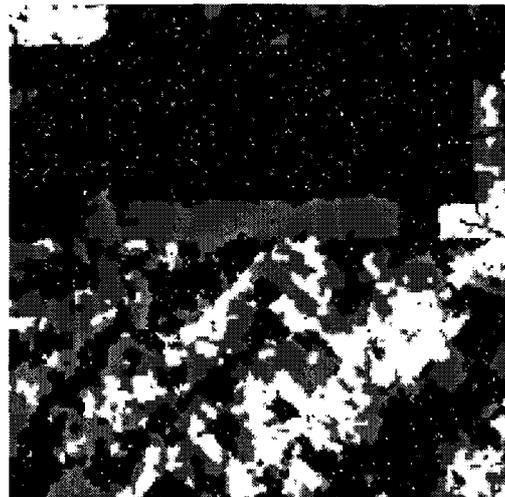


Fig III.3.b: Algorithme ICE gibbsien

III.3 Conclusion

Au cours de ce chapitre nous avons développé deux algorithmes de segmentation statistique basés sur une modélisation markovienne du champ de classes. Le premier algorithme utilise la méthode du maximum de vraisemblance pour l'estimation des paramètres. Le second algorithme utilise la notion de l'espérance conditionnelle. L'application des deux algorithmes sur une image de test et sur des images radar réelles a conclu à une même performance pour ces deux algorithmes. Les résultats d'estimation et de classifications obtenus sur l'image de synthèse et sur des images radar ont conclu à une nette supériorité des algorithmes globaux vis à vis des algorithmes locaux. Ces algorithmes ont présenté, cependant, le défaut d'être lents devant leurs analogues locaux.

CHAPITRE IV

SEGMENTATION LOCALE DES IMAGES MULTISPECTRALES

IV. Segmentation locale des images multispectrales

Au cours de ce chapitre, nous présentons les applications des méthodes de segmentation statistique décrites précédemment au cas multispectral. L'estimation des paramètres relatifs au champ de classes et des paramètres relatifs au champ de mesure sera effectuée par les mêmes procédures que celles retenues dans le cas monospectral. Nous chercherons en particulier à obtenir une expression des différents estimateurs par application de la méthode du maximum de vraisemblance et par la méthode de l'espérance conditionnelle. L'estimation des paramètres nous permettra d'appliquer ensuite le critère de Bayes pour la classification des images.

Ce chapitre est constitué de la manière suivante:

Le paragraphe 1 est consacré à l'estimation des paramètres d'une image multispectrale dans le cadre d'une modélisation locale et stationnaire du champ de classes. Nous donnons au cours de ce paragraphe une version des algorithmes d'estimation EM et SEM basés sur la maximisation de la vraisemblance et de l'algorithme ICE qui utilise la notion de l'espérance conditionnelle.

Le deuxième paragraphe permettra de généraliser à la modélisation non stationnaire du champ de classes, les méthodes d'estimation développées dans le cadre d'une modélisation locale stationnaire.

Nous analysons dans le troisième paragraphe l'ensemble des résultats obtenus sur une image de test et des images radar de type SIR-SAR-C/X et ERS-JERS.

Nous terminons enfin ce chapitre par une conclusion.

IV.1 Estimation des paramètres dans le modèle local stationnaire

Pour une image multispectrale à M canaux et dont le nombre de classes est égal à K , les paramètres à estimer dans le cadre d'une modélisation locale et stationnaire du champ de classes sont les probabilités a priori Π_k relatives aux classes $(\omega_k)_{k=1,\dots,K}$ et les paramètres $\theta_{m,k}$ relatifs aux distributions du champ de mesure $(f_{m,k})_{m=1..M;k=1,\dots,K}$ [MP91, MP92]. Rappelons que nous ne disposons que d'une réalisation $Y=y$

où $P_{X_s}^{Y_s=y_s}$ désigne la probabilité de X_s conditionnelle à Y_s (probabilité a posteriori). Elle est obtenue par:

$$P_{X_s}^{Y_s=y_s}(\omega_k) = \frac{\prod_k f_{Y_s}^{X_s=\omega_k}(y_s)}{\sum_{j=1}^{j=K} \prod_j f_{Y_s}^{X_s=\omega_j}(y_s)} \quad (\text{IV.6})$$

(Pour le détail des calculs voir annexe 3)

Dans ces expressions, aucun paramètre n'apparaît de manière explicite. Il est par conséquent impossible d'estimer directement l'ensemble des paramètres Θ . Nous proposons de résoudre itérativement l'ensemble de ces équations à l'aide des deux algorithmes: EM et SEM.

IV.1.a.i Algorithme EM

Les équations (IV. 4) et (IV. 5) s'adaptent à une résolution itérative par l'algorithme EM. Au cours de chaque itération, les probabilités a posteriori données par (IV. 5) sont estimées à partir des paramètres donnés au cours de l'itération précédente [DLR77]. L'algorithme EM ainsi construit converge vers un maximum local de la vraisemblance. L'algorithme EM multispectral est le suivant:

Itération 0: Initialisation des paramètres

Les distributions initiales $f_{m,k}^{(0)}$ sont non nulles et bornées pour tous les pixels du canal m .

Soient: $\Theta^{(0)} = \{\Pi_k^{(0)}, \theta_{m,k}^{(0)} : k = 1, \dots, K; m = 1, \dots, M\}$ les paramètres initiaux.

Itération n+1: Estimation - Maximisation

Nous distinguons les deux étapes de l'algorithme EM.

Estimation

Les probabilités a posteriori sont estimées à partir de l'ensemble des paramètres $\Theta^{(n)}$ obtenus au cours de l'itération n , soit:

Itération 0: Initialisation des paramètres

Les distributions initiales relatives à chacun des canaux de l'image sont non nulles et bornées pour tous les pixels de ces canaux.

Itération n+1: Estimation - Maximisation*Estimation*

Les probabilités a posteriori sont estimées par (IV.7).

Etape stochastique

Au cours de cette étape, un tirage stochastique est effectué dans l'image selon les probabilités a posteriori estimées. Nous obtenons ainsi une partition $(Q_1, \dots, Q_K)^{(n+1)}$ de S en K classes.

Maximisation

Les probabilités a priori $\Pi_k^{(n+1)}$ sont estimées à partir de la partition $(Q_1, \dots, Q_K)^{(n+1)}$ par:

$$\Pi_k^{(n+1)} = \frac{\text{card}(Q_k)}{\text{card}(S)} \quad (\text{IV.11})$$

Les paramètres $\theta_{m,k}$ relatifs aux distributions $f_{m,k}^{(n+1)}$ sont donnés par les MxK équations:

$$\sum_{s \in S} \frac{\partial \text{Log} f_{m,k}^{(n+1)}(y_{m,s})}{\partial (\theta_{m,k}^{(n+1)})} I_{\omega_k}(x_s) = 0 \quad (\text{IV.12})$$

Arrêt de l'algorithme

Le critère d'arrêt reste la vérification de l'inégalité:

$$\frac{|L^{(n+1)} - L^{(n)}|}{|L^{(n+1)}|} < \varepsilon \quad (\text{IV.13})$$

où ε est un nombre inférieur à 1.



Les estimateurs $\theta_{m,k}^d$ relatifs à chacune des réalisations obtenues sont donnés par la solution des équations de vraisemblance:

$$\sum_{s \in \mathcal{S}} \frac{\partial \text{Log} f_{k,d}(y_{m,s})}{\partial \theta_{m,k}^d} I_{\omega_k}(x_s) = 0 \quad (\text{IV.14})$$

Etape 2:

L'espérance conditionnelle de $\hat{\theta}_{m,k}$ est approchée par:

$$\theta_{m,k}^{(n+1)} = \frac{1}{D} \sum_{d=1}^D \theta_{m,k}^d \quad (\text{IV.15})$$

La loi de la distribution est obtenue de manière unique en fonction des paramètres estimés.

Critère d'arrêt

Le critère d'arrêt choisi est la vérification de l'inégalité (IV.13).

IV.2 Estimation des paramètres dans le modèle local non stationnaire

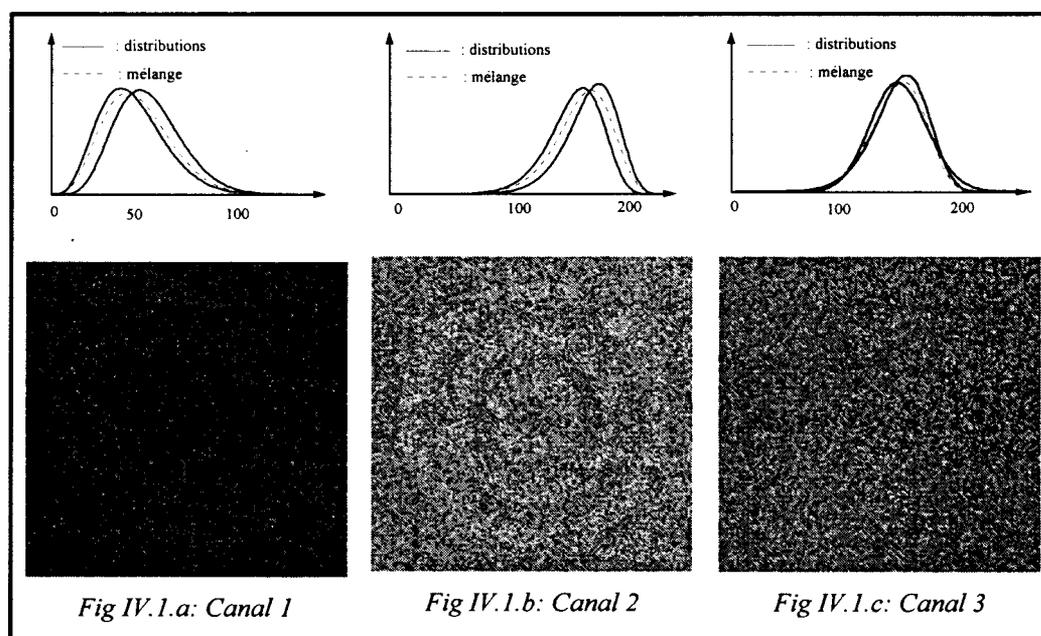
Rappelons que le modèle local non stationnaire tient compte de la variation des probabilités a priori en fonction de la position du site s . La distribution des vecteurs aléatoires Y_s dépend également de ce site, puisqu'elle s'écrit en fonction des probabilités a priori $\Pi_{k,s}$. Cette distribution est en effet donnée par le mélange:

$$f_{Y_s}(y_s) = \sum_{k=1}^K \Pi_{k,s} f_k(y_s)$$

L'estimation des probabilités $\Pi_{k,s}$ est alors effectuée, comme dans le cas monospectral en deux étapes:

et la classe ω_2 au reste de l'image. Les probabilités a priori définies dans le modèle local et stationnaire sont estimées par la fréquence d'apparition de chacune des deux classes sur un pixel quelconque de l'image. Le nombre de canaux choisis est égal à trois. Ce nombre correspond aux trois images de la cible bruitée (figures: IV.2.a, IV.2.b et IV.2.c). Le champ de mesure est obtenu par simulation sur chacun des trois canaux de deux distributions appartenant au système de distributions de Pearson. Les paramètres correspondants aux six distributions de l'image multispectrale sont résumés dans le tableau IV.1.

Afin de tester la capacité de nos algorithmes à identifier un mélange de distributions de formes différentes, nous avons choisi des distributions de formes voisines pour chacun des trois canaux. Le troisième canal de l'image multispectral n'est autre que l'image de la cible bruitée que nous avons segmentée dans le cas monospectral.



		Algorithmes											
		Paramètres Th			EM			SEM			ICE		
		C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
Classe 1	μ	50	150	130	49	150	130	49	150	130	49	149	130
	σ^2	400	500	500	394	490	505	397	497	502	394	491	503
	β_1	0,4	0,46	0,28	0,45	0,52	0,35	0,46	0,48	0,33	0,5	0,51	0,34
	β_2	3,46	3,85	3,51	3,56	3,9	3,62	3,6	3,84	3,62	3,57	3,88	3,62
	Loi	1	6	6	1	6	6	1	6	6	1	6	6
	Π	0,63			0,57			0,61			0,59		
Classe 2	μ	60	160	130	59	157	130	59	159	130	59	159	130
	σ^2	400	500	500	405	512	493	395	506	496	405	511	495
	β_1	0,26	0,54	0,03	0,24	0,5	0,01	0,25	0,49	0,02	0,24	0,5	0,01
	β_2	3,26	4	3,02	3,2	3,88	2,96	3,2	3,94	2,96	3,2	3,9	2,95
	Loi	1	6	6	1	6	1	1	6	1	1	6	1
	Π	0,37			0,43			0,39			0,41		
Itérations					200			200			200		

Tab IV.2: Paramètres de l'image de synthèse estimés par les algorithmes locaux

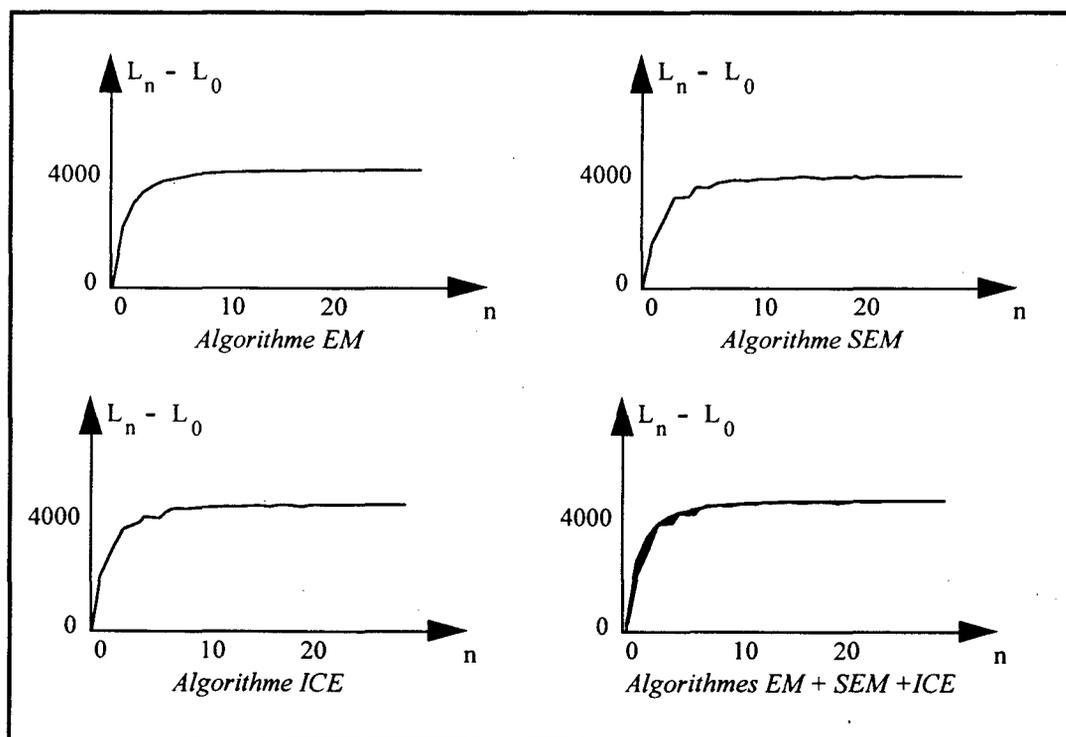


Fig IV.2: Evolution de la vraisemblance en fonction des itérations

IV.3.b Segmentation d'une image radar SIR-C/X-SAR

L'image multispectrale présentée sur les figures IV.4 correspond à un ensemble de trois vues (ou canaux) prises de la région de Pennsylvanie aux USA. Les trois vues correspondent séparément à l'une des trois bandes de fréquences C, L ou X. Elles admettent, toutes les trois le même type de polarisation VV. Ce choix de l'ensemble des canaux de l'image assure l'indépendance des variables aléatoires qui décrivent le champ de mesure.

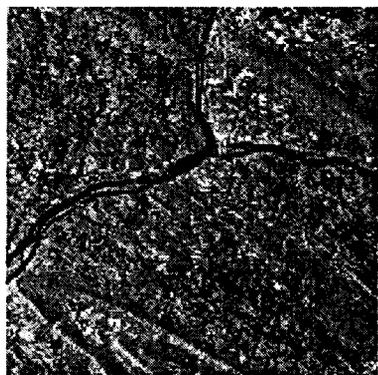


Fig IV.4.a: Canal C

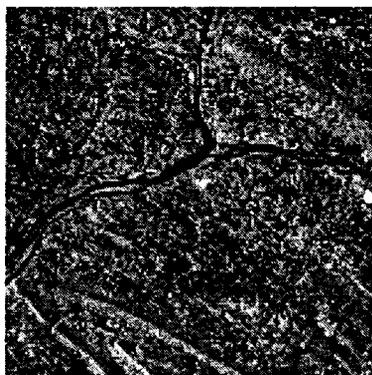


Fig IV.4.b: Canal L

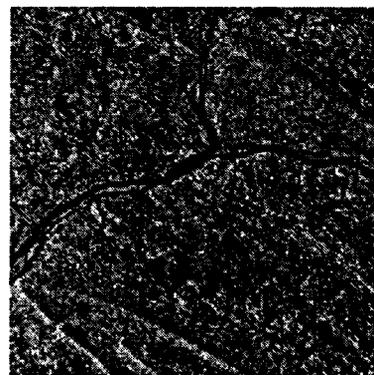


Fig IV.4.c: Canal X

IV.3.b.i Estimation des paramètres

Le tableau IV.4 résume l'ensemble des paramètres obtenus par les différents algorithmes locaux sur l'image radar. Nous remarquons d'une part une performance quasi-identique des trois algorithmes EM, SEM et ICE. D'autre part, l'image présente des distributions marginales de formes variable d'un canal à un autre. Ceci montre l'intérêt particulier à utiliser des capteurs différents pour chacun des trois canaux, car la diversité dans la forme des distributions reflète la richesse de l'image imagée sur des bandes de fréquence différentes. Les détails ignorés par l'un des capteurs se trouvent pris en compte par un autre, contribuant ainsi à une meilleure discrimination des différentes textures de l'image.

La figure IV.5 représente l'évolution de la log-vraisemblance du mélange en fonction des itérations de chacun des algorithmes. Les trois algorithmes locaux évoluent de manière similaire. Par ailleurs, nous remarquons un faible fléchissement de la croissance des différents algorithmes. Ceci est dû à l'existence de deux modes prononcés dans l'histogramme relatif à chacun des canaux. La vraisemblance se

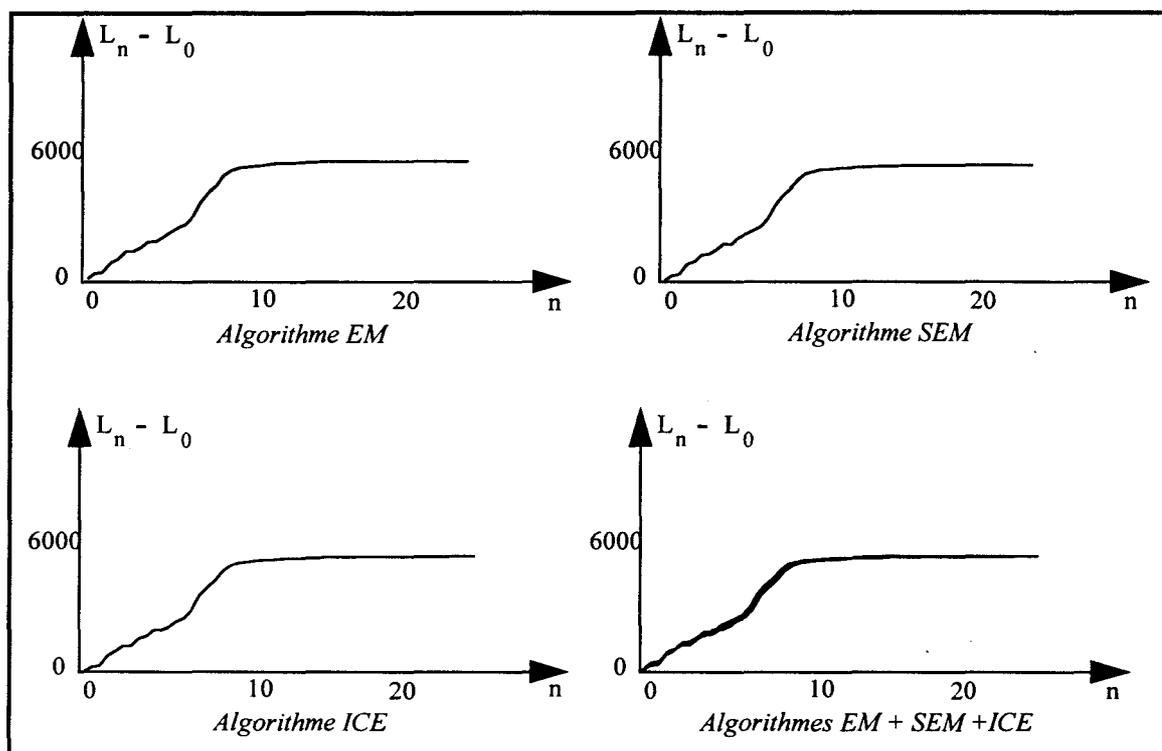


Fig IV.5: Evolution de la vraisemblance estimée par les différents algorithmes locaux

IV.3.b.ii Classification de l'image

Les performances identiques qu'ont présentées les algorithmes EM, SEM et ICE au cours de l'étape d'estimation expliquent les résultats identiques obtenus au cours de l'étape de classification. Nous obtenons, toutefois, une meilleure cohésion des classes de l'image en employant l'algorithme SEM adaptatif.

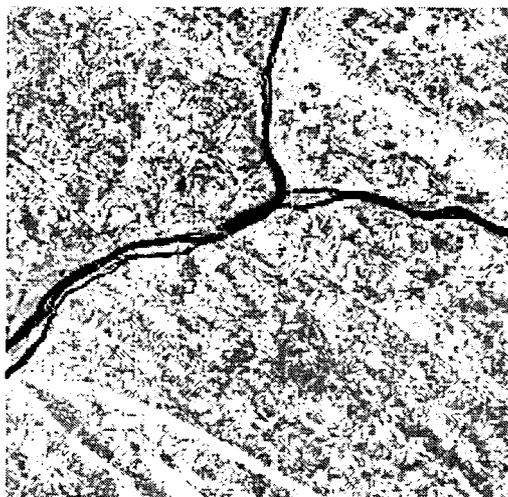


Fig IV.6.a: Algorithme EM

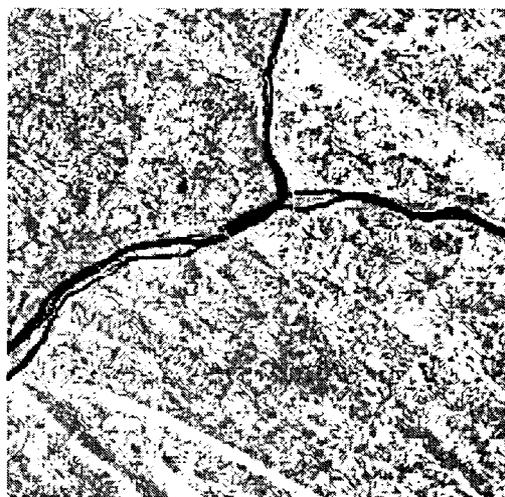


Fig IV.6.b: Algorithme SEM

champ de mesure estimés par les algorithmes EM, SEM et ICE sont proches les uns des autres. Les probabilités a priori obtenues sont également identiques.

		Algorithmes					
		EM		SEM		ICE	
		C1	C2	C1	C2	C1	C2
Classe 1	μ	44	40	45	40	45	40
	σ^2	353	216	354	215	359	213
	β_1	1,5	0,7	1,64	0,6	1,6	0,7
	β_2	6,9	4,6	7,16	4,4	7,18	4,8
	Loi	4	4	4	4	4	4
	Π	0,14		0,14		0,14	
Classe 2	μ	63	70	63	70	63	70
	σ^2	389	432	386	427	388	427
	β_1	0,35	0,3	0,3	0,28	0,3	0,3
	β_2	3,47	3,3	3,4	3,35	3,4	3,4
	Loi	1	1	1	1	1	1
	Π	0,26		0,26		0,26	
Classe 3	μ	81	106	81	106	81	106
	σ^2	678	967	679	964	678	966
	β_1	0,5	0,48	0,5	0,48	0,5	0,4
	β_2	4,18	3,9	4,16	3,9	4,1	3,9
	Loi	4	6	4	6	4	6
	Π	0,6		0,6		0,6	
Itérations		100		100		100	

Tab IV.5: Paramètres de l'image forestière estimés par les algorithmes locaux

IV.3.c.ii Classification de l'image

La segmentation de l'image par les algorithmes locaux stationnaires n'est pas performante. Nous remarquons en revanche une meilleure discrimination des classes en appliquant l'algorithme SEM adaptatif.

CHAPITRE V

SEGMENTATION GLOBALE DES IMAGES MULTISPECTRALES

V. Segmentation globale des images multispectrales

Ce chapitre est consacré à la segmentation des images multispectrales dans l'hypothèse d'une modélisation globale du champ de classes. Nous cherchons à approcher l'ensemble des paramètres modélisant la loi du champ de classes et l'ensemble des paramètres modélisant les distributions relatives au processus de mesure par les méthodes qui ont été employées au cours du chapitre 3. Nous généralisons au cas multispectral, les algorithmes EM et ICE gibbsiens développés dans le cas monospectral. Pour vérifier la robustesse des méthodes employées, nous proposons de tester l'ensemble des algorithmes sur une image de synthèse et sur des images radar multispectrales acquises par des capteurs différents.

L'organisation de ce chapitre est la suivante:

Au cours du paragraphe 1, nous développons une version de l'algorithme EM gibbsien valable dans le cas multispectral. Nous procédons ensuite à l'estimation de l'ensemble des paramètres par la méthode de l'espérance conditionnelle en développant une version de l'algorithme ICE gibbsien.

Nous analysons dans le paragraphe 2 l'ensemble des résultats obtenus par nos algorithmes sur une image de synthèse et sur des images radar.

V.1 Estimation des paramètres

La segmentation globale des images multispectrales requiert la connaissance du paramètre α définissant la loi du champ de classes, ainsi que les paramètres $\theta_{m,k}$ relatifs au champ de mesure. Rappelons que la loi du couple (X, Y) n'est plus définie à partir de la loi des variables aléatoires individuelles (X_s, Y_s) en raison de la dépendance mutuelle entre les variables aléatoires X_s .

Les variables aléatoires $Y_{m,s}$ sont supposées indépendantes conditionnellement à une réalisation particulière de X_s , la densité de Y conditionnelle à la réalisation $X=x$ s'écrit:

$$f_Y^{X=x}(y) = \prod_{s \in S} \prod_{m=1}^{m=M} f_{Y_{m,s}}^{X_s=x_s}(y_{m,s}) \quad (\text{V.1})$$

L'estimation des paramètres $\theta_{m,k}$ est obtenue en maximisant la vraisemblance L_ζ . Ils sont par conséquent solutions de l'équation:

$$\sum_{d=1}^D \sum_{s \in S} \frac{\partial \text{Log} f_{Y_{m,s}}^{X_s = \omega_k}(y_{m,s})}{\partial \theta_{m;k}} = 0 \quad (\text{V.4})$$

où les distributions $f_{Y_{m,s}}^{X_s = \omega_k}$ appartiennent au système de Pearson.

La maximisation de la vraisemblance par rapport au paramètre α étant impossible à effectuer. Il est cependant possible d'obtenir une approximation de ce paramètre en minimisant l'erreur quadratique moyenne entre les probabilités conditionnelles paramétriques et celles estimées par les fréquences, soit:

$$\begin{cases} \sum_{k=1}^{k=K} \sum_j \frac{\partial P(\omega_k / V_j)}{\partial \hat{\alpha}} (P(\omega_k / V_j) - \hat{n}(\omega_k / V_j)) = 0 \\ \sum_{k=1}^{k=K} \sum_j \frac{\partial^2 (P(\omega_k / V_j) - \hat{n}(\omega_k / V_j))^2}{\partial \hat{\alpha}^2} > 0 \end{cases} \quad (\text{V.5})$$

Où $\hat{n}(\omega_k / V_j)$ désigne la fréquence d'apparition de la classe ω_k et de son voisinage V_j dans l'échantillon ζ .

La solution des équations (V.4) et (V.5) est approchée par l'algorithme EM Gibbsien. Le déroulement de cet algorithme est le suivant:

Itération 0: Initialisation des paramètres

On se donne un ensemble de distributions $f_{Y_{m,s}}^{X_s = \omega_k^{(0)}}$ bornées ainsi qu'un paramètre α .

Itération n+1: Estimation - Maximisation

Nous distinguons les deux phases de l'algorithme EM

Estimation

Simulation d'un ensemble de réalisations $\xi^{(n+1)} = \{x_s^d; s \in S, d = 1, \dots, D\}$ du champ de classes par l'échantillonneur de Gibbs. Nous utilisons pour cela les probabilités conditionnelles $P_{X_{d,s}/y_s}^{V_{d,s}}^{(n+1)}$ calculées à partir des paramètres estimés

V.1.b Algorithme ICE gibbsien

L'algorithme ICE multispectral gibbsien est analogue dans sa construction à l'algorithme ICE monospectral. La seule différence se situe dans l'estimation des paramètres relatifs au champ de mesure. Au lieu d'estimer ces quantités sur un seul canal, nous l'estimons ici pour chacun des M canaux de l'image multispectrale.

Itération 0: Initialisation

On se donne un ensemble de paramètres $\theta_k^{(0)}$ pour la distribution de Y conditionnelle à X ainsi qu'un paramètre $\alpha^{(0)}$ pour la distribution du champ X.

Itération n+1: Espérance conditionnelle - Estimation

Les deux phases de cet algorithme sont respectivement la phase d'estimation des probabilités conditionnelles grâce à la simulation de l'échantillon représentatif ζ et la phase d'estimation de l'espérance conditionnelle relativement aux différents paramètres.

Estimation

Simulation d'un ensemble de réalisations $\xi^{(n+1)} = \{x_s^d; s \in S, d = 1, \dots, D\}$ par la même procédure utilisée dans l'algorithme EM.

Espérance conditionnelle

L'espérance conditionnelle des différents paramètres modélisant la loi de (X,Y) est effectuée en deux étapes:

Etape 1:

Calcul des fréquences empiriques $\hat{n}^d(\omega_k / V_s)$ à partir de chacune des D réalisations du champ de classes et estimation d'un ensemble de paramètres α^d par (V.5).

Les estimateurs $\theta_{m,k}^d$ relatifs à chacune des réalisations obtenues sont donnés par:

$$\sum_{s \in S} \frac{\partial \text{Log} f_{Y_{m,s}}^{x_s^d = \omega_k}(y_{m,s})}{\partial \theta_{m,k}^d} I_{\omega_k}(x_s^d) = 0$$

Paramètres estimés par les différents algorithmes

Les deux algorithmes EM et ICE globaux permettent une meilleure approche des valeurs théoriques issues de la simulation (Tab V.3). En dépit de l'initialisation arbitraire des algorithmes, nous obtenons ici une convergence rapide. Cette rapidité peut être expliquée par la contrainte à l'égalité qui a contribué à accélérer les algorithmes EM, SEM et ICE locaux. L'énergie définissant la loi a priori est sensiblement proche de celle obtenue dans le cas monospectral.

		Algorithmes								
		Paramètres Th			EM global			ICE global		
		C1	C2	C3	C1	C2	C3	C1	C2	C3
Classe 1	μ	50	150	130	50	50	130	50	150	130
	σ^2	400	500	500	397	497	501	397	496	500
	β_1	0,4	0,46	0,28	0,41	0,47	0,30	0,41	0,47	0,30
	β_2	3,46	3,85	3,51	3,47	3,86	3,56	3,48	3,85	3,56
	Loi	1	6	6	1	6	6	1	6	6
Classe 2	μ	60	160	130	60	160	130	60	160	130
	σ^2	400	500	500	397	495	499	397	495	499
	β_1	0,26	0,54	0,03	0,27	0,55	0,03	0,27	0,56	0,03
	β_2	3,26	4	3,02	3,26	4,0	2,95	3,26	4,02	2,94
	Loi	1	6	6	1	6	1	1	6	1
Energie		~			-0,6			-0,6		
Itérations		~			100			100		

Tab V.1: Paramètres de l'image de synthèse estimés par les algorithmes globaux

V.2.a.ii Segmentation de l'image

Les images suivantes représentent l'image de la cible multispectrale segmentée par les algorithmes EM et ICE globaux. Nous remarquons en particulier le faible taux d'erreur obtenu par ces algorithmes, ceci compense leur relative lenteur par rapport à leurs analogues locaux. Les algorithmes de segmentation globaux sont supérieurs à l'algorithme SEM adaptatif, vu le meilleur respect des bordures que présente l'image segmentée.

V.2.b.ii Segmentation de l'image

Les résultats de classification obtenus par les deux algorithmes gibbsiens sont pratiquement identiques car les paramètres obtenus par ces derniers sont très proches. Les algorithmes globaux présentent une meilleure cohésion des classes en comparaison avec les algorithmes locaux.



Fig V.2.a: Algorithme EM Gibbsien



Fig V.2.b: Algorithme ICE Gibbsien

V.2.c Segmentation de la zone forestière

Nous envisageons de comparer dans ce paragraphe les performances des algorithmes monospectraux et multispectraux en segmentant l'image multispectrale de la forêt. Nous estimons tout d'abord les paramètres de l'image par les algorithmes EM et ICE avant d'appliquer le critère de classification de Bayes.

V.2.c.i Estimation des paramètres

L'ensemble des paramètres estimés par les algorithmes sont résumés dans le tableau V.3. Ces paramètres sont obtenus en utilisant comme initialisation les paramètres correspondants à l'image segmentée par l'algorithme SEM adaptatif. Les algorithmes convergent approximativement vers les mêmes paramètres. Notons la plus faible valeur de l'énergie estimée dans le cas multispectral en comparaison avec celle obtenue dans le cas monospectral.

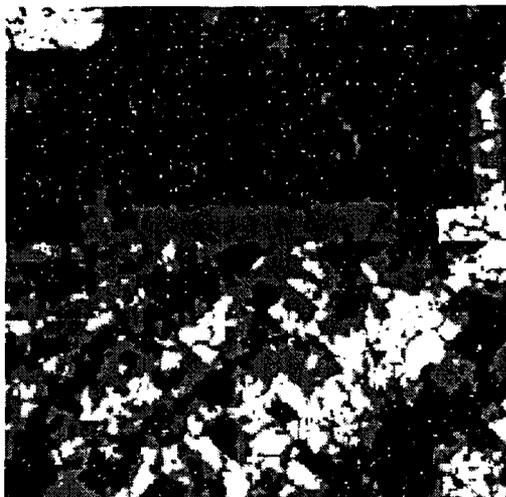


Fig V.3.a: Algorithme EM Gibbsien

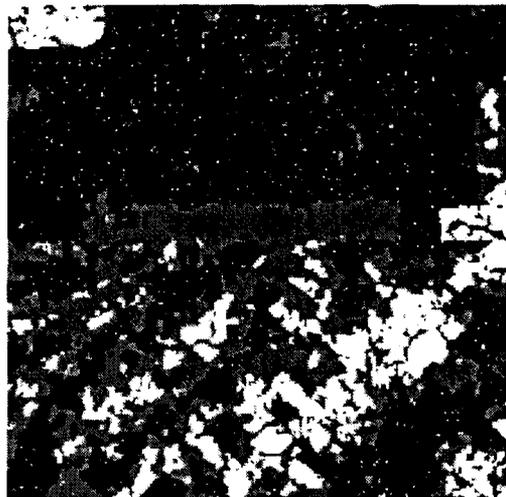


Fig V.3.b: Algorithme ICE Gibbsien

V.3 Conclusion

Nous avons présenté au cours de ce chapitre deux algorithmes de segmentation statistiques des images radar multispectrales basées sur une approche markovienne du champ de classes. L'algorithme EM gibbsien que nous avons construit a permis d'approcher la solution des équations de la vraisemblance dans le cas markovien et multispectral. L'algorithme ICE gibbsien a constitué une autre approche pour l'estimation des différents paramètres en cherchant itérativement l'espérance conditionnelle de leurs estimateurs respectifs. L'application des deux algorithmes sur une image de test a permis de valider la robustesse des méthodes employées. Leur application sur quelques images multispectrales réelles de type SIR-C/X-SAR et de type JERS-ERS1 a de nouveau confirmé la supériorité de la modélisation globale markovienne vis à vis de la modélisation locale.

CONCLUSION GENERALE

CONCLUSION GENERALE

Notre travail a été consacré à la segmentation statistique des images radar. Nous avons traité séparément le problème de segmentation des images radar monospectrales et celui des images radar multispectrales.

Après avoir introduit le modèle hiérarchique généralisé pour la modélisation statistique des images radar, nous avons modélisé le champ de classes en distinguant une approche locale qui admet l'indépendance de ses composantes et une approche globale markovienne qui, quant à elle, prend en compte leur dépendance.

Le processus de retrodiffusion au sein de l'ensemble des textures de l'image a été modélisé par l'ensemble des distributions unimodales du système de Pearson. Cette modélisation a permis de tenir compte de la forme variable de la distribution en passant d'une texture à une autre de la même image. Pour estimer les paramètres et la loi des distributions de ce système, nous avons proposé une méthode d'estimation basée sur la maximisation de la vraisemblance. La comparaison de cette méthode avec celle des moments a montré sa meilleure robustesse pour les échantillons de faibles tailles.

Nous avons ensuite examiné le problème de la segmentation des images radar monospectrales dans le cadre d'une approche locale du champ de classes. La solution de ce problème a nécessité l'estimation des paramètres relatifs à la loi du champ de classes et ceux relatifs à la loi du champ de mesure. Les algorithmes que nous avons développés ont permis une estimation non supervisée de l'ensemble de ces paramètres. Les résultats obtenus sur une image de test et des images radar ont par ailleurs montré les limitations de l'algorithme EM vis à vis des deux algorithmes SEM et ICE et la supériorité de la modélisation locale non stationnaire vis à vis de la modélisation locale stationnaire.

La segmentation des images monospectrales dans le cadre d'une modélisation globale pour le champ de classes a été réalisée en développant une version des algorithmes EM et ICE, valable pour un champ de classes markovien. Les résultats obtenus par

ANNEXES

Annexe 1: Equations de vraisemblance dans le cas monospectral

Nous désirons résoudre les équations de vraisemblance pour un mélange de distributions, dans le cas où les composantes relatives au champ de classes sont indépendantes.

Soit:

$$y = (y_s)_{s \in S} \quad (A1.1)$$

la réalisation particulière du champ des intensité

La loi de chaque variable aléatoire Y_s est donnée le mélange:

$$f_{Y_s} = \sum_{k=1}^K \Pi_k f_{Y_s}^{X_s = \omega_k} \quad (A1.2)$$

L'estimation des paramètres de ce mélange par la méthode du maximum de vraisemblance revient à maximiser la quantité:

$$L_y = \sum_{s \in S} \text{Log} f_{Y_s}(y_s) \quad (A1.3)$$

avec la contrainte à l'égalité donnée par:

$$\sum_{k=1}^K \Pi_k = 1 \quad (A1.4)$$

Pour tenir compte de la contrainte à l'égalité (A1.4), nous introduisons le coefficient de Lagrange λ dans l'expression de la vraisemblance:

$$L_y = \sum_{s \in S} \text{Log} f_{Y_s}(y_s) + \lambda \left(1 - \sum_{k=1}^K \Pi_k\right) \quad (A1.5)$$

La maximisation de (A1.5) par rapport au paramètre Π_k permet d'écrire le paramètre λ sous la forme:

$$\lambda = \sum_{s \in S} \frac{f_{Y_s}^{X_s = \omega_k}(y_s)}{f_{Y_s}(y_s)} \quad (A1.6)$$

Annexe 2: Initialisation des paramètres dans les méthodes locales

La méthode de l'histogramme cumulé permet de fournir une estimation de la moyenne et de la variance des distributions marginales. Cette méthode d'initialisation est généralement employée pour des distributions suivant une loi normale.

Soit H l'histogramme cumulé de l'image, la moyenne μ_{1k} de la classe ω_k est estimée par l'abscisse correspondant à la valeur $1/(k+1)$ de H , soit:

$$\mu_{1k} = H^{-1}\left(\frac{1}{k+1}\right) \quad (\text{A2.1})$$

Les variances sont supposées les mêmes pour chacune des K classes de l'image. Elles sont estimées par:

$$\sigma^2 = (\mu_1 - \mu_2)^2 \quad (\text{A2.2})$$

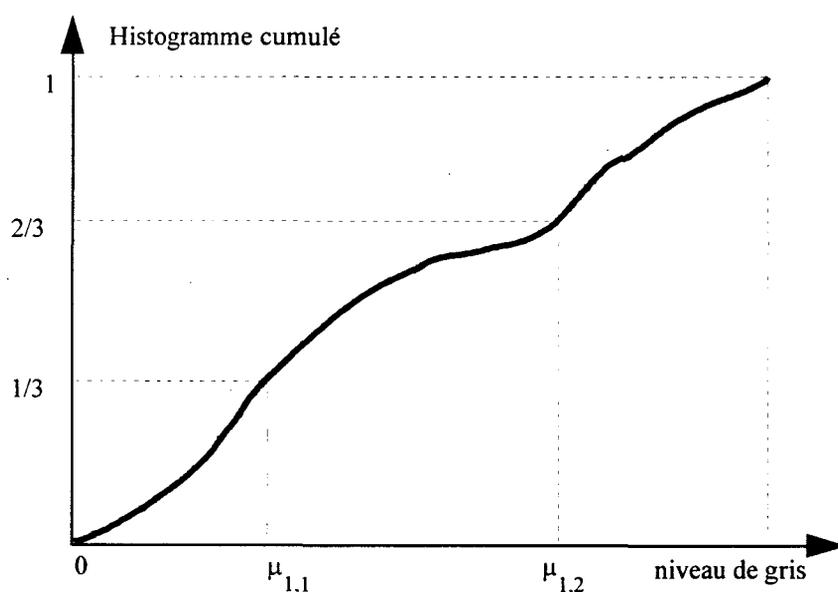


Fig A2.1: Méthode de l'histogramme cumulé

Annexe 3: Equations de vraisemblance dans le cas multispectral

L'estimation des paramètres par la méthode du maximum de vraisemblance revient à maximiser la quantité:

$$L = \sum_{s \in S} \text{Log} f_{Y_s}(y_s) \quad (\text{A3.1})$$

où: $Y_s = (Y_{1,s}, \dots, Y_{M,s})^t$

avec la contrainte à l'égalité:

$$\sum_{k=1}^K \Pi_k = 1 \quad (\text{A3.2})$$

D'autre part, les probabilités a posteriori sont de la forme:

$$P_{X_s}^{y_s}(\omega_k) = \frac{\Pi_k \left(\prod_{m=1}^M f_{m,k}(y_{m,s}) \right)}{f_{Y_s}(y_s)} \quad (\text{A3.3})$$

Les distributions $f_{m,k}$ dépendent des paramètres $\theta_{k,m}$. Par dérivation de (A3.1), nous obtenons les MxK équations suivantes:

$$\sum_{s \in S} \frac{\Pi_k \left(\prod_{j \neq m} f_{j,k}(y_{j,s}) \right) \frac{\partial f_{m,k}(y_{m,s})}{\partial \hat{\theta}_{m,k}}}{f_{Y_s}(y_s)} = 0 \quad (\text{A3.4})$$

La résolution directe de ces équations est impossible analytiquement, d'où le recours à des méthodes numériques itératives : algorithmes EM et SEM.

Pour tenir compte de la contrainte à l'égalité (A3.2), nous introduisons le coefficient de Lagrange λ dans l'expression de la vraisemblance:

$$L = \sum_{s \in S} \text{Log} f_{Y_s}(y_s) + \lambda \left(1 - \sum_{k=1}^K \Pi_k \right) \quad (\text{A3.5})$$

Nous obtenons alors l'expression suivante pour Π_k

REFERENCES BIBLIOGRAPHIQUES

Bibliographie

- [ABH92] O. Allagnat, J.M.Boucher, D.C.He, W. Pieczynski; *Hidden Markovfields and unsupervised segmentation of images*; Proc. International Conference on Pattern Recognition, La Haye, Pays Bas, septembre 1992.
- [Ant 86] P. Anthouard; *Imagerie RADAR à ouverture synthétique: applications aéoroportées et spatiales*; Premier colloque image, traitement, synthèse, technologie et applications, Biarritz mai 1984.
- [All 84] T.D. Allan; *Seasat measurement of wind and waves on selected passes over Jasin*; T.D Allan Ed, 1983, pp. 147-154.
- [ARR 81] W.R. Alpers, D.R. Ross and C.L. Ruffenach; *On the dectectability of ocean surface waves by real and synthetic aperture radar*; In Journal of geophysical research, Vol. 87, C5, pp. 6481-6498, July 1981.
- [CD 86] G. Celeux and J. Diebolt; *L'algorithme SEM: Un algorithme d'apprentissage probabiliste pour la reconnaissance de mélange de densités*; Revue de statistique appliquée, Vol. 34, N° 2, 1986.
- [Cha89] B. Chalmond; *An iterative Gibbsian technique for reconstruction of m-array images*; in Pattern recognition, Vol. 22, N° 6, pp. 747-761, 1989.
- [DD88] P.A Devijver and M. Dekesel; *Champs aléatoires de Pickard et modélisation d'images digitales*; Traitement de signal, Vol. 5, N° 5, 1988.
- [DH 73] R. Duda and P. Hart; *Pattern Classification and Scene Analysis*; Wiley Interscience 1973.
- [DJ 89] R.C. Dubes, A.K. Jain; *Random fieald models in image analysis*; Journal of applied Statistics, Vol.16, N°2, 1989.
- [DP 96] M. Craig Dobson and Leland E. Pierce; *Knowledge-Based Land-Cover Classification Using ERS1/JERS-1 SAR Composites*; in IEEE Trans. on Geoscience and Remote Sensing; Vol. 34, N°1, pp. 86-99, January 1996.

- [HD 96] Z. S. Haddad and P. Dubois; *Bayesian estimation of soil parameters from Radar backscatter data*; IEEE Transactions on Geoscience and Remote Sensing, Vol. 34, N°1, pp. 76-82, January 1996.
- [JK 70] Johnson and Kotz; *Distributions in Statistics: Continuous Univariate Distributions*; pp. 9-15. Wiley, 1970.
- [JS 96] Y. Jhung and Ph.H Sawain; *Bayesian contextual classification based on modified M-Estimates an Markov fields*; IEEE Transactions on Geoscience and Remote Sensing, Vol. 34, N°34, pp. 67-75, January 1996.
- [KDH88] P.A. Kelly, H. Derin and K.D. Hartt; *Adaptive segmentation of speckled images using a hierarchial model*; IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 36 N°10, pp. 1628-1641, October 1988.
- [Lau 89] H. Laur; *Analyse d'images radar en télédétection: Discriminateur radiométriques et texturaux*; Thèse de doctorat de Université Paul Sabatier, Toulouse, 1989.
- [LD 89] S. Lakshmann and H. Derin; *Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing*; IEEE Transactions on Pattern Analysis and machine Intelligence, Vol PAMI-11, N°8, AUG 1989.
- [Lev95] E. Levitan; *Image-Modeling Gibbs Priors*; IEEE Graphical Models and Image Processing , Vol. 57, N°2, pp. 117-130, March 1995
- [LLN90] A. Lopes, H. Laur and E. Nezry; *Statistical distribution and texture in multilook and complex SAR images*; in Proceeding of IGARR'S symp, Washington, Vol. 3, pp. 2427-2430, May 90.
- [LNT92] A. Lopes, E. Nezry, R. Touzi and H. Laur; *Adaptive processing of multilook complex SAR images*; In Proceeding of IGARR'S, Clear Lake (USA), Vol. 3, pp. 890-892, 26-29 May 1992.
- [LNT93] A. Lopes, E. Nezry, R. Touzi and H. Laur; *Structure detection and statistical adaptive speckle filtering in SAR images*; International Journal of Remote Sensing, Vol. 3, pp. 2409-2412, 1993.

- [Pie94b] W. Pieczynski; *Champs de Markov cachés et estimation conditionnelle itérative*; Traitement du Signal, Vol. 11, N° 2, pp. 141-153, 1994.
- [Pos87] J.G. Postaire; *De l'image à la décision*; DUNOD informatique, 1987.
- [PP 95] A. Peng and W. Pieczynski; *Adaptive mixture estimation and unsupervised local Bayesian image segmentation*; Graphical Model and Image Processing, Vol. 57 N° 5, pp. 389-399, Sept 1995.
- [Quel93] H. Ch. Quelle; *Segmentation Bayésienne non supervisée en imagerie RADAR*; Thèse de doctorat de l'université de Rennes, 1993.
- [QBP92a] H. Ch. Quelle, J.M Boucher and W. Pieczynski; *Adaptive parameter estimation and unsupervised image segmentation*; In Proceeding of IGARSS'92.
- [QBP92b] H. Ch. Quelle, J.M Boucher and W. Pieczynski; *Local parameter estimation and unsupervised segmentation of SAR images*; IEEE Proceeding of IGARSS92, pp. 1111-1113.
- [QWK96] P. Queru, F. Wattrelot, L. Khaterine, J.P. Dubus; *Segmentation des cellules endothéliales de la cornée par programmation dynamique*; Innovation et Technologie en Biologie et Médecine, Vol. 17 N° 3, pp. 177-193, Juin 1996.
- [RP 89] Y. Rungsunseri, W. Pieczynski et C. Roux, *Stratégie contextuelle et extraction des primitives pour la segmentation des images multispectrales*. GRETSI Juan Les Pins 1989.
- [Sch76] Schroeder, *Analyse d'un mélange de distributions de même type*. Revue de Statistique Appliquée, N° 1, pp. 24, 1976
- [SCD91] C. Soubiran, G. Celeux, J. Diebolt and C. Robert, *Analyse de mélanges gaussiens pour de petits échantillons: Application à la cinématique stellaire*, Rev Statistique Appliquées, N° 1, pp. 17-36, 1991
- [ST 96] A. Solberg and T. Taxt, *A Markov random field model classification of multisource satellite imagery*, IEEE Trans. on Geoscience and Remote Sensing, Vol. 34, N° 1, pp. 100-113, January 1996
- [Tas 89] Ph. Tassi, *Méthodes statistiques*, pp. 117 - 223, Economica, 1989

NOTATIONS, TABLEAUX ET FIGURES

Liste des notations

Géométrie de l'image

- s Site particulier de l'image de coordonnées relatives (i,j).
 S Ensemble des sites de l'image.
 N Taille de l'image.
 M Nombre de canaux de l'image multispectrale.

Champs aléatoires définissant l'image

- ω_k Etiquette attribuée à la classe d'indice k.
 Ω Ensemble des classes ω_k .
 K Nombre d'éléments de l'ensemble de classes Ω .
 X Champ de classes.
 X_s Composante du champ de classes située au site s.
 Y Champ des intensités.
 Y_s Composante du champ des intensités au site s.
 $Y_{m,s}$ Composante du champ des intensités située au site s relativement au canal m.
 ξ Champ de mesure.
 f_{Y_s} Densité de probabilité de la composante Y_s du champ Y.
 f_Y Densité de probabilité du champ Y.
 $f_{Y_s}^{x_s}$ Densité de la composante Y_s conditionnelle à x_s .
 $f_{Y_s}^X$ Densité de la composante Y_s conditionnelle à la réalisation $X=x$ du champ de classes.
 $\Pi_{k,s}$ Probabilité a priori de la classe ω_k au site s.
 Π_k Probabilité a priori de la classe ω_k indépendamment de la position du site.
 $P_{X_s}^{y_s}$ Probabilité a posteriori de X_s sachant y_s .

Liste des tableaux

Chapitre I

Tableau I.1	Paramètres théoriques des distributions univariées utilisées.
Tableau I.2	Valeurs du biais pour la distribution de première espèce.
Tableau I.3	Valeurs du biais pour la distribution de seconde espèce.
Tableau I.4	Valeurs de la variance pour la distribution de première espèce.
Tableau I.5	Valeurs de la variance pour la distribution de seconde espèce.

Chapitre II

Tableau II.1	Paramètres théoriques du mélange associé à l'image de synthèse.
Tableau II.2	Paramètres initiaux du mélange.
Tableau II.3	Paramètres du mélange estimés par les algorithmes locaux
Tableau II.4	Taux d'erreur obtenus par les différents algorithmes.
Tableau II.5	Paramètres de l'image de Brest estimés par les algorithmes locaux.
Tableau II.6	Paramètres de l'image de la forêt estimés par les algorithmes locaux.

Chapitre III

Tableau III.1	Paramètres de l'image de synthèse estimés par les algorithmes globaux.
Tableau III.2	Taux d'erreur obtenus par les algorithmes globaux.
Tableau III.3	Paramètres de l'image de Brest estimés par les algorithmes globaux.
Tableau III.4	Paramètres de l'image forestière par les algorithmes globaux.

Chapitre IV

Tableau. IV.1	Paramètres théoriques de l'image de synthèse multispectrale.
Tableau. IV.2	Paramètres de l'image de synthèse multispectrale estimés par les algorithmes locaux.

Liste des figures

Chapitre I

- Figure I.1 Principe d'un radar imageur à visée latérale.
Figure I.2 Correspondance pixel - cellule de résolution.
Figure I.3 Antenne du SIR-C/X-SAR.
Figure I.4 Schéma simplifié de l'antenne SIR-C/X-SAR.
Figure I.6 Description du modèle hiérarchique dans le cas multispectral.
Figure I.6.a Voisinage d'ordre 1.
Figure I.6.b Voisinage d'ordre 2.
Figure I.7.a Cliques associés au voisinage d'ordre 1.
Figure I.7.b Cliques associés au voisinage d'ordre 2.
Figure I.8 Champs markovien simulé par l'échantillonneur de Gibbs.
Figure I.9 Graphe de Pearson.
Figure I.10 Distribution bêta de première espèce en U.
Figure I.11 Distributions bêta de première espèce en L et en J.
Figure I.12 Distributions bêta de première espèce en cloche.
Figure I.13 Distributions bêta de seconde espèce en cloche et en L.
Figure I.14 Distribution de type 4.
Figure I.15 Positions de la distribution à l'intérieur d'une surface.
Figure I.16 Positions de la distribution au voisinage d'une droite ou d'un point.

Chapitre II

- Figure II.1.a Echantillons représentatifs des deux classes.
Figure II.1.b Echantillons représentatifs d'une seule classe.
Figure II.1.c Un seul échantillon est représentatif d'une classe.
Figure II.2 Fort recouvrement avec des probabilités a priori voisines.
Figure II.3 Fort recouvrement avec des probabilités a priori disproportionnées.
Figure II.4 Image de la cible originale.
Figure II.5 Image de la cible bruitée.

Chapitre III

- Figure III.1.a Image de la cible segmentée par l'algorithme EM gibisien.
- Figure III.1.b Image de la cible segmentée par l'algorithme ICE gibisien.
- Figure III.2.a Image de Brest segmentée par l'algorithme EM gibisien.
- Figure III.2.b Image de Brest segmentée par l'algorithme ICE gibisien.
- Figure III.3.a Image de la forêt segmentée par l'algorithme EM gibisien.
- Figure III.3.b Image de la forêt segmentée par l'algorithme ICE gibisien.

Chapitre IV

- FigureIV.1.a Canal 1 de l'image de synthèse.
- FigureIV.1.b Canal 2 de l'image de synthèse.
- FigureIV.1.c Canal 3 de l'image de synthèse.
- Figure IV. 2 Evolution de la vraisemblance de l'image de synthèse.
- Figure IV.3.a Image de synthèse segmentée par l'algorithme EM stationnaire.
- Figure IV.3.b Image de synthèse segmentée par l'algorithme SEM stationnaire.
- Figure IV.3.c Image de synthèse segmentée par l'algorithme ICE stationnaire.
- Figure IV.3.d Image de synthèse segmentée par l'algorithme adaptatif.
- Figure IV.4.a Canal C de l'image SIR-C/X-SAR.
- Figure IV.4.b Canal L de l'image SIR-C/X-SAR.
- Figure IV.4.c Canal X de l'image SIR-C/X-SAR.
- Figure IV. 5 Evolution de la vraisemblance de l'image SIR-C/X-SAR.
- Figure IV.6.a Image SIR-C/X-SAR segmentée par l'algorithme EM stationnaire.
- Figure IV.6.b Image SIR-C/X-SAR segmentée par l'algorithme SEM stationnaire.
- Figure IV.6.d Image SIR-C/X-SAR segmentée par l'algorithme SEM adaptatif.
- Figure IV.7.a Canal JERS de la zone forestière.
- Figure IV.7.b Canal ERS1 de la zone forestière.
- Figure IV.8.a Image de la forêt segmentée par l'algorithme EM stationnaire.
- Figure IV.8.b Image de la forêt segmentée par l'algorithme SEM stationnaire.
- Figure IV.8.c Image de la forêt segmentée par l'algorithme ICE stationnaire.
- Figure IV.8.d Image de la forêt segmentée par l'algorithme SEM adaptatif.

SUMMARY

Our work deals with statistical segmentation of SAR images. We propose a generalised hierarchical model which use the Pearson system to modelize the backscattering from each type of observed nature. We introduce the notion of a generalised mixture and propose methods for estimating it, with applications to unsupervised statistical image segmentation. We show that classical mixture estimation algorithms can be adapted to such situation once as we dispose of a method of recognition of each component separately. We develop some algorithms using likelihood maximisation and iterative conditional expectation to estimate the mixture components. We apply the algorithms proposed to the problem of unsupervised Bayesian image segmentation of both monospectral and multispectral images.

Key words: radar images, statistical modelization, Pearson system, mixture distributions, EM, SEM and ICE algorithms.

RESUME

Cette thèse est consacrée à la segmentation statistique non supervisée des images radar. L'originalité de notre étude réside dans l'adoption du système de distributions de Pearson pour modéliser les scènes naturelles homogènes. Nous généralisons les méthodes d'estimation employées pour un mélange de distributions de même loi à un mélange de distributions de lois différentes et appartenant au système de Pearson. Des algorithmes d'estimation basés sur la maximisation de la vraisemblance et sur la notion de l'espérance conditionnelle ont été développés afin d'accomplir cette tâche. Les méthodes d'estimation développées pour les images radar monospectrales ont été généralisées pour le cas des images radar multispectrales. Nous étudions la robustesse des algorithmes proposés à travers la segmentation d'images de synthèse et d'images radar de natures différentes.