

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'enseignement Supérieur et de la Recherche Scientifique

Université de MOSTAGANEM

Faculté des Sciences et Sciences de l'Ingénieur

Mémoire de magister

PRESENTE A L'UNIVERITE DE MOSTAGANEM

Spécialité: Electronique

Option : Signaux & Applications

Par

M^{elle} MEHIDI Aicha

Thème

**Analyse d'images par champs de
Markov**

Application mammographie

Soutenu le devant le jury :

● **Président:**

● **Rapporteurs:** Mme Malika MIMI

Professeur

Université de Mostaganem

Mr Abdelmalik TALEB AHMED

Professeur

Université de Valenciennes France

● **Examineurs :**

**Promotion
2008 -2009**

*A mon très cher père, qui m'a beaucoup encouragé,
A ma très chère mère pour son sacrifice depuis mon
enfance,*

*A mes frères et mes sœurs et à tout qui porte le nom
MEHIDI,*

*Et à toutes les personnes qui m'ont aidé de près ou
de loin.*

Remerciements

Cette petite portion de mémoire est consacrée à remercier ceux que sans leurs aide , leur soutien et encouragements, ce travail n'aurait jamais vu le jour, nos sincères remerciements s'adressent à :

Mes parents, mes frères, pour leurs aide et leur soutien durant la préparation de ce mémoire.

Je remercie particulièrement prof. Mme Malika MIMI et prof. Mr Abdelmalik TALEB AHMED à avoir accepté de m'encadrer, m'orienter et m'aider à concevoir ce mémoire de Magister depuis le début de notre travail.

Je tiens également à remercier tous les membres de jury de m'avoir accepté de juger mon travail.

Mes remerciements s'adressent aux membres du laboratoire signaux et applications, Dr. Mme Naima RACHID et aux étudiants de ma promotion de Magister.

Je n'oublie pas à remercier tous mes amis et ceux qui par leurs aides, par leurs conseils, leurs discussions, ou leurs encouragements m'ont aidé à terminer ce travail.

Analyse d'images par champs de Markov

Présenté par : M^{elle} MEHIDI Aicha

Thèse Magister

- **Page de garde**
- **Dédicace**
- **Remerciement**
- **Résumé**
- **Abstract**
- **Sommaire**
- **Table de matière**
- **listes des tables**
- **listes des figures**
- **listes des notations**
- **INTRODUCTION GENERALE**
- **Chapitre I : POSITION DU PROBLEME**
- **Chapitre II : MODELISATION MARKOVIENNE**
- **Chapitre III : SEGMENTATION NON SUPERVISEE D'IMAGES**
- **Chapitre IV : SIMULATION ET DISCUSSION DES RESULTATS**
- **Conclusions & perspective**
- **Annexes**
- **Bibliographies.**

SOMMAIRE

Résumé	A
Abstract	C
Listes des tableaux	E
Listes des figures	G
Liste des symboles	J
Introduction Générale	1
Chapitre I : Position du problème	6
Chapitre II : Modélisation markovienne	21
Chapitre III : Segmentation non supervisée d'image	53
Chapitre IV : Simulation et discussion des résultats	81
Chapitre V : Conclusion Générale	95
Annexe	100
Bibliographie	105

Table des matières

Résumé	A
Abstract	C
Listes des tableaux	E
Listes des figures	G
Liste des symboles	J
Introduction Générale	1
Chapitre I	6
I.1. Introduction	7
I.2. Segmentation d'images	7
I.2.1. Principe de segmentation	8
I.2.2. Quelques méthodes de segmentation	9
I.2.2.1. Le seuillage	10
I.2.2.2. La croissance de régions	10
I.2.2.3. La coalescence	10
I.2.2.4. Support vecteur machine (SVM)	11
I.2.2.5. L'analyse factorielle discriminante	12
I.2.2.6. Les réseaux de neurones artificiels	12
I.2.2.7. Les modèles déformables	13
I.2.2.8. Le template	13
I.2.2.9. Autres méthodes	14
I.3. Les images à traiter	14
I.3.1. Principe de la mammographie	14
I.3.2. Les caractéristiques des images mammographies	16
I.4. Généralités sur les champs de Markov cachés	18
I.5. Conclusion	20
Chapitre II	21
II.1. Introduction	22

II.2. Champ de Markov et distribution de Gibbs	22
II.2.1. Système de voisinage et cliques	22
II.2.1.1. Système de voisinage	22
II.2.1.2. Clique	23
II.2.2. Champs de Markov aléatoires	25
II.2.3. Champs de Gibbs	25
II.2.4. Equivalence champ de Markov – champ de Gibbs	26
II.2.5. Modèles markoviens classiques	27
II.2.5. 1. Modèle d’Ising	27
II.2.5. 2. Modèle de Potts	28
II.3. Simulation d’un champ de Markov	29
II.3.1. Algorithme de Métropolis	29
II.3.2. Echantillonneur de Gibbs	30
II.3.3. Conclusion	33
II.4. Modélisation markovienne	33
II.4.1. Champs de Markov cachés	33
II.4.2. Le graphe d’un champ de Markov cachée	35
II.4.3. Champs de Markov cachés à bruit indépendant	35
II.4.4. Champs de Markov cachés en segmentation d’images	37
II.4.5. Spécification de $U(x/y)$ pour le problème de segmentation d’image	38
II.4.6. La réalisation d’un modèle de champ de Markov caché	39
II.4.7. Discussion	42
II.4.7.1. Role du choix des cliques	42
II.4.7.2. Role de la constante de couplage β	42
II.5. Estimation bayésienne	43
II.5.1. Estimateur Maximum A Posteriori (MAP)	44
II.5.2. Estimateur du Mode Conditionnel Itéré (ICM)	45
II.5.3. Estimateur du Mode des Marginales à posteriori (MMP)	46
II.5.4. Comparaison	47
II.6. Conclusion	52

Chapitres III :	53
III.1. Introduction	54
III.2. Synoptique de la segmentation	54
III.3. Segmentation non supervisée	55
III.3.1. Initialisation	55
III.3.2. Estimation des paramètres	56
III.3.2.1. Principe de l'algorithme ICE	57
III.3.2.2. L'algorithme ICE	58
III.3.2.3. Critère d'arrêt	60
III.3.3 L'algorithme MPM pour la segmentation d'image	60
III.4. Critères d'évaluation	61
III.4.1. Le rapport signal sur bruit crête (PSNR)	61
III.4.2. Matrice de confusion	62
III.5. Résultats sur d'images synthétiques	65
III.5.1. Présentation des resultants	66
III.5.2. Analyse et discussion	67
III.6. Conclusion	80
 Chapitre IV :	 81
IV.1. Introduction	82
IV.2. Critère d'évaluation d'un résultat de segmentation	82
IV.2.1. Définition du critère	82
IV.2.2. Valeur de l'uniformité intraclasse normalisée	83
IV.3. Résultats expérimentaux	83
IV.3.1. Présentation des résultats	84
IV.3.2. Analyse et discussion	85
IV.4. Conclusion	94
 Chapitre V : Conclusion Générale	 95
V.1. Récapitulatif	96
V.2. Perspectives	98

Annexe	100
Bibliographie	105

RÉSUMÉ



La mammographie permet de visualiser le sein ainsi que d'éventuelles lésions à son niveau. Elle est actuellement le meilleur examen de dépistage du cancer du sein. En effet, la mammographie permet de détecter, chez une femme qui n'a aucun symptôme, des tumeurs très petites non perçues à la palpation. C'est également un examen de diagnostic que le médecin ou le gynécologue prescrit quand il détecte à la palpation une grosseur anormale ou si la patiente présente des symptômes persistants tels qu'une induration du sein, une augmentation anormale de volume, un nodule, une rougeur de la peau, un écoulement anormal du mamelon ...

L'image représente une richesse inestimable vu la diversité des informations qu'elle recèle en elle à travers ces nombreux contours et détails. L'image médicale est donc une source de diagnostic très importante pour le médecin traitant. Les images obtenues à partir des appareils numériques, doivent être donc interprétées correctement.

Notre projet s'inscrit dans ce but afin de proposer une procédure de segmentation non supervisée, dont le formalisme s'appuie sur la théorie des champs de Markov cachés. L'originalité de cette procédure est qu'elle prend en compte des informations structurelles sous la forme de contraintes spatiales flexibles. Les premiers résultats sont satisfaisants.

Mots clés: Les champs de Markov cachés, Echantillonneur de Gibbs, Algorithme ICE, Algorithme MPM, Segmentation non supervisée, Image mammographie

ABSTRACT



The mammography permits to visualize the breast as well as of possible lesions to his/her level. It is currently the best exam of tracking of the breast cancer. Indeed, the mammography permits to detect, at a woman whom has no symptom, of the tiny tumors non discerned to the palpation. It is also an exam of diagnosis that the physician or the gynecologist prescribes when he detects to the palpation an abnormal size or if the patient presents obstinate symptoms as an induration of the breast, an increase abnormal of volume, a nodule, a redness of skin, an out-flow abnormal of the nipple...

The picture represents an inestimable wealth seen the diversity of the information that it conceals in it through these numerous contours and details. The medical picture is therefore a source of very important diagnosis for the physician. The pictures gotten from the numeric devices must be interpreted therefore correctly.

Our project appears in this goal in order to present an unsupervised segmentation method which the formalism is relied on hidden Markov field. The originality of this method is, it takes into account structural information processed as flexible spatial. Our first results are very satisfactory.

Key words: Hidden Markov field, Gibbs sampler, Algorithm ICE, Algorithm MPM, unsupervised segmentation, Image mammography.

LISTES DES TABLEAUX

Chapitre I :	Page
Chapitre II :	
Chapitres III :	
Table 1 : Modèle de matrice de confusion à 4 entrées ($N = \sum_i a_{i+} = \sum_j a_{+j}$ est le nombre total de pixels dans les échantillons).	63
Table 2 : Matrice de confusion pour le résultat de la classification dans figure (d) de l'image synthétique (1) avec flou.	68
Table 3 : L'évaluation de résultat de segmentation de l'image synthétique (1) avec flou.	68
Table 4 : Matrice de confusion correspondant à la classification faite par champ de Markov caché sur l'image synthétique (1) avec flou avec le nombre de Pixel.	69
Table 5 : Matrice de confusion du Tab-III.4 avec toutes les valeurs exprimées en pourcentage.	70
Table 6 : Matrice de confusion pour le résultat de la classification dans figure (d) de l'image synthétique (1) sans flou.	71
Table 7 : L'évaluation de résultat de segmentation de l'image synthétique (1) sans flou.	71
Table 8 : Matrice de confusion correspondant à la classification faite par champ de Markov caché sur l'image synthétique (1) sans flou avec le nombre de Pixel.	72
Table 9 : Matrice de confusion du Tab-III.8 avec toutes les valeurs exprimées en pourcentage.	73
Table 10 : Matrice de confusion pour le résultat de la classification dans figure (d) de l'image synthétique (2) avec flou.	74
Table 11 : L'évaluation de résultat de segmentation de l'image synthétique (2) avec flou.	74
Table 12 : Matrice de confusion correspondant à la classification faite par champ de Markov caché sur l'image synthétique (2) avec flou avec le nombre de Pixel.	75

Table 13: Matrice de confusion du Tab-III.12 avec toutes les valeurs exprimées en pourcentage.	76
Table 14 : Matrice de confusion pour le résultat de la classification dans figure (d) de l'image synthétique (2) sans flou.	77
Table 15 : L'évaluation de résultat de segmentation de l'image synthétique (2) sans flou.	77
Table 16 : Matrice de confusion correspondant à la classification faite par champ de Markov caché sur l'image synthétique (2) sans flou avec le nombre de Pixel.	78
Table 17 : Matrice de confusion du Tab-III.16 avec toutes les valeurs exprimées en pourcentage.	79

Chapitre IV :

Table 1 : Caractéristiques des images utilisées.	84
Table 2 : Nombre d'itération de chaque algorithme.	86
Table 3 : L'évaluation de résultat de segmentation de l'image mammographie sain	87
Table 4 : Nombre d'itération de chaque algorithme.	88
Table 5 : L'évaluation de résultat de segmentation de l'image mammographie (1).	89
Table 6 : Nombre d'itération de chaque algorithme.	90
Table 7 : L'évaluation de résultat de segmentation de l'image mammographie (2).	91
Table 8 : Nombre d'itération de chaque algorithme.	92
Table 9 : L'évaluation de résultat de segmentation de l'image mammographie (3).	93

Chapitre V :

LISTES DES FIGURES

Chapitre I :	Page
Figure 1 : Les images de simulation. a) cas sain et b, c et d) présence d'anomalie.	17
 Chapitre II :	
Figure 1 : Relation de voisinage aux 4 ou 8 plus proches voisins.	23
Figure 2 : Cliques associées à un système de voisinage en 4 et en 8-connexité.	24
Figure 3 : Orientation d'un spin à l'état d'équilibre.	28
Figure 4 : Exemple de simulation d'un champ de Markov à 4 classes par échantillonneur de Gibbs pour différentes valeurs de β selon la probabilité a priori (4 pixels voisins).	31
Figure 5 : Exemple de simulation d'un champ de Markov à 4 classes par échantillonneur de Gibbs pour différentes valeurs de β selon la probabilité a priori (8 pixels voisins).	32
Figure 6 : La structure d'un modèle de champ de Markov caché.	35
Figure 7 : Exemple de simulation d'un champ de Markov cachée à 4 classes par échantillonneur de Gibbs pour différentes valeurs de β selon la probabilité a posteriori d'une Image mammographie (4 pixels voisins).	40
Figure 8 : Exemple de simulation d'un champ de Markov cachée à 4 classes par échantillonneur de Gibbs pour différentes valeurs de β selon la probabilité a posteriori d'une Image mammographie (8 pixels voisins).	41
Figure 9 : Résultats de classification par ICM pour l'image mammographie correspond à différentes valeurs de β (4 pixels voisins).	48
Figure 10 : Résultats de classification par ICM pour l'image mammographie correspond à différentes valeurs de β (8 pixels voisins).	49
Figure 11 : Résultats de classification par MPM pour l'image mammographie correspond à différentes valeurs de β (4 pixels voisins).	50
Figure 12 : Résultats de classification par MPM pour l'image mammographie correspond à différentes valeurs de β (8 pixels voisins).	51

Chapitre III :

Figure 1 : Organigramme de la procédure de segmentation d'images.	54
Figure 2 : (a) Image synthétique (1) originale; (b) Image bruitée avec flou et (c) Image bruitée sans flou.	65
Figure 3 : (a) Image synthétique (2) originale; (b) Image bruitée avec flou et (c) Image bruitée sans flou.	65
Figure 4 : Résultat de segmentation en niveau de gris de l'image synthétique (1) avec flou : (a) image originale, (b) image bruitée avec flou, (c) selon algorithme en Kmeans et (d) selon algorithme en champ de Markov caché	68
Figure 5 : Résultat de segmentation en niveau de gris de l'image synthétique (1) sans flou : (a) image originale, (b) image bruitée avec flou, (c) selon algorithme en Kmeans et (d) selon algorithme en champ de Markov caché.	71
Figure 6 : Résultat de segmentation en niveau de gris de l'image synthétique (2) avec flou : (a) image originale, (b) image bruitée avec flou, (c) selon algorithme en Kmeans et (d) selon algorithme en champ de Markov caché.	74
Figure 7 : Résultat de segmentation en niveau de gris de l'image synthétique (2) sans flou : (a) image originale, (b) image bruitée avec flou, (c) selon algorithme en Kmeans et (d) selon algorithme en champ de Markov caché.	77

Chapitre IV :

Figure 1 : Résultat de segmentation en niveau de gris de l'image mammographie sain.	86
Figure 2 : Résultat de segmentation en couleur de l'image mammographie sain.	87
Figure 3 : Résultat de segmentation en niveau de gris de l'image mammographie (1) 'présence d'anomalie '.	88
Figure4 : Résultat de segmentation en couleur de l'image mammographie (1) 'présence d'anomalie '.	89

Figure5 : Résultat de segmentation en niveau de gris de l'image mammographie (2) 'présence d'anomalie '.	90
Figure 6 : Résultat de segmentation en couleur de l'image mammographie (2) 'présence d'anomalie '.	91
Figure 7 : Résultat de segmentation en niveau de gris de l'image mammographie (3) 'présence d'anomalie '.	92
Figure 8 : Résultat de segmentation en couleur de l'image mammographie (3) 'présence d'anomalie '.	93

Chapitre V :

Figure 1 : Organigramme de la procédure de la détection des anomalies	98
---	-----------

LISTE DES SYMBOLES

I	Une image
R_i	Régions
α	Paramètre d'interaction
β	Paramètre associé aux singletons
η	Paramètres associés au bruit
C	Ensemble de cliques
c	Clique
$s = (i, j)$	Un point de coordonnées (i,j) d'une image.
δ	Fonction de Dirac
E	Espérance mathématique
γ	Fonction de partition
Ω	Ensemble de classes
ω_i	Classe
φ	Fonction potentielle
$P(\Omega)$	Ensemble des parties de Ω
$p(x)$	Probabilité
R	Ensemble des réels
S	Ensemble des pixels
N	La taille de S
s	Site ou pixel
U	Fonction énergie
Z	Constante de normalisation
V	Système de voisinage
V_s	Ensemble de voisins de s
X	Champ caché
X_s, Y_s	Variables aléatoires
Y	Champ observé
x	Réalisation de X .
y	Réalisation de Y .
f_k	Densité de probabilité conditionnelle.
μ_s	La moyenne.
σ_s	L'écart type.

k	La valeur de kappa
GU	la mesure d'uniformité intraclasse.
NU	la mesure d'uniformité intraclasse normalisée.
BI	Bruit indépendant
CG	Champ de Gibbs
CM	Champ de Markov
CMC	Champ de Markov Caché
DC	Données Complètes
DI	Données Incomplètes
ECI	Estimation conditionnelle Iterative
MAP	Maximum a posteriori
ICM	Mode Conditionnel Itéré
MPM	Mode des Marginales à posteriori
LBG	Linde-Buzo-Gray
SVM	Support vecteur machine
ART	Adaptive Resonance theory
CC	Cranio- Caudale
MLO	Medio- Laterale Oblique
MCMC	Monte Carlo Markov Chains

INTRODUCTION GENERALE



INTRODUCTION

Le travail présenté ici s'inscrit dans le contexte général de l'aide au traitement et à l'interprétation des données. Les applications concernent plus particulièrement l'imagerie médicale. Ceci est déroulé au sein de l'équipe «Signals and Systems», dont les activités de recherche concernent l'analyse d'images. Parmi les motivations du groupe «Signals and Systems», on trouve l'opération de segmentation des images mammographie par le champ de Markov caché.

De nos jours, le diagnostique par imagerie est un outil inestimable dans la médecine. L'image mammographie (1989), l'image à résonance magnétique nucléaire (RMN en 1971), le scanner hélicoïdal (CT en 1990), la tomographie (1927) et d'autres modalités d'imagerie médicale fournissent des moyens efficaces pour la configuration de l'anatomie d'un sujet, la détection d'anomalies, et à faire des mesures sans intervention directe. Ces technologies ont augmenté la connaissance des anatomies normaux et malades pour la recherche médicale. Ils sont un composant critique dans le diagnostic et la planification du traitement.

Avec la croissance de la taille et du nombre d'images médicales, l'utilisation d'ordinateurs pour faciliter leur traitement et leur analyse est devenue nécessaire. En particulier les algorithmes de segmentation d'image qui jouent un rôle crucial dans plusieurs applications d'imagerie médicales en automatisant la délimitation de structures anatomiques et d'autres régions d'intérêt.

Les méthodes de segmentations performantes varient largement selon l'application spécifique, la modalité de l'image et d'autres facteurs. Chaque modalité d'image a ses propres caractéristiques. Il n'y a actuellement aucune méthode de segmentation simple qui rapporte des résultats acceptables pour chaque image médicale. Les méthodes qui existent sont plus général et peuvent être appliqué à une variété de données. Cependant, les méthodes qui sont spécialisées à des applications particulières peuvent souvent réaliser la meilleure performance en prenant en considération des connaissances a priori. Le choix d'une approche appropriée à un problème de segmentation peut donc être un dilemme difficile.

La segmentation d'image est une opération d'autant plus importante qu'elle est présente à la base de bon nombre de chaînes d'analyse d'image. Elle consiste à découper les données en régions homogènes, chaque région étant repérée par une étiquette (numéro, couleur,...). Le travail réalisé au cours de cette thèse concerne l'étude et la mise en œuvre d'un modèle de segmentation à l'aide d'une approche markovienne.

Les champs de Markov sont des outils probabilistes assez largement utilisés en traitement d'images. Leur intérêt majeur est de pouvoir modéliser les dépendances spatiales des variables aléatoires, dont les réalisations modélisent les quantités observées ou recherchées, de façon relativement simple et se prêtant bien aux divers traitements. Divers modèles fondés sur les champs markoviens ont été proposés, parmi lesquels les champs de Markov cachés (CMC), très utilisés en segmentation d'images. Il est important d'abord de préciser que la modélisation par les champs de Markov cachés n'est pas en elle-même une méthode de segmentation, mais un modèle statistique qui peut être employé dans des méthodes de segmentation. On peut noter aussi que les variables cachées n'ont pas, a priori d'existence physique dans le phénomène observé; mais, elles sont avant tout utilisées pour créer des modèles flexibles. Cependant, après analyse des données en regard du modèle, elles trouvent souvent une interprétation concrète a posteriori [Dur 2003].

Les champs de Markov cachés permettent, grâce à leur structure, de prendre en compte les dépendances spatiales entre les différents pixels de l'image et d'incorporer une information a priori sur la segmentation [Pey 2001]. Une difficulté associée aux champs de Markov cachés est le choix approprié des paramètres qui contrôlent la force d'interactions spatiales. De plus, ce modèle exige d'habitude des algorithmes de temps de calcul intensifs.

Dans l'imagerie médicale, ils sont typiquement employés pour tenir compte du fait que la plupart des pixels appartiennent à la même classe comme leurs pixels voisins. Ils sont largement utilisés non seulement pour modéliser les classes de segmentation, mais aussi pour modéliser les inhomogénéités d'intensité qui peuvent apparaître par exemple dans des images mammographies. Ainsi, les champs de Markov cachés peuvent être utilisés pour modéliser l'existence des zones homogènes. Il s'agit donc du modèle approprié pour étudier la segmentation d'images avec un apprentissage non supervisé.

Ce manuscrit est constitué de quatre chapitres organisés comme suit :

- Le premier chapitre est introductif et présente la problématique dans sa généralité. Nous commençons par définir la nature et les objectifs de l'opération de segmentation, ainsi que les notations utilisées au cours des chapitres suivants et une présentation de différents modèles de segmentation qui ont suscité de nombreux travaux dans ce domaine au cours de ces dernières années. Nous introduisons également les images mammographie sur lesquelles nous appliquons le modèle de segmentation markovien. Ce chapitre s'achève par un aperçu très général sur les champs de Markov cachés.
- Le second chapitre présente le modèle de segmentation par les champs de Markov cachés. Au début du chapitre nous précisons les notations qui y sont attachées. Nous développons ensuite les champs markovien cachés pour la segmentation d'images. Le modèle que nous définissons ici met en jeu le contexte spatial de l'image. Les champs de Markov cachés permettent d'avoir une interprétation des interactions tant locales que globales dans une grille de variables aléatoires.
- Les problèmes de segmentation d'image et l'apprentissage non supervisé des paramètres pour les champs de Markov cachés sont l'objectif du troisième chapitre. Nous présentons deux algorithmes originaux. Pour le premier sujet, il s'agira du mode des marginales à posteriori (MPM) qui consiste à estimer les informations cachées. Quant à la seconde question, nous adapterons l'algorithme d'espérance conditionnelle itérative (ICE) à l'estimation des paramètres a priori et d'attaches aux données. Ici, c'est une étape supérieure et un niveau de difficulté accru, car il s'agit de réaliser le même objectif que la segmentation supervisée mais en estimant conjointement les paramètres. Dans ce chapitre, nos différents algorithmes seront testés sur des problèmes de classification non supervisée d'images synthétiques et pour évaluer nos résultats de segmentation, nous proposons d'exploiter deux critères d'évaluations : l'un le rapport signal sur bruit crête (PSNR). C'est un critère global, l'autre la matrice de confusion. C'est une matrice ou un tableau affichant les statistiques de la précision de classification d'une image, notamment le degré de classification erronée parmi les diverses classes.

- Au sein du quatrième chapitre, nous expérimentons les champs markoviens cachés que nous avons étudié aux chapitres précédents sur des données réelles. Nous traitons d'images mammographies. Puisque aucune vérité terrain n'est disponible pour ces images pour évaluer nos résultats de segmentation, nous offrons d'exploiter le critère d'évaluation de Nazif et Levine.

Ce manuscrit s'achève par un récapitulatif des résultats obtenus avec le champ de Markov caché et par l'évocation de quelques perspectives et extensions possibles de nos travaux que nous envisageons dans l'avenir.

CHAPITRE 1

POSITION DU PROBLÈME

Dans ce chapitre, nous allons décrire les problèmes abordés au cours de cette étude. Pour arriver à l'élaboration d'un algorithme il faut d'abord connaître le problème. Dans ce cas, Nous exposons la problématique concernant la segmentation d'images en générale en définissant la nature, les objectifs de l'opération de segmentation, et en passant en revue de quelques modèles développés au cours de ces dernières années dans le cadre de la segmentation d'images. Nous parlons brièvement des principales caractéristiques des images mammographies traitées au sein de ce travail. Nous terminons par un aperçu très général sur les champs de Markov cachés.

I.1. Introduction

La segmentation reste un des problèmes fondamentaux de l'analyse d'image. Elle consiste à diminuer considérablement le nombre d'entités présentes dans l'image. Ainsi, le rôle de la segmentation est, à partir de toutes les données, de retrouver l'information utile et de la coder sous une forme de liste de primitives, plus facilement exploitable par un futur processus d'interprétation. Ici, nous nous intéressons plus particulièrement à la segmentation en régions non supervisée d'images mammographies. Tous les paramètres de l'algorithme, ainsi que le nombre d'étiquettes, sont estimés. Nous utilisons des modèles probabilistes associés à la théorie bayésienne. Ils offrent un cadre mathématique cohérent pour l'extraction de primitives en introduisant conjointement : des connaissances a priori sur les primitives et une modélisation statistique du lien entre les observations et les primitives. Les algorithmes fondés sur une modélisation statistique conjointe des régions et des niveaux de gris, notamment ceux s'appuyant sur les champs de Markov cachés, reposent sur la minimisation énergie. Cette fonction prend simultanément en compte la vraisemblance de l'appartenance du pixel à une région considérant son niveau de gris, et les régions auxquelles appartiennent les pixels voisins. Cette fonction effectue un compromis entre la fidélité à l'image initiale et la régularité des régions segmentées.

I.2. Segmentation d'images

La segmentation d'images est une opération présente à la base de nombreuses chaînes de traitements et d'interprétation de données. On distingue en général deux niveaux dans l'analyse d'images [CoP 1995] : les traitements de bas niveau et ceux de haut niveau. Dans les traitements de bas niveau, il s'agit d'extraire des attributs de l'image à partir des grandeurs physiques observées, que ces attributs soient réalisés ou non. Elles regroupent les approches qui essaient de décrire l'image, telles que la compression, la restauration, et la détection de contours ... etc. C'est à dire une description de l'image par rapport aux structures qui la composent. Une fois ces attributs de bas niveau extraits, ils peuvent être interprétés par un traitement de haut niveau. La reconnaissance d'objets et la reconstruction du mouvement s'inscrivent dans cette partie. Cela dit, l'intérêt ici est de trouver des objets qui ont une forme ou des caractéristiques qui sont connues. Le processus de segmentation, dont il est question au cours de ce travail, appartient aux approches de type bas niveau.

I.2.1. Principe de la segmentation

L'information véhiculée par une image est portée par bien d'autres données que les seuls niveaux de gris retournés par chaque pixel. Cette grandeur physique est la plus simple, et la plus souvent utilisée, mais les modèles de segmentation que nous utilisons peuvent s'étendre à des critères différents comme des attributs de texture par exemple, le tout étant que les attributs considérés puissent être caractérisés par des valeurs numériques. Une fois ces attributs caractérisés, le principe de la segmentation consiste à réaliser une partition de l'image I en sous-ensembles R_i , appelés régions. Ainsi, les ensembles qui composent une segmentation doivent satisfaire :

$$\begin{aligned} \forall i, R_i &\neq \emptyset \\ \forall i \neq j, R_i \cap R_j &= \emptyset \\ I &= \bigcup_{i=1}^K R_i \end{aligned}$$

Une région est un ensemble de pixels connexes ayant des propriétés communes qui les différencient des pixels des régions voisines [CoP 1995]. Aujourd'hui, vu la simplicité de la définition de l'objectif de la segmentation d'images et la richesse des travaux effectués dans ce domaine, on pourrait penser que la segmentation d'images est un problème en grande partie résolu. Il n'en est rien, probablement à cause de la très grande diversité des applications et des approches mathématiques mises en œuvre sur des ordinateurs toujours plus puissants permettant l'exploration de nouvelles approches autrefois trop coûteuses en temps de calcul.

De nombreuses méthodes ont été développées pour la segmentation d'images [FUM 1981]. Deux familles importantes sont : la segmentation par recherche de frontières et celle par recherches de régions.

Les méthodes par recherche de frontières sont des méthodes de détection de contours correspondant à une variation d'intensité ou à une discontinuité entre deux ensembles connexes de pixels. Ensuite, les contours obtenus doivent être fermés de façon à obtenir une partition de l'ensemble des pixels de l'image.

L'approche pour la segmentation par recherche de régions consiste à regrouper les pixels ayant des propriétés communes et à leur attribuer une même étiquette.

Le processus de segmentation nécessite de déterminer les paramètres de classes caractérisant les différents attributs de l'image. Lorsque les paramètres sont connus au préalable (estimés par un algorithme ou donnés par un opérateur), nous parlerons de segmentation supervisée, au contraire de la segmentation non supervisée où les paramètres sont déterminés au sein même de la segmentation.

Les segmentations supervisées consistent à déterminer des frontières de décision linéaires ou non- linéaires afin de segmenter les données. Les méthodes de segmentation linéaires ne sont généralement pas applicables à des données non linéairement séparables, puisque les frontières de décision linéaires obtenues par ces méthodes ne prennent pas correctement en compte la répartition souvent complexe des données.

Les méthodes non- supervisées sont très intéressantes car elles ne supposent pas d'étapes d'apprentissage ou la mise à disposition d'un ensemble de données préalablement étiquetées. De plus elles ne se heurtent pas au problème de généralisation et/ou de pertinence de l'ensemble d'apprentissage. Cependant leur utilisation est souvent délicate et spécifique à un type de traitement.

I.2.2. Quelques méthodes de segmentation

Dans cette section, nous décrivons brièvement plusieurs approches sur la segmentation d'images médicales. Nous définissons chaque méthode, fournissons une vue d'ensemble sur sa mise en oeuvre et discutons ses avantages et ses inconvénients. Nous divisons des méthodes de segmentation en neuf catégories :

- Les approches de seuillage,
- Les approches par croissance de régions,
- Les approches de coalescence,
- Les approches de support vecteur machine (SVM),
- Les approches de l'analyse factorielle discriminante,
- Les réseaux de neurones artificiels,
- Les modèles déformables,
- Les approches de template,
- Autres méthodes.

I.2.2.1. Le seuillage

Une procédure de seuillage tente de déterminer une valeur d'intensité, appelée seuil, qui sépare les classes désirées. La segmentation est alors réalisée en groupant tous les pixels d'intensité plus grand que le seuil dans une classe et tous les autres pixels dans une autre classe. La détermination de plus qu'une valeur de seuil est un processus appelé multiseuillage. Souvent, le seuillage est un des moyens simple et efficace pour obtenir une segmentation des images. Ses limitations principales sont que dans sa forme la plus simple seulement deux classes sont produites et il ne peut pas être appliqué aux images à canaux multiples. De plus, le seuillage ne tient pas compte des caractéristiques spatiales d'une image. Cela le rend sensible au bruit et des inhomogénéités d'intensité, qui peuvent arriver dans les images mammographies [FUM 1981].

I.2.2.2. La croissance de régions

Dans sa forme la plus simple, la segmentation par croissance de régions exige un point de graine choisi manuellement par un opérateur et extrait tous les pixels connectés à la graine initiale avec la même valeur d'intensité.

Il est employé en particulier pour la délimitation de structures petites, simples tel que les tumeurs et les lésions. Son inconvénient primaire est qu'il exige l'interaction manuelle pour obtenir le point de graine. Ainsi, pour chaque région qui doit être extraite, la graine doit être plantée.

L'accroissement de régions peut être aussi sensible au bruit, cela cause des régions extraites ayant des trous ou qu'elles deviennent même déconnectées[CoM 2002].

I.2.2.3. La coalescence

Les méthodes de coalescence sont nombreuses et très souvent utilisées pour leur facilité et leur rapidité en temps de calcul. La méthode des K-moyennes [HaT 2001] est par exemple, une méthode basée sur un regroupement (clustering) de points en fonction d'une métrique d préalablement définie (distance euclidienne, distance de Bhattacharya, distance $L1$...). L'observation est découpée en plusieurs groupes tels que les éléments d'un même groupe soient les plus proches possibles et ceux de groupes différents soient les plus différents possibles au sens de la métrique d .

La méthode des K-moyennes n'introduit aucune contrainte spatiale entre l'élément courant et ses voisins dans l'image et est basée uniquement sur le choix de la métrique d . Bien que les algorithmes de cluster n'exigent pas de données d'apprentissage, ils exigent une segmentation initiale. Ces algorithmes de cluster n'incorporent pas directement la modélisation spatiale et peuvent donc être sensible aux inhomogénéités d'intensité et le bruit. Cependant, ce manque de modélisation spatiale peut fournir des avantages significatifs pour le calcul rapide [MDP 1994].

L'algorithme LBG (Linde-Buzo-Gray) [LBG 1980] consiste à "découper" successivement l'observation à l'aide de l'algorithme des K-moyennes (avec la métrique d). La principale différence avec la méthode précédente réside dans la construction des regroupements dans l'image. En effet, le nombre de classes croît progressivement dans l'algorithme des LBG (via une suite de découpage des regroupements déjà établis) alors qu'il est fixé pour les K-moyennes. Comme pour les K-moyennes, l'algorithme LBG n'introduit aucune contrainte spatiale entre l'élément courant et ses voisins dans l'image.

I.2.2.4. Support vecteur machine (SVM)

La méthode des Support Vecteur Machine (SVM) [CST 2000] est une approche non-linéaire élégante. En effet, les SVM utilisent un noyau (simple fonction analytique), ou une combinaison de noyaux simples, afin de linéariser les données et obtenir un hyperplan séparant les classes. La méthode SVM est rapide et souple d'utilisation notamment grâce à la construction de noyaux particuliers et spécifiques à une problématique donnée mais reste supervisée.

Par ailleurs, les algorithmes génétiques [Gol 1989] sont une méthode d'optimisation possible se calquant sur la théorie de l'évolution biologique. Ces algorithmes peuvent explorer un très grand espace de solutions et trouver une solution minimisant une fonction d'erreur (fonction d'adaptation) par une série de croisements, de mutations des individus constituant la population. Les individus minimisant la fonction d'adaptation sont conservés tandis que ceux ne s'adaptant pas au problème sont éliminés lors d'une étape de sélection. Le choix de la fonction d'adaptation permet une genericité totale dans l'utilisation de ces algorithmes évolutionnaires mais les algorithmes génétiques ont l'inconvénient d'être lents et leur utilisation nécessite une calibration correcte de leurs nombreux paramètres.

I.2.2.5. L'analyse factorielle discriminante

L'analyse factorielle discriminante [Vol 1985], par exemple, sépare linéairement les données en les projetant dans un espace minimisant la variance intra-classes tout en maximisant la variance inter-classes. Cette méthode est particulièrement rapide et les frontières de décisions obtenues discriminent linéairement les nuages de points. L'inconvénient majeur de cette méthode est la nécessité de disposer d'un ensemble d'apprentissage complet et pertinent afin de déterminer les frontières de décision entre les classes. La généralisation de la classification obtenue à des données non classifiées est, en outre, généralement difficile.

I.2.2.6. Les réseaux de neurones artificiels

Les réseaux de neurones artificiels ont été utilisés avec succès dans plusieurs applications de reconnaissance de formes. Ils sont capables d'agir comme des classificateurs. La puissance des réseaux de neurones artificiels tient non pas dans la capacité de calcul du processeur, mais plutôt dans le nombre de leurs inter-connexions et dans le parallélisme de l'algorithme.

Les réseaux de neurones artificiels sont des réseaux de traitement parallèle où leurs nœuds simulent un apprentissage biologique. Chaque nœud dans les réseaux de neurones artificiels est capable d'exécuter des calculs élémentaires. Il est réalisé par l'adaptation de poids assignés aux connexions entre nœuds. Les réseaux de neurones artificiels représentent un paradigme pour une machine d'apprentissage et peut être employé dans une variété de voies pour segmentation d'images. L'utilisation la plus largement appliquée dans l'imagerie médicale est comme étant un classificateur, où les poids sont déterminés en employant des données d'apprentissage et le réseau neuronal est alors employé pour segmenter de nouvelles données. Le réseau neuronal peut aussi être employé d'une façon non supervisée comme une méthode de cluster, aussi bien que pour des modèles déformables.

À cause de beaucoup d'interconnexions employées dans un réseau neuronal, l'information spatiale peut facilement être incorporé dans ses procédures de segmentation [CMM 2002].

I.2.2.7. Les modèles déformables

Les modèles déformables délimitent les frontières de région en employant des courbes paramétriques fermées ou des surfaces qui se déforment sous l'influence de forces internes et externes.

Pour délimiter une frontière d'objet dans une image, une courbe fermée ou une surface doivent d'abord être placés près de la frontière désirée et ensuite permet de subir un processus de relaxation itératif. Des forces internes sont calculées à l'intérieur de la courbe ou de la surface pour le tenir lisse partout dans la déformation. Les forces externes sont d'habitude tirées de l'image pour guider la courbe ou la surface vers des caractéristiques d'intérêt désirées. Les principaux avantages des modèles déformables sont leur capacité de produire directement des courbes paramétriques fermées ou surfaces dans des images. Cela fournit une robustesse au bruit et aux faux contours. Un inconvénient est qu'ils exigent une interaction manuelle pour placer un modèle initial et choisir des paramètres appropriés [CMM 2002].

I.2.2.8. Le template

Dans cette approche, on utilise une forme géométrique aussi appelée patron (template) que l'on va tenter de localiser sur l'image. Le principe est de connaître exactement (ou presque) la forme de l'objet que l'on recherche et de parcourir toute l'image pour placer le template à l'endroit le plus vraisemblable. Ceci est beaucoup plus contraignant que d'avoir besoin de savoir combien d'éléments sont présents dans l'image. Il faut, dans ce cas, connaître la forme précise de ce que l'on recherche. Il existe des travaux sur des templates déformables qui se rapprochent des Snakes. On localise le meilleur endroit pour placer le template dans l'image puis on le laisse se déformer pour correspondre au mieux à l'image. On soumet le template pour cela à une énergie du même genre que l'énergie externe d'un Snake (énergie imposée par l'image). Un avantage de ces approches est que les étiquettes sont transférées aussi bien que la segmentation. Ils fournissent aussi un système standard pour l'étude morphométrique des propriétés [CMM 2002].

I.2.2.9. Autres méthodes

L'intelligence artificielle propose une série de méthodes dérivées des approches neuronales afin de classifier les données de manière non-supervisée. C'est le cas par exemple des algorithmes ART (Adaptive Resonance theory) [GrC 1991] et des réseaux de Kohonen [Koh 1982] permettant d'auto organiser les connaissances en structure tendant à résoudre le problème de la stabilité-plasticité. La stabilité correspond à la capacité du système à organiser les données tandis que la plasticité correspond à la capacité du système à appréhender de nouvelles données. Enfin la morphologie mathématique propose également des méthodes de segmentation basées sur des opérateurs dérivés des fermetures et ouvertures morphologiques [Soi 2004].

Ces approches ne prennent généralement pas en compte le bruit de l'image, pouvant être porteur d'information, qui est filtré dans une étape préalable à la segmentation.

I.3. Les images à traiter

Depuis plusieurs années, les connaissances médicales, le diagnostic et les thérapies se transforment de manière durable sous l'influence des progrès technologiques dans le domaine de l'imagerie. Les développements de l'imagerie permettent, entre autres, l'aide au diagnostic, le suivi de l'évolution des pathologies, la détection précoce, la planification et la simulation d'interventions.

I.3.1. Principe de la mammographie

La mammographie a pour but de déceler au plus tôt des anomalies avant même qu'elles n'aient provoqué des symptômes cliniques. Elle peut permettre, ainsi, de détecter des cancers bien avant qu'ils ne soient palpables. Pour réaliser cet examen, le manipulateur utilise un mammographe. Cet appareil se compose d'un tube à rayons X et d'un système de compression du sein.

L'examen consiste à comprimer à tour de rôle les deux seins puis à les exposer à une faible dose de rayons-X. La compression est obtenue par une pelote de compression spécialement conçue pour ce type d'exploration et réglée par un manipulateur spécialisé pour qu'elle soit indolore ou parfaitement tolérée par la patiente. Cette compression permet l'étalement des tissus mammaires ce qui facilite la visualisation des structures du sein et

permet de réduire la dose de rayons-X délivrée. La mammographie est le principal outil de dépistage et de diagnostic ; c'est un examen radiographique (par rayons X) des deux seins, généralement sous deux incidences pour (CC Cranio- Caudale et MLO pour Medio- Laterale Oblique). On considère qu'en moyenne le cancer est visible sur la mammographie de un à trois ans avant le premier signe clinique.

La mammographie reste l'outil principal de dépistage pour plusieurs raisons [Gui 1996] :

- c'est un examen sensible (85 à 97 %) et spécifique (70 à 80 %),
- le coût moyen d'un examen mammographique est faible comparé à d'autres techniques d'imagerie comme l'IRM. De plus l'appareillage est mobile et peu encombrant,
- l'examen n'est pas traumatisant. En effet les tissus du sein sont mous et perméables aux rayons X, rendant faibles les doses d'irradiation. D'autre part, la compression du sein, nécessaire pour obtenir une meilleure uniformité de la densité optique, ne doit pas en général être douloureuse,
- la résolution spatiale d'une radiographie est de l'ordre de quelques dizaines de microns, ce qui permet de visualiser les micro-calcifications, contrairement aux autres techniques d'imagerie ;
- les mammographies numériques sont de plus en plus utilisées, ce qui permet de disposer du signal directement sous forme numérique, avec les avantages associés en terme d'archivage et de consultation de dossier médical à distance.

Pour autant, la mammographie numérique reste coûteuse pour des performances au niveau des diagnostics équivalents à celles des techniques conventionnelles.

Il y a beaucoup de maladies qui affectent le sein, les plus connus sont : le fibroadénomes et le kystiques, qui forment 90% des lésions du sein [Haa 1971]. Les kystes, sont des tumeurs bénignes contenant un corps liquide ou mou, dans la plupart des cas, il sont résultat d'un blocage dans les conduits de lait dus à une inflammation, il sont bien tracés et légèrement mobile. Le fibroadénomes, qui apparaît généralement chez les jeune femmes, sont également des tumeurs bénignes et se composent de glande comme tissu, kystes et fibroadénomes, se développent lentement à une tache, serrant les pièces voisines de côté mais ne les envahissant pas.

Les tumeurs malignes (les cancers), s'écartent rapidement, d'un point à un autre, envahissant et détruisent les tissus environnants. Les cancers sont mal tracés, immobiles et pleins. Le taux de croissance des tumeurs malignes a été rapporté, avec doublement du taux de 30 jours à 300 jours. La détection tôt et le traitement du cancer sont nécessaires pour réduire le taux de mortalité [Roe 1986] et il est important que les techniques d'imagerie soient en mesure de détecter le cancer en premier lieu et distinguer les lésions bénignes et malignes. La capacité de détecter de petits cancers tôt est importante si une technique d'imagerie doit être employée comme méthode pour scanner. Actuellement la mammographie, examine le sein par radiographie faible puissance, c'est donc la technique principale de l'imagerie. Les études ont prouvé à travers l'examen clinique que cette méthode est meilleure pour la détection des cancers moins de 2 centimètres de diamètre [Roe 1986]. D'autres techniques plus récentes tels que la micro-onde, l'imagerie par thermographie, radiographie ultrasonique et l'infrarouge, n'ont pas été évalué entièrement comme la mammographie, mais constituent des méthodes additionnelles utiles d'imagerie à employer

- chez les femmes avec le tissu glandulaire dense par lequel la détection des lésions par mammographie est difficile,
- fournissent des informations additionnelle qui contribuent dans le diagnostique.
- en tant que méthodes de scannage préliminaires pour identifier les femmes de gros risque qui peuvent alors être exposées à la mammographie. Ceci réduirait le nombre de femmes exposées aux rayons X et le risque lié à ceci.

Un excellent examen complet des méthodes d'imagerie de sein a été donné par [Jon 1982].

I.3.2. Les caractéristiques des images mammographies

Quand l'imagerie mammographie a été présentée, la communauté scientifique n'y a accordé que peu d'intérêt, sans pressentir les améliorations qu'elle allait apporter dans l'aide au diagnostic ou au confort du malade. Actuellement, elle a pris une importance considérable dans l'ensemble des modalités d'imagerie diagnostique. L'analyse de la mammographie nécessite des clichés de bonne qualité et une approche systématique de l'analyse, en deux étapes : Lors de la première étape, le but est de détecter une région du sein qui a un aspect différent par rapport aux régions voisines ou au sein controlatéral. La seconde étape consiste à disséquer cet aspect pour distinguer les variations plus ou moins physiologiques ou au contraire classer l'image comme suspecte.

La majorité de recherche dans les méthodes de segmentation d'image médicale se concentre sur la mammographie, particulièrement la détection de cancer. La plupart de ces méthodes sont basées sur l'information de l'intensité et leur souci principal est la présence d'inhomogénéités d'intensité des images mammographies. Ainsi les images mammographies sont très sensibles aux distances résiduelles résultant du mode d'acquisition. Pour corriger ces imperfections et réduire ces hétérogénéités, nous avons exploité une méthode de segmentation automatique basée sur les champs de Markov cachés. Cette méthode permet d'extraire automatiquement les régions de l'image et permet aussi de corriger les hétérogénéités des images grâce à une optimisation de l'entropie (énergie).

Donc ce travail s'articule autour de deux principales thématiques d'application, l'extraction des régions et la réduction des inhomogénéités des images mammographies suivantes :

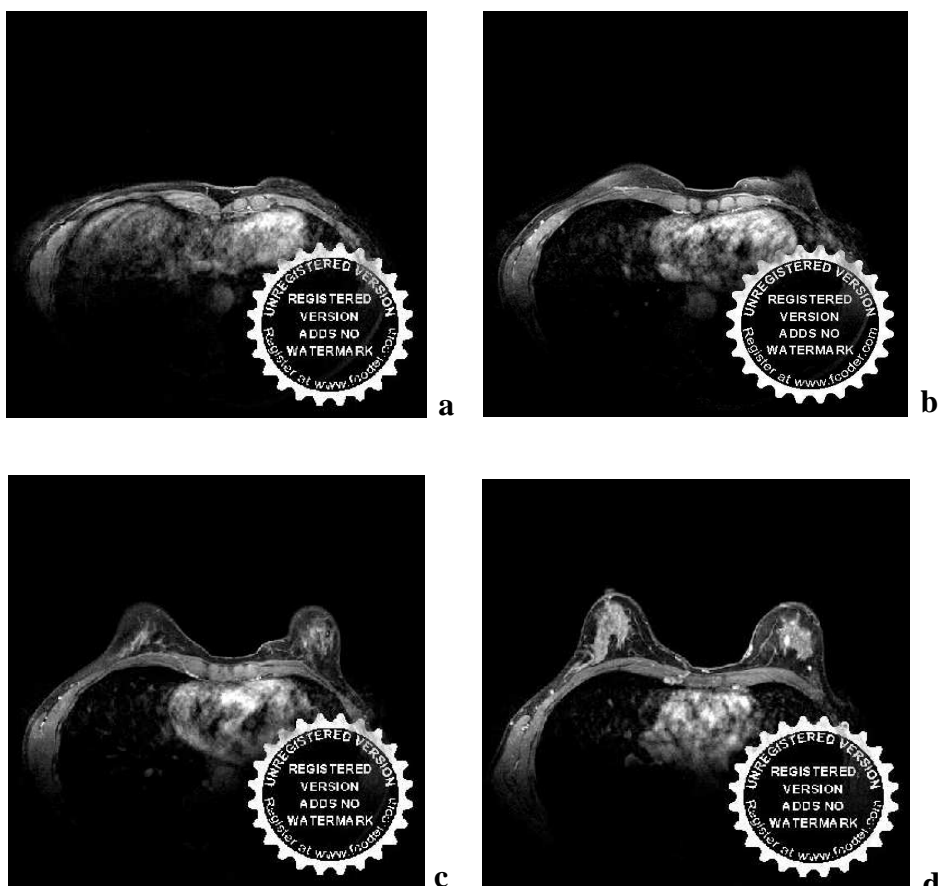


FIG.I.1 - Les images de simulation. a) cas sain et b, c et d) présence d'anomalie.

I.4. Généralités sur les champs de Markov cachés

Les champs de Markov cachés ont connu un essor considérable, en traitement d'images, à partir des années quatre-vingts. Leur intérêt est de pouvoir prendre en compte l'information contextuelle dans une image, de manière mathématiquement rigoureuse et pour les contextes de très grande taille. En segmentation d'images, problématique traitée dans cette thèse, on considère deux champs aléatoires $X = (X_s)_{s \in S}$ et $Y = (Y_s)_{s \in S}$, où S est l'ensemble des pixels. Chaque X_s prend ses valeurs dans un ensemble fini de $\Omega = \{\omega_1, \dots, \omega_k\}$ et chaque Y_s prend ses valeurs dans l'ensemble des nombres réels \mathfrak{R} . Le problème de la segmentation est celui de l'estimation de $X = (X_s)_{s \in S}$ à partir de $Y = (Y_s)_{s \in S}$. Ce problème ne peut alors être traité directement à cause du nombre trop grand de réalisations possibles de $X = (X_s)_{s \in S}$.

En effet, si N est la taille de S , qui est couramment de l'ordre de 256×256 (mais peut également atteindre 1024×1024), le nombre de réalisation de $X = (X_s)_{s \in S}$ est de k^N , ce qui ne peut pas être traité directement pour toute loi du couple (X, Y) . On est ainsi obligé de considérer des lois (X, Y) particulières et les champs de Markov cachés sont largement utilisés pour cette modélisation. Après les articles fondateurs [Bes 1974, Bes 1986, DeE 1987, GeG 1984, MMP 1987], des centaines de publications ont été présentées sur diverses applications des champs de Markov cachés, en imagerie et il est impossible de les citer toutes. Mentionnons les livres suivants [Cha 2000, Guy 1995, Li 1995, Mai 2003, Win 2003], incluant chacun une riche bibliographie, ou encore les articles de synthèse [Per 1998, Pie 2003]. Citons également quelques exemples d'articles traitant des généralités ou des applications particulières en imagerie médicale [MMS 2002], en vision [KeH 1998, HPB 1994], en télédétection [Ric 2005, PSZ 2005], en imagerie radar [DMP 1997, TNT 2004, TuR 2005], ou encore en imagerie sonar [MCP 2000, RPB 2003]. De plus, diverses publications montrent que les résultats obtenus avec ce type de modèles peuvent être, dans certaines situations, spectaculaires.

Le succès des champs de Markov cachés est dû au fait qu'il est possible, dans leur cadre, de mettre en place des estimateurs bayésiens de X , à partir de Y , grâce à des techniques générales de simulation dites "méthodes de Monté Carlo par chaînes de Markov" (MCMC de l'anglais "Monte Carlo Markov Chains" [Rob 1991], parmi d'autres). Les techniques MCMC

sont applicables à condition que la loi de X conditionnelle à $Y = y$ (sa loi "a posteriori") soit de Markov. Les champs de Markov cachés sont souvent incorporés dans des algorithmes de segmentation dans un cadre bayésien. La segmentation est alors obtenue en maximisant la probabilité a posteriori sachant les données de l'image en employant des méthodes itératives comme l'ICM (pour «Mode Conditionnel Itéré»), le MAP (pour «Maximum A Posteriori») et le MPM (pour «Mode des Marginales a Posteriori»).

Nous avons vu dans le paragraphe précédent qu'il était possible de fabriquer des estimateurs optimaux pour résoudre un problème de segmentation, pourvu que nous ayons à notre disposition la distribution a priori et la distribution des observations conditionnellement (à posteriori). Malheureusement, nous avons très rarement une connaissance aussi complète de la structure probabiliste du problème ; tout au plus pouvons nous avoir quelques vagues connaissances générales, comme des formes paramétriques pour la priori et à posteriori par exemple qui diminueront le nombre de paramètres à estimer, et/ou un ensemble d'échantillons représentatifs des objets de chaque classe. Un nouveau problème consiste alors à utiliser au mieux ces informations pour construire un estimateur.

L'apprentissage des paramètres d'une distribution est un problème classique des statistiques et est généralement résolu grâce à l'estimation du Maximum de Vraisemblance des données. Mais il est nécessaire de bien distinguer deux cas : celui avec données complètes et celui avec données incomplètes. Le premier problème correspond à l'estimation des paramètres d'un modèle à l'aide d'une (ou plusieurs) réalisation(s) de celui-ci. Le second met en jeu l'estimation des paramètres de la distribution jointe à posteriori, en ayant à notre disposition uniquement une partie des données qui seront ici ce que nous avons appelées précédemment les observations y .

Dans le travail qui suit, il s'agit de segmentation non supervisée en s'appuyant sur une modélisation probabiliste. On considère alors que l'image observée et l'image segmentée sont des réalisations de variables aléatoires liées et les données à restaurer étant les étiquettes assignées à chaque pixel dans l'image segmentée. La richesse des modèles probabilistes permet de saisir les caractéristiques globales et locales des images. Parmi ceux-ci, nous nous sommes intéressés à des champs de Markov cachés, Ce modèle est intéressant car il utilise réellement l'information spatiale dont on dispose a priori sur l'image que l'on observe.

I.5. Conclusion

Les futures recherches dans la segmentation d'images médicales lutteront vers l'amélioration de la performance, de la précision et de la vitesse de calcul des méthodes de segmentation, aussi bien que la réduction de la quantité d'interaction manuelle. Probablement la question la plus importante entourant l'utilisation de la segmentation d'images est son application dans les cliniques. Il est peu probable que l'automatisation des méthodes de segmentation ne remplacera jamais les médecins mais, ils deviendront probablement des éléments cruciaux dans l'analyse d'image médicale.

Actuellement, la mammographie est l'examen de choix. Les tumeurs et les lésions sont démontrées par la mammographie mieux que par aucune autre modalité. L'objectif initial de ce travail est de faciliter l'exploitation des images mammographie, afin d'apporter une aide au diagnostic et au suivi médical.

L'étude des différentes techniques de segmentation d'images montre que la théorie la plus utilisée pour modéliser l'information extraite des images est sans conteste la théorie des probabilités basée sur les champs de Markov cachés. Ils introduisent des informations contextuelles avec une régularisation spatiale.

Tandis que les autres méthodes de segmentation d'images, qu'on a vu, comportent certaines limites. Ces méthodes ne respectent pas la cohérence spatiale des classes dans l'image.

Cependant, en segmentation d'images les méthodes statistiques de segmentation utilisant les champs de Markov cachés donnent des résultats satisfaisants dans de nombreuses situations. Dans ce contexte, nous montrons qu'il est possible d'utiliser, dans la segmentation des images mammographies, la puissance de cette modélisation markovienne cachée au niveau de la prise de décision.

L'objet du chapitre suivant sera donc de présenter la modélisation markovienne cachée.

CHAPITRE 2

MODELISATION MARKOVIENNE

Dans [Li 1995] et [DuJ 1989], nous apprenons que le but de la modélisation mathématique en traitement d'images est d'extraire certaines caractéristiques intrinsèques de l'image par quelques paramètres pour mieux comprendre la nature du phénomène qui a construit l'image.

Le but de ce chapitre est d'introduire la théorie des champs markoviens d'une manière générale et de présenter les algorithmes classiques de simulation. A cet effet, nous commençons par rappeler quelques concepts généraux concernant les distributions de champs de Markov, comme le système de voisinage et équivalence champ de Markov – champ de Gibbs. On décrit par la suite les différentes méthodes classiques de simulation, tel que l'échantillonneur de Gibbs ou l'algorithme de Metropolis. Enfin, nous présentons ensuite la modélisation par les champs de Markov cachés, ainsi que les différentes techniques d'estimation des paramètres du modèle. Diverses simulations ont été effectuées dont certain feront l'objet d'une comparaison.

II.1. Introduction

La modélisation de l'image a des buts variés parmi lesquels se trouve la description de l'image avec peu de paramètres, ce qui permet de comprendre la nature du phénomène d'origine et ses restrictions naturelles. Les champs de Markov introduisent la prise en compte de la notion de dépendance spatiale dans l'image et permettent de pondérer la portée de l'influence de ces dépendances, offrant ainsi un cadre mathématique cohérent pour divers traitements. En particulier, les techniques d'acquisition de données (optique, radar, sonar, scanner,...) créent des superpositions entre les mesures faites sur des pixels voisins, engendrant ainsi une dépendance spatiale qui rendent d'autant plus attrayante et nécessaire l'utilisation des champs markoviens. Ainsi, l'intérêt de l'utilisation des champs en segmentation d'images est de mieux modéliser l'image. Ceci permet d'obtenir de bons résultats de segmentation d'images qui, associée aux champs de Markov et à la théorie bayésienne.

II.2. Champ de Markov et distribution de Gibbs

La théorie des champs de Markov est une branche de la théorie des probabilités pour l'analyse des dépendances spatiales de phénomènes physiques.

Dans cette section nous commençons par rappeler quelques notions de base, tels le système de voisinage et clique. Ensuite, nous présentons la définition mathématique des champs de Markov et distribution de Gibbs, ainsi que l'équivalence entre champs de Markov et champs de Gibbs (théorème de Hammersley – Clifford).

II.2.1. Système de voisinage et cliques

II.2.1.1. Système de voisinage

En analyse d'images, l'image est représentée par une grille $S = \{s = (i, j); 0 \leq i \leq n-1; 0 \leq j \leq m-1\}$ rectangulaire finie bidimensionnelle de taille $n \times m$, contenant un ensemble de "sites" $s \in S$ appelés également "pixels".

On appelle un système de voisinages toute famille $V = (V_s)_{s \in S}$ vérifiant les propriétés suivantes :

- $s \notin V_s$, i.e. un site n'est pas voisin à lui même ;
- $s \in V_t \Leftrightarrow t \in V_s$, qui signifie que la relation de voisinage est symétrique

A titre d'exemple, l'ensemble des "voisins" d'un site s peut être défini comme un ensemble de sites situé dans un rayon r : $V_s = \{t \in S / d(s,t) \leq r, t \neq s\}$, où $d(s,t)$ est la distance euclidienne entre le pixel s et le pixel t et r est à valeur dans l'ensemble des entiers [Li 1995].

Notons que les systèmes de voisinage les plus usuels, représentés en **FIG.II.1**, sont :

- Système "4 - connexité", également appelé voisinage du premier ordre ;
- Système "8 - connexité", également appelé voisinage du second ordre.

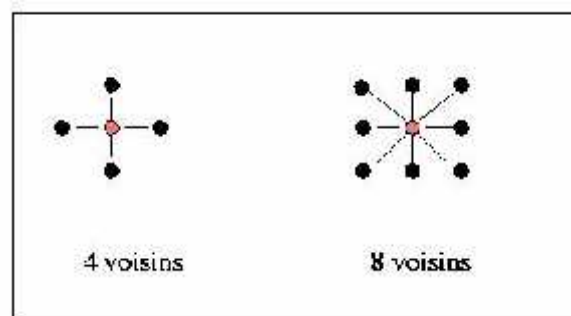


FIG.II.1 - Relation de voisinage aux 4 ou 8 plus proches voisins.

A partir d'un système de voisinage, un système de clique est déduit.

II.2.1.2. Clique

Un sous ensemble c de S est appelé clique relative au système de voisinage V , si c est un singleton ou si tous les sites distincts de c sont voisins. Une clique est dite d'ordre p , si elle contient p éléments, en d'autres termes, l'ordre d'une clique est par définition son cardinal. **FIG.II.2** illustre la notion de cliques dans le cas de voisinage en 4-connexité et en 8-connexité. On notera C l'ensemble des cliques associées à un système de voisinage.

Nous allons définir l'énergie de l'image à partir des cliques, car un potentiel est associé à chaque clique dont la valeur dépend de l'état des sites qui la constituent. L'énergie globale U de l'image est donnée par :

$$U = \sum_{c \in C} U_c \dots\dots\dots (II.1)$$

et l'énergie locale en un site s dans l'image est :

$$U_s = \sum_{c \in C / s \in c} U_c \dots\dots\dots (II.2)$$

interprétée comme la somme des potentiels de toutes les cliques dont s fait parti.

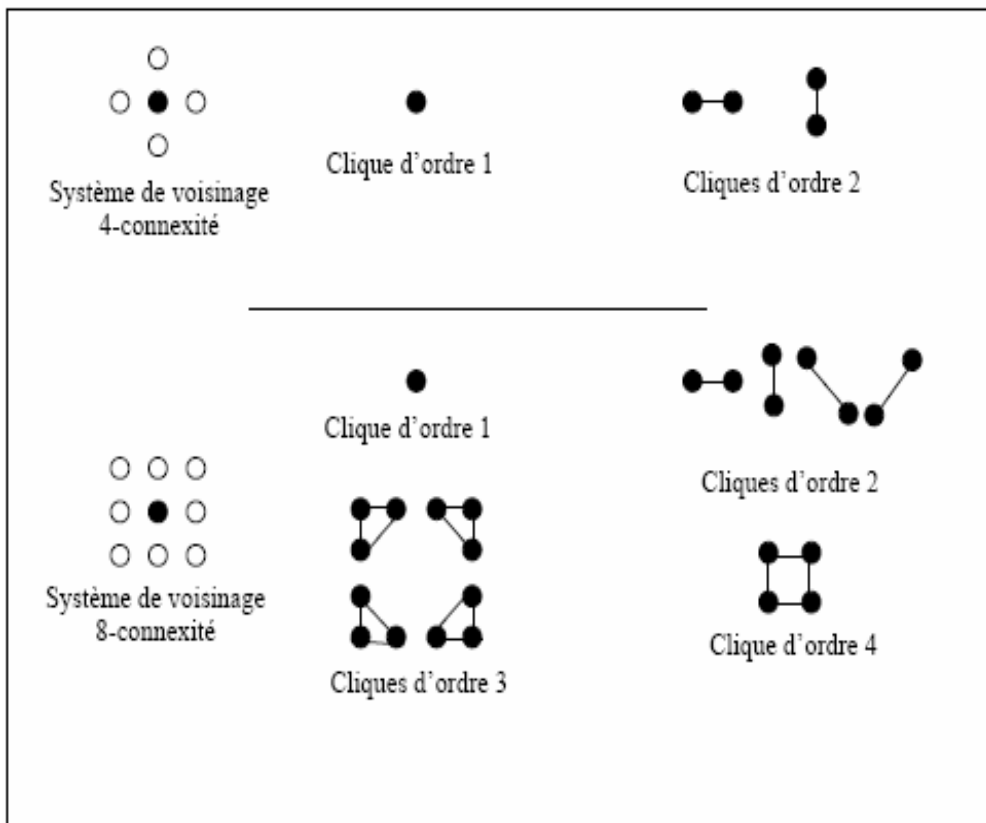


FIG.II.2 - Cliques associées à un système de voisinage en 4-connexité et en 8-connexité

II.2.2. Champs de Markov aléatoires

Soit S un ensemble de pixels et $X = (X_s)_{s \in S}$ une famille de variables aléatoires, définie sur S , où chaque X_s est à valeurs dans un ensemble fini de classes $\Omega = \{\omega_1, \dots, \omega_k\}$. Une telle famille sera dans la suite appelée "champ aléatoire". Soit un système de voisinage V . On notera par la même lettre p les diverses lois de probabilité liées à X .

On dit qu'un champ X est un champ de Markov relativement à V , si et seulement si les deux propriétés suivantes sont vérifiées [Li 1995] :

- Positivité :

$$\forall x \in \Omega^{Card(s)}, p(x) > 0 \text{ (Contrainte de positivité)} \dots \dots \dots \text{(II.3)}$$

- Markovianité :

$$p(x_s / x_t, t \neq s) = p(x_s / x_t, t \in V_s) \dots \dots \dots \text{(II.4)}$$

II.2.3. Champs de Gibbs

Soit S un ensemble de pixels muni d'un système de voisinage V , et soit C l'ensemble des cliques c associées à V . Le champ X est un champ de Gibbs (CG), si et seulement si sa loi est définie par :

$$p(x) = \frac{1}{Z} \exp(-U(x)) \dots \dots \dots \text{(II.5)}$$

avec :

$$U(x) = \sum_{c \in C} \varphi_c(x_c), \dots \dots \dots \text{(II.6)}$$

qui est dite "fonction énergie", et :

$$Z = \sum_{x \in \Omega^{Card(s)}} \exp(-U(x)) \dots \dots \dots \text{(II.7)}$$

Une telle loi est dite une "distribution de Gibbs". Notons que φ_c sont des applications de $\Omega^{Card(c)}$ dans \mathfrak{R} appelées "fonction potentiel" avec x_c la restriction de x à c et Z est une constante de normalisation, également appelée "fonction de partition" qui fait de $p(x)$ une probabilité. Etant donné que cette constante est une somme sur l'ensemble de toutes les configurations possibles du champ X , son calcul est dans la pratique généralement impossible. Par exemple, pour une image binaire ($Card(\Omega) = 2$) de taille 256×256 pixels, le calcul de Z nécessite une somme sur $2^{256 \times 256}$ configurations possibles, ce qui est rédhibitoire.

II.2.4. Equivalence champ de Markov – champ de Gibbs

Comme nous l'avons évoqué précédemment, le champ de Markov (CM) est caractérisé par sa propriété locale (Markovianité) tandis que le champ de Gibbs (CG) est caractérisé par sa propriété globale (distribution de Gibbs). Le théorème de Hammersley-Clifford (1971), dont une démonstration peut être vue dans [Bes 1974], établit l'équivalence entre ces deux propriétés.

Théorème :

Soit S un ensemble de pixels muni d'un système de voisinage V . Un champ X sur S est un champ de Markov relativement à V , si et seulement si X est un champ de Gibbs de potentiel associé à V .

L'intérêt pratique de ce théorème est qu'il permet d'accéder d'une manière simple à une forme exploitable des probabilités jointes et ces, en spécifiant les fonctions potentielles φ_c définies sur les restrictions x_c de x aux clique c . Afin de calculer la probabilité jointe du champ de Markov qui est une distribution de Gibbs, il est nécessaire d'évaluer la fonction de partition (II.5). Comme la somme englobe l'ensemble des configurations, ce calcul est en général impossible. Cependant, il reste possible de simuler des réalisations de ce champ à partir de ses caractéristiques locales, grâce à des méthodes de relaxation que nous allons voir dans la section suivante. C'est cette possibilité de simulations, qui s'avère être un puissant outil, qui est à l'origine des champs de Markov en imagerie.

II.2.5. Modèles markoviens classiques

Des nombreux modèles classiques sont des champs markoviens définis par leurs fonctions de potentiels et leur voisinage. Nous allons présenter deux exemples : le modèle d’Ising et le modèle de Potts, qui est la généralisation du premier pour un espace d’états plus large.

II.2.5.1. Modèle d’Ising

Ce modèle, qui a été développé à partir de l’étude du ferromagnétisme en physique statistique [Isi 1925], utilise un espace binaire de descripteurs, correspondant à celui des spins $\Lambda = \{1, -1\}$ comme le montre **FIG.II.3**.

La fonction d’énergie pour ce modèle en tenant compte d’un voisinage 4 connexe est :

$$U(x) = \sum_{c \in C_1} U_c(x_s) + \sum_{c \in C_2} U_c(x_s, x_t) \dots\dots\dots (II.8)$$

La clique C_1 est constituée d’un seul spin et le potentiel associé est de la forme $U_c(x_s) = -\alpha \cdot x_s$, où α est un champ magnétique externe. Le potentiel associé aux cliques C_2 est donné par :

$$U_c(x_s, x_t) = \begin{cases} \beta & \text{si } x_s \neq x_t \\ 0 & \text{si } x_s = x_t \end{cases} \dots\dots\dots (II.9)$$

où β est la constante de couplage entre s et t . l’énergie totale s’écrit alors,

$$U(x) = -\sum_{c \in C_1} \alpha \cdot x_s - \sum_{c \in C_2} \beta \cdot x_s x_t \dots\dots\dots (II.10)$$

β régularise le modèle : lorsque $\beta > 0$, les configurations d’énergie minimale sont celles pour lesquelles les sites ont le même état, et si $\beta < 0$ les configurations privilégiées sont celles où les sites sont d’états opposés.

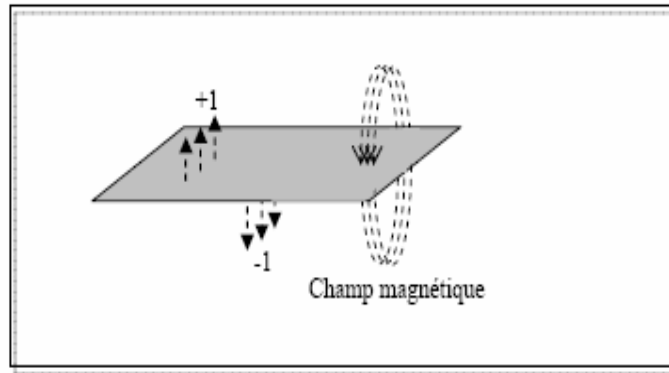


FIG.II.3- Orientation d'un spin à l'état d'équilibre

II.2.5.2. Modèle de Potts

Le modèle de Potts, [Wu 1982], est une extension du modèle d'Ising, pour un espace de n valeurs $\{1, \dots, n\}$, et la topologie est indifféremment la 4- ou la 8-connexité dans une image 2D. Seuls les potentiels d'ordre 2 sont définis par :

$$U_c(x_s, x_t) = -\beta x_s x_t = \begin{cases} \beta & \text{si } x_s \neq x_t \\ 0 & \text{si } x_s = x_t \end{cases} \dots\dots\dots (II.11)$$

Il est très utilisé pour la classification d'images. Dans [SiR 1993], les auteurs utilisent le modèle de Potts pour une classification en tenant compte des différentes valeurs de β .

Lorsque $\beta > 0$, les configurations les plus probables correspondent à des sites voisins ayant des descripteurs égaux, ce qui dans le cas du traitement d'image donne des réalisations avec de grandes zones homogènes (au sens de la similarité entre descripteurs). La taille de ces régions est gouvernée par la valeur de la constante de couplage.

Il est possible de raffiner ce modèle en considérant des valeurs de β différentes suivant les directions explorées dans la clique (vertical/horizontal en 4-connexité par exemple), et donc de privilégier la détection de zones homogènes directionnelles.

II.3. Simulation d'un champ de Markov

La simulation des réalisations des champs de Markov est un outil qui peut s'avérer très utile pour résoudre un problème dont on ne peut pas proposer une solution analytique. En effet, dans le cas de la segmentation d'images, où les paramètres du modèle sont à estimer, certaines méthodes d'estimation des paramètres demandent la réalisation du champ X . Les algorithmes d'échantillonnage, tels que l'algorithme de Métropolis [MRR 1953], ou l'échantillonneur de Gibbs [GeG 1984], que nous décrivons ci après, permettent effectivement la simulation des champs des Markov.

II.3.1. Algorithme de Métropolis

L'algorithme de Métropolis est un algorithme itératif qui permet de simuler des réalisations d'un champ markovien X . On considère une suite x^n de réalisations de X , où chaque x^n est obtenu à partir de x^{n-1} en balayant l'ensemble des pixels et en considérant, en chaque pixel s , la variation de l'énergie locale dU (on entend par énergie locale, la restriction de l'énergie au site s et ses voisins, donnée par (II.12)). Si dU est négative, la classe associée au site s est échangée par celle obtenue par tirage aléatoire. Dans le cas contraire, deux possibilités peuvent se présenter.

Les étapes de cet algorithme sont décrites comme suit :

- Initialiser une première carte "image" x^0 d'une façon arbitraire.
- A chaque itération n :
- Balayer l'ensemble des sites $s \in S$, et en chaque site s :
 - Effectuer un tirage aléatoire d'une variable k selon la loi uniforme dans Ω ,
 - Calculer la variation d'énergie locale dU par :

$$\Delta U = U(x_s = k, x_{V_s}) - U(x_s^n, x_{V_s}) \dots \dots \dots (II.12)$$

- Si $dU \leq 0$: le changement est accepté $x_s^{n+1} = k$
- Sinon, le changement est accepté ($x_s^{n+1} = k$) ou refusé ($x_s^{n+1} = x_s^n$) selon les probabilités $\exp(-\Delta U(x_s^n, x_{V_s}))$ ou $1 - \exp(-\Delta U(x_s^n, x_{V_s}))$

II.3.2. Echantillonneur de Gibbs

Le principe de l'algorithme de l'échantillonneur de Gibbs ressemble à celui de l'algorithme de Métropolis; cependant, la différence réside dans l'étape de mise à jour qui est fondée sur la probabilité conditionnelle au lieu de la variation d'énergie.

Cet algorithme se résume à :

- Initialiser une première carte "image" x^0 d'une façon arbitraire.
- A chaque itération n :
- Balayer l'ensemble des sites $s \in S$, et en chaque site s :
 - Calculer les probabilités $p(x_s^n / x_{V_s})$;
 - Effectuer un tirage aléatoire d'une variable k dans Ω , selon des probabilités conditionnelles et poser $x_s^{n+1} = k$.

On obtient ainsi une suite x^0, x^1, \dots, x^N de réalisations aléatoires du champ X , dont la loi converge vers (II.5) [Win 2003].

Remarque

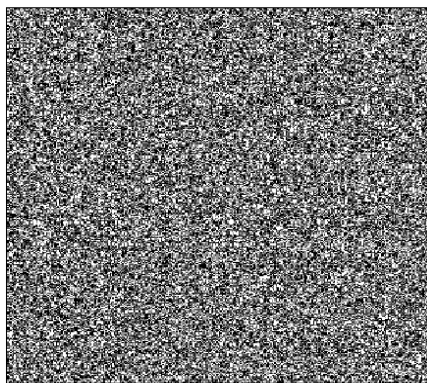
Il existe deux types d'échantillonneurs de Gibbs :

- (i) à balayage séquentiel : dans ce cas, le balayage s'effectue séquentiellement, par exemple dans l'ordre $1-2-\dots-(n \times m)$.
- (ii) à balayage aléatoire : le parcours se fait aléatoirement, à condition de balayer tous les sites de l'image en temps fini avec la probabilité égale à 1.

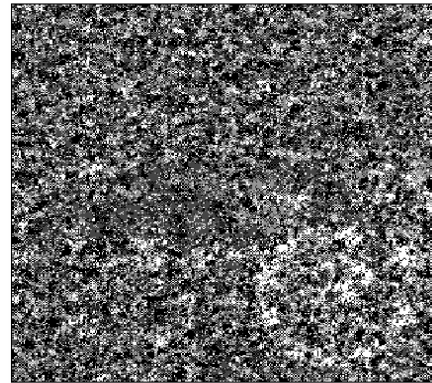
Nous présentons sur les figures **FIG.II.4** d'une réalisation de champs de Markov de loi $p(x)$ définie par :

$$p(x) = \gamma \exp\left(-\beta_H \sum_{s \sim t} (1 - \delta(x_s, x_t)) + \beta_V \sum_{s \sim t} (1 - \delta(x_s, x_t))\right) \dots \dots \dots \text{(II.13)}$$

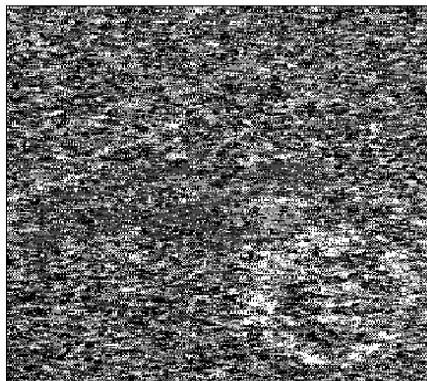
avec les paramètres d'interaction β_H correspondant aux cliques horizontales, β_V aux cliques verticales, les paramètres permettant de générer les disproportions de chaque classe dans l'image, respectivement. Ces simulations correspondent à différentes valeurs de β selon la probabilité a priori. Elles ont été réalisées par l'échantillonneur de Gibbs utilisant 20 balayages déterministes.



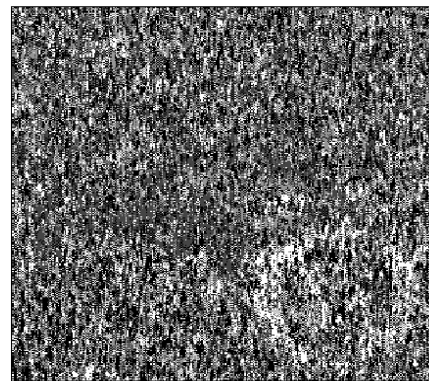
$$\beta_H = \beta_V = 0.5$$



$$\beta_H = \beta_V = 1.0$$



$$\beta_H = 2, \beta_V = 0.3$$



$$\beta_H = 0.3, \beta_V = 2$$

FIG.II.4. - Exemple de simulation d'un champ de Markov à 4 classes par échantillonneur de Gibbs pour différentes valeurs de β selon la probabilité a priori (4 pixels voisins).

Nous montrons sur les figures **FIG.II.5** d'un accomplissement de champs de Markov de loi $p(x)$ définie par :

$$p(x) = \gamma \exp\left(-\beta \sum_{s \sim t} (1 - \delta(x_s, x_t))\right) \dots\dots\dots (II.14)$$

Ces simulations correspondent à différentes valeurs de β selon la probabilité a priori. Elles ont été réalisées par l'échantillonneur de Gibbs utilisant 20 balayages déterministes.

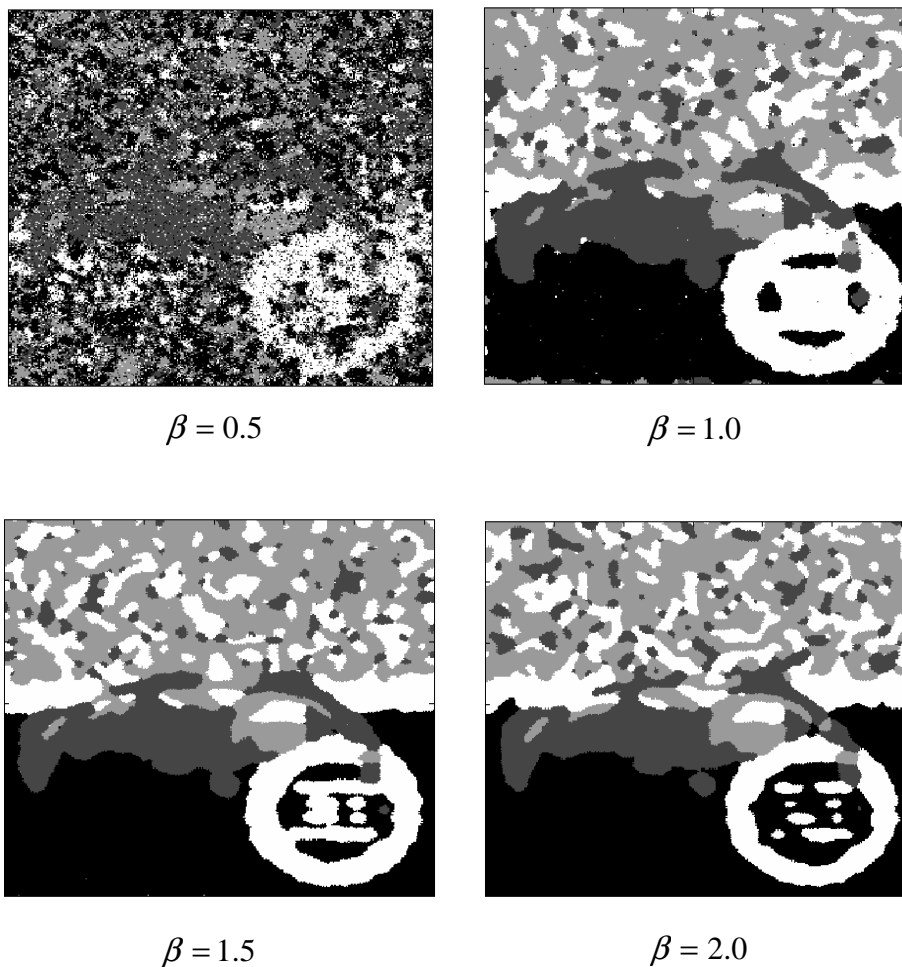


FIG.II.5. - Exemple de simulation d'un champ de Markov à 4 classes par échantillonneur de Gibbs pour différentes valeurs de β selon la probabilité a priori (8 pixels voisins).

Remarque

La différence avec l'échantillonneur de Gibbs réside dans le tirage au sort du nouveau niveau de gris (ou descripteur), au lieu de considérer la loi définie par tous les descripteurs. Comme on ne considère que la variation énergétique entre les 2 configurations, l'algorithme de Metropolis soit plus rapide à chaque étape que l'échantillonneur de Gibbs, qui lui nécessite le calcul de la fonction de partition locale. Mais la convergence peut être plus lente car le taux d'acceptation est strictement inférieur à 1 (les transitions ne sont pas toujours acceptées, contrairement au cas de l'échantillonneur de Gibbs).

II.3.3. Conclusion

Les algorithmes décrits dans cette section permettant donc la simulation des réalisations d'un champ de Markov. Nous optons, pour la suite de cette thèse, pour l'échantillonneur de Gibbs.

II.4. Modélisation markovien

L'objet de cette section est de présenter la modélisations par les champs de Markov cachés (CMC) que l'on appellera également "modèle classique", nous commençons par la définition de ce modèle, ainsi le graphe d'indépendance et champs de Markov cachés à bruit indépendant. Nous présentons ensuite l'application de ce modèle en segmentation d'image. Enfin nous exposons une réalisation de ce dernier.

II.4.1. Champs de Markov cachés

Soit S un ensemble de pixels de cardinal N ($N = \text{Card}(S)$), et $X = (X_s)_{s \in S}$, $Y = (Y_s)_{s \in S}$ deux champs aléatoires, l'un caché et l'autre observé. Chaque variable aléatoire X_s est à valeurs dans l'ensemble fini de classes $\Omega = \{\omega_1, \dots, \omega_k\}$ et chaque variable aléatoire Y_s est à valeurs dans \mathfrak{R} . En segmentation d'images, l'objectif est de retrouver le champ inobservé X , à partir des données observées Y qui représentent la version bruitée du champ X . A cet effet, nous devons définir la loi de (X, Y) , ce qui est classiquement obtenu à partir de la loi de X et celle de Y conditionnelle à X .

Dans le cas de la modélisation par champs de Markov cachés (CMC), X est supposé markovien. Sa loi est alors donnée par :

$$p(x) = \gamma \exp(-U(x)) \dots\dots\dots (II.15)$$

avec γ une constante de normalisation, appelée également "fonction de partition", et U "la fonction énergie". Celle-ci se décompose en une somme de fonctions φ_c , appelées "fonctions de potentiel", associées à l'ensemble des cliques. Ainsi, $U(x)$ s'écrit :

$$U(x) = \sum_{c \in C} \varphi_c(x_c) \dots\dots\dots (II.16)$$

La fonction de partition γ est généralement impossible à calculer du fait du nombre trop important des réalisations possibles du champ X ; cependant, il reste possible de simuler une réalisation de ce champ à partir de la loi conditionnelle de X_s sachant X_{V_s} tel que X_{V_s} est la restriction du champ X au voisinage V_s . L'échantillonneur de Gibbs que nous allons utiliser au cours de cet thèse et qui a été détaillé au (§ II.3.2), permet cette simulation lorsque la loi conditionnelle est connue. Cette loi est donnée par :

$$p(x_s / x_{V_s}) = \frac{\exp- \sum_{c \in C, s \in c} \varphi_c(x_s, x_{c-\{s\}})}{\sum_{\omega \in \Omega} \exp- \sum_{c \in C, s \in c} \varphi_c(x_s = \omega, x_{c-\{s\}})} \dots\dots\dots (II.17)$$

Ainsi, dans le cas des champs de Markov cachés, le champ inobservé X est supposé markovien, ce qui justifie l'appellation CMC. La loi de X étant définie, il reste à définir la loi conditionnelle $p(y/x)$.

II.4.2. Le graphe d'un champ de Markov cachée

La FIG.II.6 montre le graphe d'indépendance associé aux champs de Markov cachés usuellement définis sur une grille bidimensionnelle. Les variables aléatoires X_s forment un champ de Markov caché [Ros 1999].

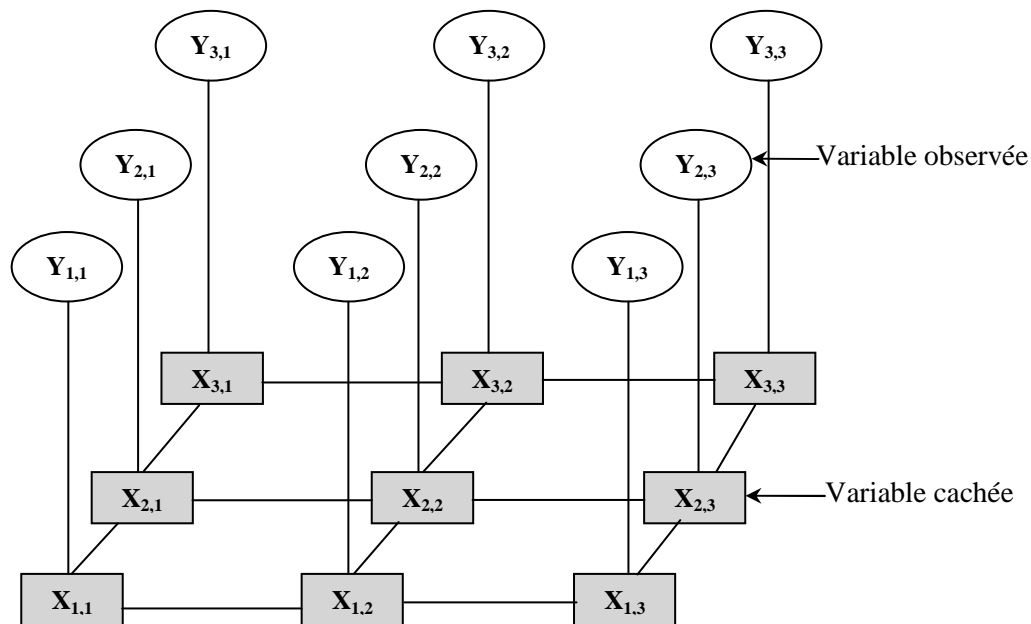


FIG.II.6 - La structure d'un modèle de champ de Markov caché.

II.4.3. Champs de Markov cachés à bruit indépendant

Les champs de Markov cachés à bruit indépendant (CMC-BI) sont la modélisation la plus simple des CMC. La probabilité conditionnelle $p(y/x)$ vérifie les deux hypothèses suivantes :

- **H1** : $p(y_s / x) = p(y_s / x_s)$ pour chaque $s \in S$
- **H2** : $p(y/x) = \prod_{s \in S} p(y_s / x)$

La deuxième hypothèse signifie que les Y_s sont indépendants conditionnellement à X .

Du fait de H1 et H2, la loi de $p(y/x)$ s'écrit sous la forme :

$$\begin{aligned}
 p(y/x) &= \prod_{s \in S} p(y_s / x_s) \\
 &\dots\dots\dots (II.18) \\
 &= \exp \left[\sum_{s \in S} \log(p(y_s / x_s)) \right]
 \end{aligned}$$

La loi du couple (X,Y) résulte des formules (II-15), (II-16) et (II-18). Elle s'écrit comme suit :

$$\begin{aligned}
 p(x/y) &= p(x)p(y/x) \\
 &\dots\dots\dots (II.19) \\
 &= \gamma \exp \left[- \sum_c \varphi_{c \in C}(x_c) + \sum_{s \in S} \log(p(y_s / x_s)) \right]
 \end{aligned}$$

La probabilité *a posteriori* $p(x/y)$ du champ X est alors donnée par :

$$\begin{aligned}
 p(x/y) &= \gamma(y) \exp(-U(x, y)) \\
 &\dots\dots\dots (II.20) \\
 &= \gamma(y) \exp \left[- \sum_{c \in C} \varphi_{c \in C}(x_c) - \sum_{s \in S} \log(p(y_s / x_s)) \right]
 \end{aligned}$$

Avec

$$\gamma(y) = \sum_{x \in \Omega^N} \exp \left[- \sum_{c \in C} \varphi_{c \in C}(x_c) + \sum_{s \in S} \log(p(y_s / x_s)) \right] \dots\dots\dots (II.21)$$

Cette formulation montre que la distribution *a posteriori* du champ X reste une distribution de Gibbs, ce qui signifie que la loi du champ caché conditionnellement à l'observation est également une loi de Markov. Cela rend possible la simulation des réalisations de X selon sa probabilité *a posteriori*, et permet ainsi d'utiliser différentes méthodes de segmentation bayésienne.

II.4.4. Champs de Markov cachés en segmentation d'images

Le problème posé ici est de trouver une réalisation x , modélisée par un processus markovien X , correspondant à une image d'étiquettes, chaque étiquette correspondant à un objet présent dans y : c'est le problème de segmentation d'images.

Dans un cadre bayésien, on peut rechercher la configuration \hat{x} maximisant la probabilité $p(X = x/Y = y)$, qui s'écrit :

$$p(\mathbf{X} = \mathbf{x} | \mathbf{Y} = \mathbf{y}) = \frac{p(\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{x})p(\mathbf{X} = \mathbf{x})}{p(\mathbf{Y} = \mathbf{y})} \dots\dots\dots (II.22)$$

- Le premier terme du numérateur décrit la probabilité d'observation y , sachant que l'image idéal est x : il modélise donc l'acquisition de l'image sous l'hypothèse d'indépendance des pixels,

$$p(\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{x}) = \prod_{s \in S} p(\mathbf{Y}_s = \mathbf{y}_s | \mathbf{X}_s = \mathbf{x}_s) \dots\dots\dots (II.23)$$

- Le second terme du numérateur décrit la probabilité d'existence de l'image idéale x , qui dans le cadre qui nous intéresse, répond à l'hypothèse markovienne selon un système de voisinage V et un modèle ϕ dépendant de l'application (Dans le cas de segmentation).

$$p(\mathbf{X} = \mathbf{x}) = \frac{e^{-U(x)}}{\sum_{z \in \Omega} e^{-U(z)}} \dots\dots\dots (II.24)$$

- Le dénominateur est constant, et en particulier indépendant de x . Ainsi

$$\begin{aligned}
 p(\mathbf{X} = \mathbf{x} \mid \mathbf{Y} = \mathbf{y}) &= K \cdot p(\mathbf{Y} = \mathbf{y} \mid \mathbf{X} = \mathbf{x}) p(\mathbf{X} = \mathbf{x}) \\
 &= K \cdot e^{\ln(P(\mathbf{Y}=\mathbf{y}|\mathbf{X}=\mathbf{x}))-U(\mathbf{x})} \\
 &= K \cdot e^{-U(\mathbf{x}/\mathbf{y})} \dots\dots\dots (II.25)
 \end{aligned}$$

où

$$U(x/y) = -\sum_{s \in S} p(\mathbf{Y}_s = \mathbf{y}_s \mid \mathbf{X}_s = \mathbf{x}_s) + \sum_{c \in C} U_c(x) \dots\dots\dots (II.26)$$

La distribution $p(X = x/Y = y)$ est donc une distribution de Gibbs, et par le théorème d’Hammersley - Clifford, le champ X conditionnellement à y est un champ de Markov.

II.4.5. Spécification de $U(\mathbf{x}/\mathbf{y})$ pour le problème de segmentation d’image

Dans ce cas, X est défini sur l’espace des étiquettes, et Y sur l’espace de ses descripteurs. Le terme $p(Y = y/X = x)$ traduit donc la probabilité d’observation de l’image y connaissant l’appartenance de chaque pixel à un objet présent dans la scène. En supposant que y_s ne dépend que de x_s , on peut écrire :

$$p(\mathbf{Y} = \mathbf{y} \mid \mathbf{X} = \mathbf{x}) = \prod_{s \in S} p(\mathbf{Y}_s = \mathbf{y}_s \mid \mathbf{X}_s = \mathbf{x}_s) = \prod_{s \in S} f_{x_s}(y_s) \dots\dots\dots (II.27)$$

Les valeurs de probabilité conditionnelles sont données par la fréquence d’observation des niveaux de gris pour une classe donnée, et si chaque classe k suit une distribution gaussienne de moyenne m_k et d’écart-type σ_k :

$$f_{x_s}(y_s) = p(Y_s = y_s / x_s = k) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left[-\frac{(y_s - m_k)^2}{2\sigma_k^2}\right] \dots\dots\dots (II.28)$$

Si on fait encore une fois l'hypothèse markovienne sur X et si on se restreint aux cliques d'ordre 2

$$p(X = x) = \frac{e^{-\beta \sum_{s,t \in C_2} \phi(x_s, x_t)}}{e^{-\beta \sum_{z \in \Omega, s,t \in C_2} \phi(z_s, z_t)}} \dots\dots\dots (II.29)$$

et donc :

$$U(x/y) = \sum_{s \in S} \frac{(y_s - m_{x_s})^2}{2\sigma_{x_s}^2} + \log(2\pi)\sigma_{x_s} + \beta \sum_{s,t \in C_2} \phi(x_s, x_t) \dots\dots\dots (II.30)$$

où

$$\phi(x_s, x_t) = \begin{cases} 1 & \text{si } x_s \neq x_t \\ 0 & \text{si } x_s = x_t \end{cases}$$

Le terme d'ordre un exprime l'attache aux données, et le terme d'ordre deux régularise l'image. ϕ est le modèle de Potts, pour favoriser les grandes régions homogènes dans l'image.

II.4.6. La réalisation d'un modèle de champ de Markov caché

Nous présentons sur les figures **FIG.II.7** et **FIG.II.8** des réalisations de champs de Markov de loi $p(x/y)$ définie par l'équation (II.30) en 4 connexités et en 8 connexités. Ces simulations correspondent à différentes valeurs de β selon la probabilité a posteriori. Elles ont été réalisées par l'échantillonneur de Gibbs utilisant 20 balayages déterministes.

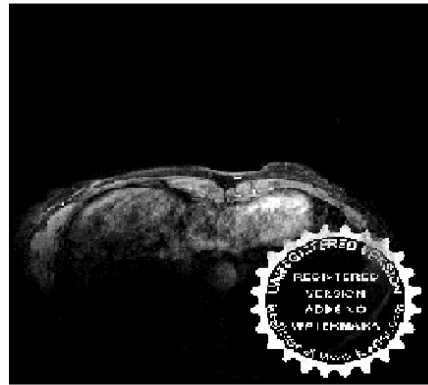
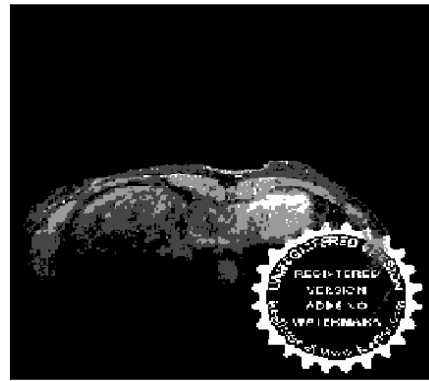


Image Originale



$\beta = 0.8$



$\beta = 1.4$



$\beta = 2.0$



$\beta = 2.5$

FIG.II.7. - Exemple de simulation d'un champ de Markov cachée à 4 classes par échantillonneur de Gibbs pour différentes valeurs de β selon la probabilité a posteriori d'une Image mammographie (4 pixels voisins).

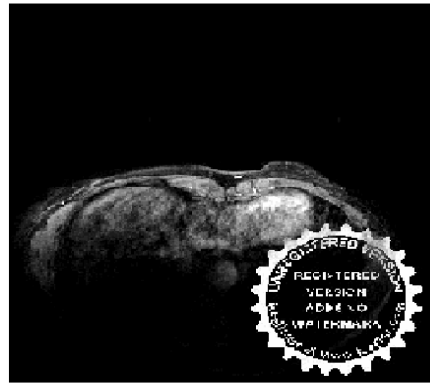
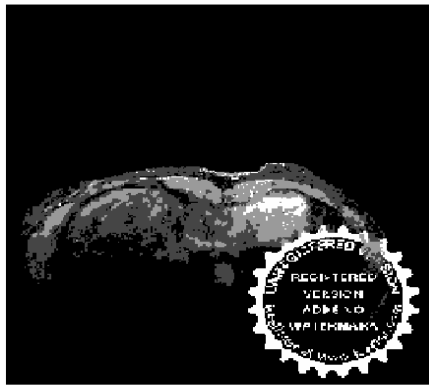
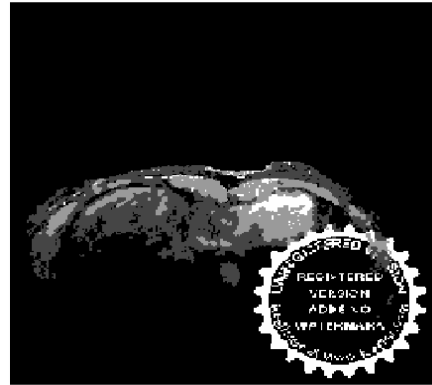


Image Originale



$\beta = 0.8$



$\beta = 1.4$



$\beta = 2.0$



$\beta = 2.5$

FIG.II.8. - Exemple de simulation d'un champ de Markov cachée à 4 classes par échantillonneur de Gibbs pour différentes valeurs de β selon la probabilité a posteriori d'une Image mammographie (8 pixels voisins).

II.4.7. Discussion

II.4.7.1. Role du choix des cliques

D'après les deux cas ci-dessus, on peut conclure qu'il faut prendre des cliques sur toutes les directions pour approcher une certaine isotropie. Nous considérons dans un premier temps à 4-connexité, on trouve un résultat un peu meilleur qu'en 8-connexité mais les contours sont marqués, par contre avec les 8 voisins, on trouve des images mieux régularisées. Nous pouvons conclure qu'un choix de voisinage plus grand donne une image plus régulière.

II.4.7.2. Role de la constante de couplage β

Lors de l'utilisation du modèle de champs de Markov caché nous devons fixer un certain nombre des coefficients du modèle. Un des coefficients à fixer est le coefficient β qui pondère la partie de l'énergie. Correspondant au modèle de Potts. Nous illustrons sur les deux cas ci-dessus le rôle de ce coefficient en choisissant différentes valeurs : $\beta = 0.8$, $\beta = 1.4$, $\beta = 2.0$ et $\beta = 2.5$.

- Le cas $\beta = 0.8$ donne un résultat un peu meilleur dans le 4- connexités, il présente des nouvelles zones blanches. Mais avec 8 pixels voisins les images sont mieux régularisés
- Le cas $\beta = 1.4$ donne un bien meilleur résultat lorsque l'on utilise le 8-connexité. Les petites zones ont disparue. On obtient des zones uniformes. On remarque aussi que le résultat obtenu avec 4- connexité est moins bon.
- Les deux cas $\beta = 2.0$ et $\beta = 2.5$ donnent des bons résultats dans les deux connexités.

La comparaison des quatre cas présentés permet d'illustrer la raison pour la quelle le coefficient β et parfois appelé facteur d'échelle. Lorsque le coefficient grandit les zones plus petites qu'une certaine échelle n'apparaissent plus dans la modélisation.

II.5. Estimation bayésienne

Soit $X = (X_s)_{s \in S}$ et $Y = (Y_s)_{s \in S}$ deux champs aléatoires comme décrit ci-dessus. Le problème de la segmentation bayésienne consiste en l'estimation de la réalisation invisible X à partir des données observées $Y = y$. Ainsi, le problème est de déterminer une estimation $\hat{x} \in \Omega$ de x à partir de y , obtenue en optimisant un certain critère.

L'estimation bayésienne nécessite la spécification d'une "fonction de coût", L définie dans $\Omega^{Card(S)} \times \Omega^{Card(S)} \rightarrow R^+$. Celle-ci possède les propriétés suivantes :

$$\forall x, \hat{x} \in \Omega^{Card(S)} \times \Omega^{Card(S)} :$$

$$\begin{aligned}
 &L(x, \hat{x}) \geq 0; \\
 &..... (II.31) \\
 &L(x, \hat{x}) = 0 \Leftrightarrow x = \hat{x}
 \end{aligned}$$

Le risque bayésien associé à la stratégie $\hat{s} : R^{Card(S)} \rightarrow \Omega^{Card(S)}$ est donné par le coût moyen $R = E[L(\hat{s}(Y), X)]$. La stratégie bayésienne \hat{s}_B est une stratégie dont le risque bayésien est minimum :

$$E[L(\hat{s}(Y), X)] = \min_{\hat{s}} E[L(\hat{s}(Y), X)]..... (II.32)$$

L'estimateur bayésienne \hat{s}_B est alors obtenu en minimisant l'espérance de cette fonction de coût conditionnellement aux observations :

$$\hat{s}_B(y) = \arg \min_{\hat{x} \in \Omega} E[L(\hat{x}(y), X) / Y = y]..... (II.33)$$

A chaque fonction de coût correspond ainsi un estimateur bayésien. Dans ce qui suit nous nous intéressons principalement à deux fonctions de coût, très utilisées, qui définissent les estimateurs les plus répandus dans la littérature, à savoir l'estimateur du "maximum a posteriori" (MAP) et l'estimateur du "mode des marginales a posteriori" (MMP), très connu

sous l'abréviation MPM (en anglais). Aussi, nous nous intéressons à l'estimateur "iterated conditional mode" (ICM) qui est une approximation du MAP.

Le choix d'un estimateur est souvent laissé à l'appréciation de l'utilisateur. Nous détaillons, ci après, les estimateurs les plus répandus, énumérés plus haut.

II.5.1. Estimateur Maximum A Posteriori (MAP)

L'estimateur MAP est associé à la fonction de coût suivante :

$$L(\hat{x}, x) = 1 - \delta(\hat{x}, x) \dots\dots\dots (II.34)$$

où la fonction δ est définie par $\delta(x_s, s_t) = 0$ pour $x_s \neq s_t$ et $\delta(x_s, s_t) = 1$ pour $x_s = s_t$. Cette fonction de coût est sévère, car elle pénalise toute différence entre deux configurations de manière identique. Autrement dit, elle pénalise de façon identique une erreur sur un site et une erreur sur plusieurs sites. Nous pouvons alors écrire :

$$\begin{aligned} E[L(\hat{x}, x) / Y = y] &= \sum_{x \in \Omega} L(\hat{x}, x) p(x / y) \dots\dots\dots (II.35) \\ &= 1 - p(\hat{x} / y) \end{aligned}$$

Par conséquent, l'estimation bayésienne est :

$$\hat{x} = \arg \min_{\hat{x} \in \Omega} [1 - p(x / y)] = \arg \max_{\hat{x} \in \Omega} [p(x / y)] \dots\dots\dots (II.36)$$

L'estimation au sens du MAP, revient donc à maximiser la probabilité a posteriori. Par ailleurs, cet estimateur présente certains inconvénients : Le premier, bien argumenté dans l'article de Bouman et Shapiro [BoS 1994], est dû aux méthodes d'approximations, du fait que le calcul de l'estimateur par MAP ne peut pas se faire directement. En effet, le calcul du MAP nécessite la minimisation d'une fonction énergie, souvent non convexe, ce qui a pour effet la présence de minima locaux.

Une minimisation exacte est impossible, d'où la nécessité de faire appel à des méthodes d'approximations, tel que le ICM, décrit dans la sous section suivante. Le deuxième, avancé par Marroquin [MVR 2001], est lié à la fonction de coût. Celle-ci n'est pas très appropriée du fait qu'elle attribue le coût "0" pour la solution parfaite uniquement et le coût "1" pour toutes les autres solutions, sans tenir compte du nombre des pixels mal classés. Ceci rend l'estimateur MAP très sélectif, fournissant ainsi de mauvais résultats, en particulier lorsque le bruit est fort [MMP 1987]. Cependant, Winkler [Win 2003] estime que les résultats insatisfaisants obtenus par segmentation MAP ne sont pas dus à l'estimateur lui-même, mais plutôt au choix de la loi a priori. Ainsi, selon Winkler, les probabilités a priori sont, en pratique, fréquemment choisies de telle sorte que l'estimateur MAP soit facilement calculable, sans correspondre nécessairement aux caractéristiques de l'image.

II.5.2. Estimateur du Mode Conditionnel Itéré (ICM)

L'estimateur du mode conditionnel itéré (ICM) a été proposé comme une méthode itérative d'approximation de l'estimateur MAP [Bes 1986]. C'est un algorithme déterministe nécessitant une bonne valeur d'initialisation. Son principe est fondé sur la maximisation des probabilités conditionnelles locales en chaque site séquentiellement. En d'autres termes, en chaque site s on retient la classe qui maximise cette probabilité.

L'algorithme ICM se résume comme suit :

- Initialiser une première carte "image" x_0 d'une façon arbitraire.
- A chaque itération n :
 - Balayer l'ensemble des sites $s \in S$ de x_{n-1} et en chaque site on calcule la probabilité conditionnelle, on pose :

$$x_s^n = \arg \max_{x_s \in \Omega} p(x_s / x_{V_s}^{n-1}) \dots \dots \dots \text{(II.37)}$$

- Refaire l'étape n jusqu'à réalisation d'un critère d'arrêt (par exemple lorsque le nombre de changements d'une étape à une autre devient suffisamment faible).

Remarque :

L'algorithme ICM converge vers le premier minimum local rencontré à partir des conditions initiales. Des travaux ont été faits pour améliorer la performance de cet algorithme, et une des voies de recherche est la quête d'une bonne initialisation.

II.5.3. Estimateur du Mode des Marginales à posteriori (MMP)

Pour éviter la sévérité de la fonction de coût de l'estimateur MAP, une autre fonction de coût moins restrictive, est associée à l'estimateur MPM. Celle ci est donnée comme suit :

$$L(\hat{x}_s, x_s) = \sum_{s \in S} 1 - \delta(\hat{x}_s, x_s) \dots\dots\dots (II.38)$$

Cette fonction consiste à pénaliser l'erreur commise en fonction du nombre de sites mal estimés (ou mal classés). On montre que l'estimation Bayésienne \hat{x} est ici obtenue suivant :

$$\hat{x}_s = \arg \max_{\hat{x}_s} p(x_s / y) \dots\dots\dots (II.39)$$

Cette estimation ressemble à l'estimation du MAP, mais opérée de façon locale. Autrement dit, on passe de la probabilité conditionnelle globale d'une configuration, à la probabilité conditionnelle en un site. Ainsi, la configuration optimale est atteinte lorsque toutes les lois marginales en chaque site sont maximisées. Cependant, le calcul direct et exact des probabilités a posteriori $p(x_s / y)$ est impossible, compte tenu du gigantisme de l'espace des configurations (§ II.2.3), mais le fait de pouvoir simuler des réalisations de X par des approximations de type Monte-Carlo décrites dans (§ II.3), permet leur estimation. En effet, on peut poser :

$$\hat{p}(x_s = \omega / y) = \frac{1_{[x_s^1 = \omega]} + \dots + 1_{[x_s^N = \omega]}}{N} \dots\dots\dots (II.40)$$

Finalement, pour la segmentation MPM, on procède de la manière suivante :

- Simuler N réalisations x^1, x^2, \dots, x^N de X selon la probabilité $Y = y$ en utilisant L'échantillonneur de Gibbs (décrit dans § II.3.2) ;
- Estimer à partir des réalisations x^1, x^2, \dots, x^N la loi de chaque X_s par les fréquences (équation (II.40)) ;
- Retenir la classe maximisant la loi ainsi obtenue. En d'autres termes choisir la classe dont le nombre d'apparitions dans les simulations est le plus grand.

II.5.4. Comparaison

Ces deux estimateurs sont ici comparés dans le cadre de segmentation. Dans le cas du MAP, les résultats sont obtenus par ICM. L'image à segmenter est une image mammographie. L'énergie a posteriori utilisée s'écrit :

$$U(x/y) = \sum_{s \in S} \frac{(y_s - m_{x_s})^2}{2\sigma_{x_s}^2} + \log(2\pi)\sigma_{x_s} + \beta \sum_{s,t \in C_2} \phi(x_s, x_t)$$

avec

$$\phi(x_s, x_t) = \begin{cases} 1 & \text{si } x_s \neq x_t \\ 0 & \text{si } x_s = x_t \end{cases}$$

L'initialisation est donnée par l'image à segmenter. Pour tous les estimateurs, on réalise 20 itérations correspondant à différentes valeurs de β en 4 - et en 8- connexités.

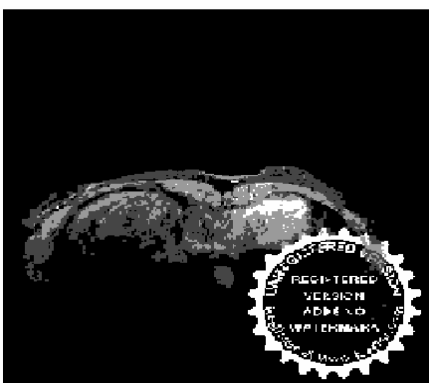
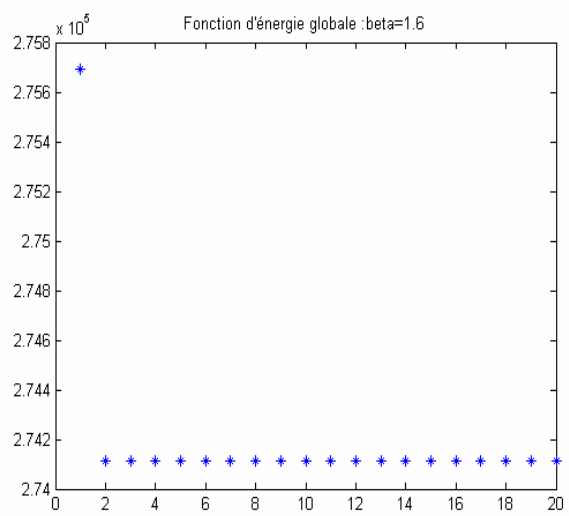
Visuellement, le meilleur résultat est obtenu par le MPM pour les différentes valeurs de β et en deux connexités, l'estimateur MPM est par définition plus local, donne un résultat plus régularisé avec une minimisation exponentielle de la fonction d'énergie global, qui augmentée avec l'augmentation de β . Par ailleurs, on constate que le résultat obtenu par le MAP du ICM est un peu meilleur qu'en MPM et moins régularisée. Nous remarquons que si les étiquettes du voisinage d'un site ne change pas, alors l'étiquette du site ne sera pas modifiée, c'est un inconvénient majeur de cette méthode qu'il exige de bon balayage. Nous optons, pour la suite de cette thèse, pour MPM.



Image Originale



ICM : $\beta = 1.6$



ICM : $\beta = 2.4$

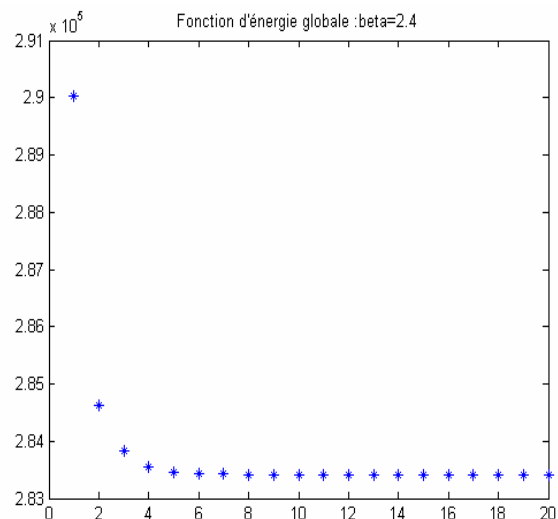


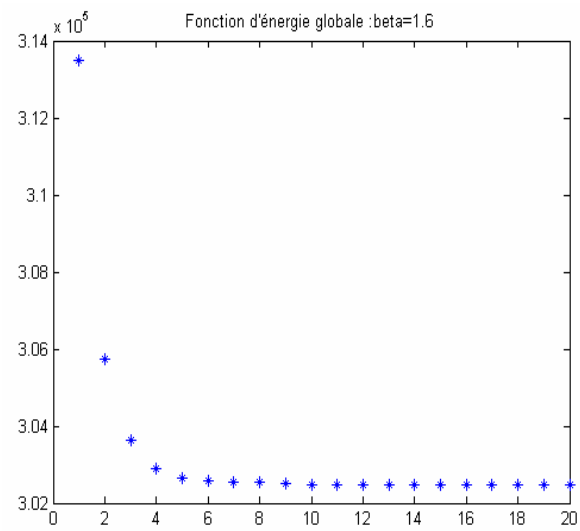
FIG.II.9. - Résultats de classification par ICM pour l'image mammographie correspond à différentes valeurs de β (4 pixels voisins).



Image Originale



ICM : $\beta = 1.6$



ICM : $\beta = 2.4$

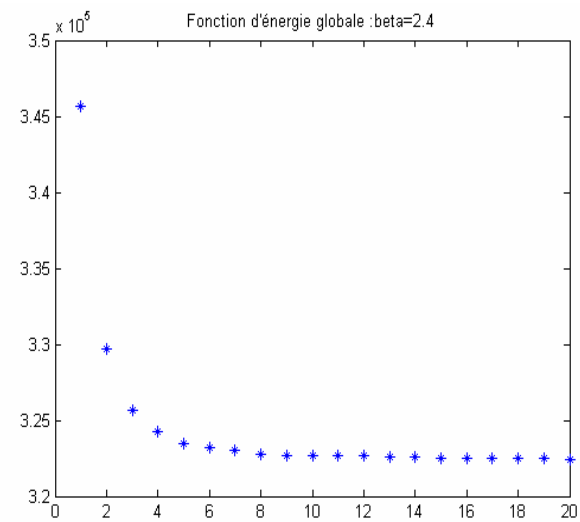


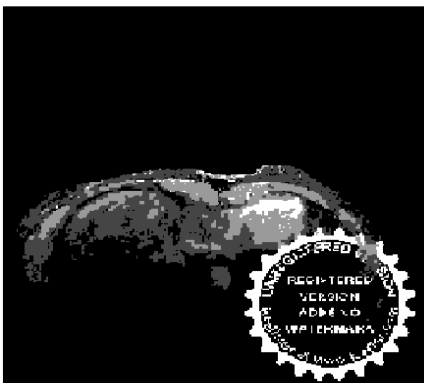
FIG.II.10. - Résultats de classification par ICM pour l'image mammographie correspond à différentes valeurs de β (8 pixels voisins).



Image Originale



MPM : $\beta = 1.6$



MPM : $\beta = 2.4$

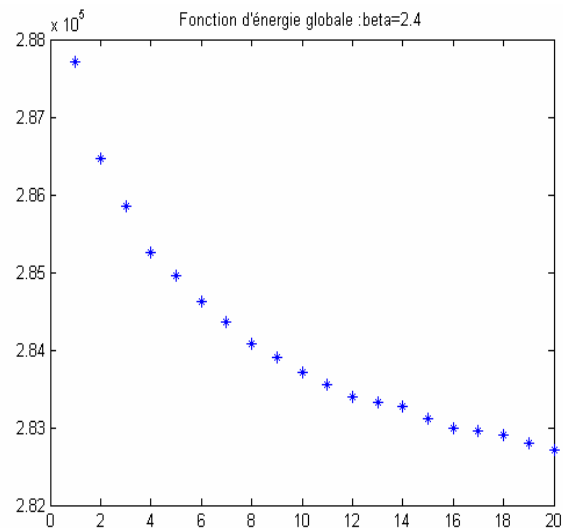
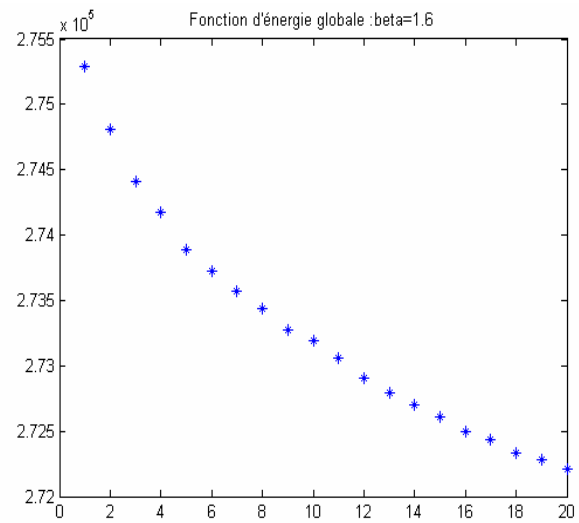


FIG.II.11. - Résultats de classification par MPM pour l'image mammographie correspond à différentes valeurs de β (4 pixels voisins).

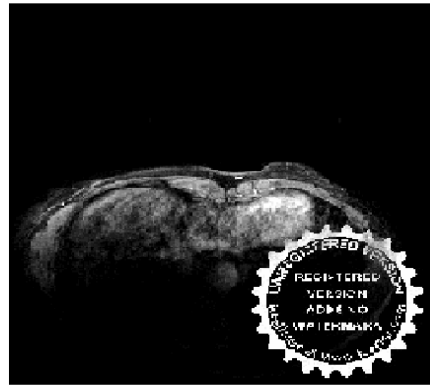
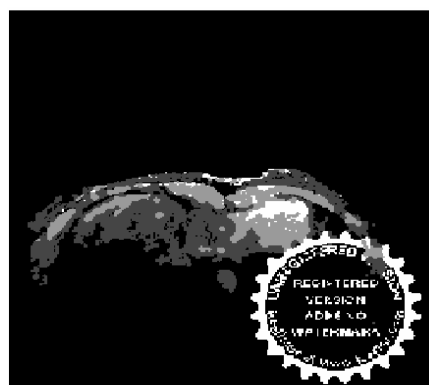


Image Originale



MPM : $\beta = 1.6$



MPM : $\beta = 2.4$

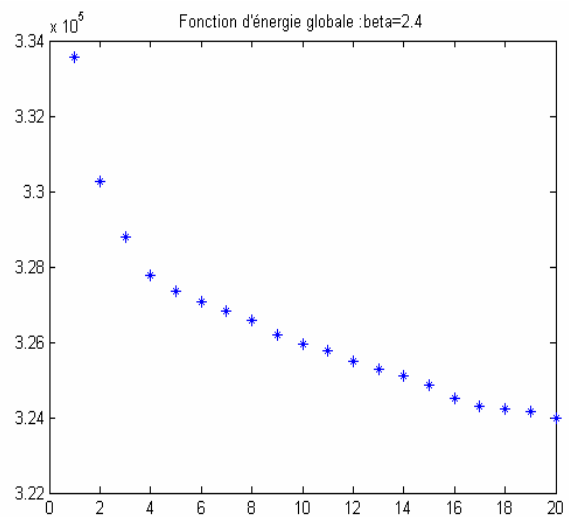
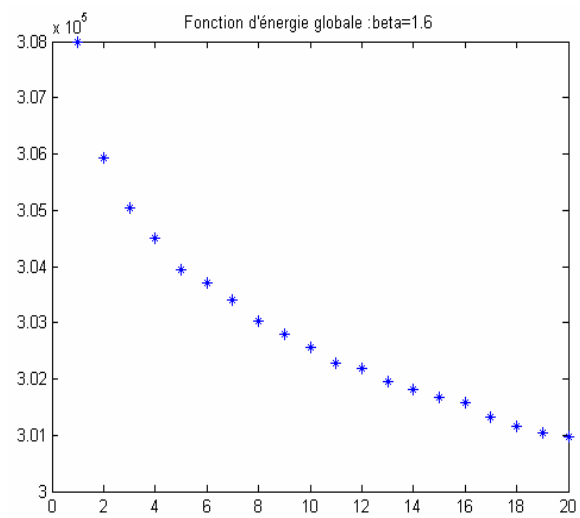


FIG.II.12. - Résultats de classification par MPM pour l'image mammographie correspond à différentes valeurs de β (8 pixels voisins).

II.6. Conclusion

La modélisation par champ de Markov caché rend compte des propriétés spatiales a priori des images observées, cette modélisation présente deux intérêts majeurs : (1) tout d'abord, il s'agit d'un modèle facilement paramétrable ; (2) à cette souplesse s'ajoute l'avantage d'une formulation immédiate du terme de probabilité conjointe du champ d'étiquettes grâce au théorème de Hammersley- Clifford qui relie propriétés locales (marginales) et probabilités globales du champ.

Nous avons exposé dans ce chapitre la problématique générale de la segmentation statistique d'images, abordée par les champs de Markov cachés. Nous avons présenté les modèles classiques, ainsi que les principaux algorithmes de segmentation bayésienne. Les principales méthodes classiques d'estimation, permettant la conception des méthodes de segmentation non supervisée.

Le chapitre suivant présente une méthode de segmentation non supervisée markovienne.

CHAPITRE 3

SEGMENTATION NON SUPERVISÉE D'IMAGE

Les enjeux majeurs en analyse d'images, fondée sur un modèle probabiliste, tels que les champs Markoviens cachés considérés dans cette thèse, sont l'estimation des paramètres et la segmentation - thèmes qui seront développés de manière plus détaillée dans ce chapitre.

Le but de ce chapitre est d'introduire les différentes étapes de segmentation non supervisée. A cet effet, nous commençons par la synoptique de la segmentation. Nous présentons ensuite l'estimation des paramètres par une méthode originale fondée sur le principe général de "Iterative Conditional Estimation" (ICE [Pie 1992]), ainsi que l'algorithme d'estimateur du Mode des Marginales à posteriori (MPM). Enfin, nous testons et évaluons nos différents algorithmes sur des problèmes de segmentation non supervisée d'images synthétiques.

III.1. Introduction

La segmentation non supervisée d'images est parmi les problèmes clé en traitement d'images. Parmi les différents modèles et approches développées, une partie des méthodes statistiques, couramment utilisées, sont fondées sur les champs de Markov cachés (CMC). Ce succès est principalement dû à l'aptitude du modèle de prendre en compte des dépendances spatiales des variables aléatoires, même lorsqu'elles sont en très grand nombre, pouvant dépasser le million. Dans un tel prototype le champ caché X est supposé markovien et doit être estimé à partir du champ observé Y .

III.2. Synoptique de la segmentation

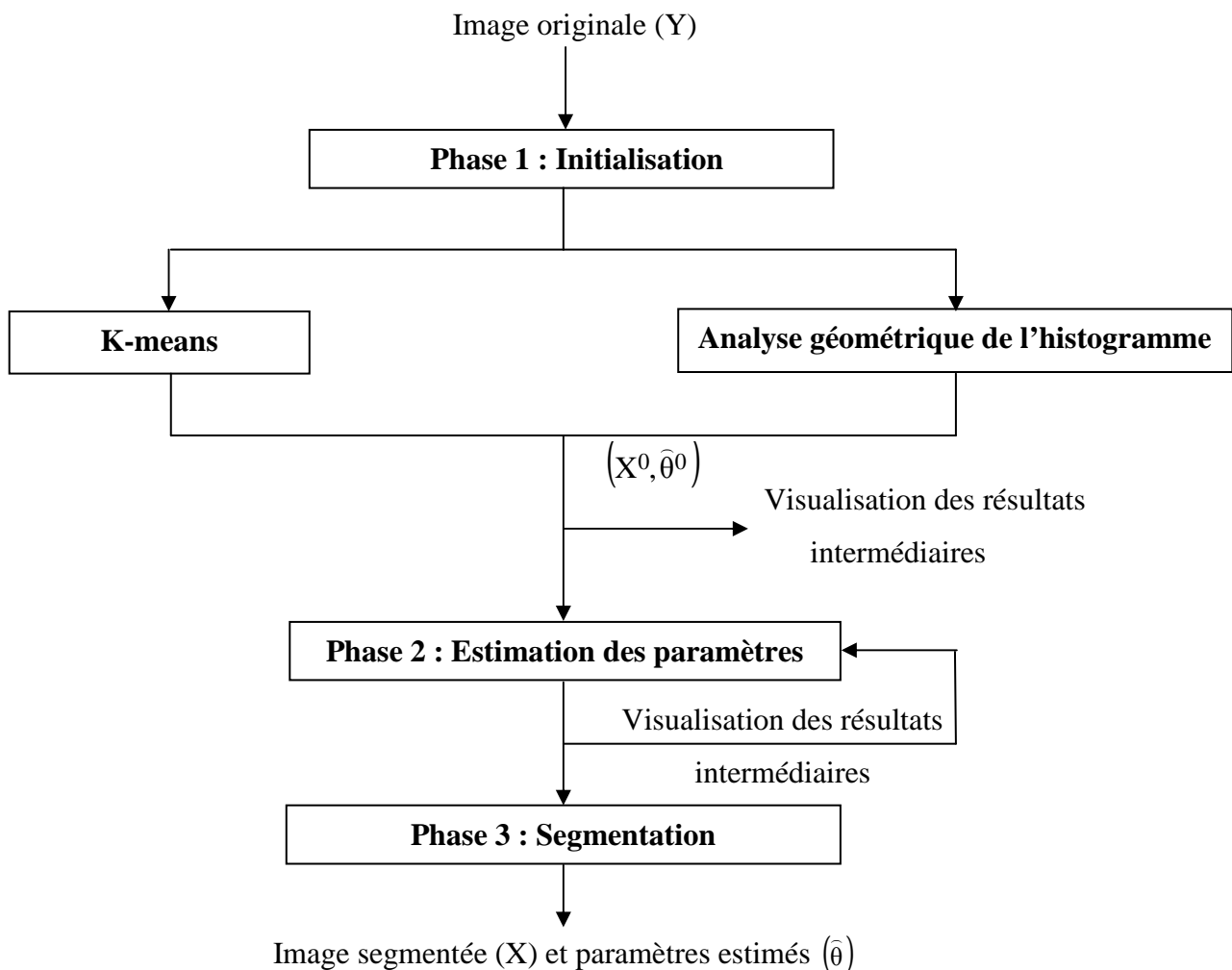


FIG.III.1 : Organigramme de la procédure de segmentation d'images

Selon la figure ci - dessus, l'algorithme global de la segmentation non supervisée se présente de la manière suivant :

Phase 1 : Initialisation

Initialisation de l'image d'entrée peut se faire de deux manières différentes :

- a. K-means.
- b. Analyse géométrique de l'histogramme.

Nous considérons seulement L'algorithme K-means.

Phase 2 : Estimation des paramètres

La phase d'estimation des paramètres est réalisée par l'algorithme ICE. Chaque itération fournit une estimation des paramètres. Quand l'estimée finale des paramètres θ_n est obtenue, on procède à la segmentation ordinaire (segmentation supervisée) avec des paramètres connus.

Phase 3 : Segmentation

Cette phase permet de segmenter l'image originale à partir des paramètres estimés θ_n (à la sortie de ICE). Dans un contexte bayésien, l'algorithme utilisé est le « Mode Marginales Posteriori » (MPM).

III.3.Segmentation non supervisée

III.3.1. Initialisation

L'objectif de cette phase est de fournir une première estimation des paramètres des modèles, en vue de la seconde phase. Pour cela, nous avons implémenter un algorithme de classification : l'algorithme des K-means. La méthode des k-means est une méthode itérative permettant de classer les pixels d'une image, en k classes ($k \in \Omega$), selon leur niveau de gris. Le pixel est affecté dans la classe pour laquelle la distance du centre de la classe au pixel est minimale.

Les étapes de l'algorithme K-means sont décrites comme suit :

- Initialiser k noyaux (centres des classes) d'une manière aléatoire ;
- A chaque itération n balayer l'image,
- Affecter x_s à la classe ω_i si

$$\|x_s - \mu(\omega_i)\| = \min_{i=1}^k \|x_s - \mu(\omega_i)\| \dots\dots\dots (III.1)$$

- Recalculer le centre de chaque classe par :

$$\mu(\omega_k) = \frac{1}{|\Omega|} \sum_{x \in \Omega} x \dots\dots\dots (III.2)$$

La phase d'initialisation est essentiellement identique dans les deux cas et les étapes qui la composent sont les suivantes :

- Lecture de l'image (Y).
- Calcul de l'histogramme de l'image.
- Segmentation initiale par l'algorithme classique des K-means.
- Estimation initiale des paramètres θ^0 à partir de l'image originale (Y) et de l'image segmentée initiale (X^0). Cette étape dépend du modèle de Markov. Elle est également utilisée dans la seconde phase (Estimation des paramètres) qui a été présentée précédemment.

L'image segmentée résultante de cette initialisation (X^0) conduit à une première estimation des paramètres du modèle (θ^0).

III.3.2. Estimation des paramètres

Dans la pratique, les paramètres de régularité et les paramètres des distributions des classes sont souvent inconnus et doivent être estimés à partir de l'observation $Y = y$. Le problème est alors double : nous ne connaissons pas les caractéristiques des classes et nous ne savons pas quels pixels sont représentatifs pour chaque classe.

Dans ce travail, nous considérons que l'attache aux données est une gaussienne dont ses paramètres sont : la moyenne et la variance des classes. Le modèle a priori est le modèle de Potts de paramètre de régularité β . Nous supposons que les différentes classes sont définies à partir des niveaux de gris uniquement, c'est à dire que nous ne tenons pas compte de critères de texturation. Nous considérons également que nous avons une seule image de données fixe.

Deux approches permettent de traiter le problème de l'estimation des paramètres : Une première approche dite supervisée suppose connue sur une partie de l'image, une réalisation de X . Une deuxième dite non supervisée. Dans ce cas, il faut déterminer uniquement à partir de l'observation Y les différents paramètres. Cette estimation plus difficile à mettre en œuvre, mais offrant à l'opérateurs une plus grande souplesse d'utilisation. L'apprentissage non supervisé d'un modèle correspond à l'estimation des paramètres le caractérisant. Ces paramètres se divisent en deux types : ceux associés au terme d'attache aux données (les probabilités conditionnelles) et ceux relevant du modèle à priori (les probabilités à priori). Une fois ces paramètres estimés, on peut procéder à l'étape de la segmentation.

Ainsi, l'apprentissage non supervisée des paramètres qui induit à une segmentation non supervisée est le thème central de cette étude.

Il y a plusieurs méthodes itératives pour l'estimation de ces paramètres, incluant l'algorithme EM (pour « Expectation-Maximisation ») [Del 1997], le SEM (pour « Stochastique EM ») [BGP 2000] [Pey 2001] et l'ICE (pour « Espérance Conditionnelle Itérative ») [PaP 1993]. Ici, Nous considérons seulement le dernier.

III.3.2.1. Principe de l'algorithme ICE

L'algorithme ICE est une méthode très générale d'estimation de paramètres. Elle a été introduite par W.Pieczynski en traitement d'images en 1996 [Char 2000] [Del 1997] [PaP 1993]. Elle apparaît plus générale car elle accepte toute forme de modélisation.

Son principe s'appuie sur l'hypothèse qu'il existe au moins un estimateur $\hat{\theta}(X, Y)$ de θ défini sur les données (X, Y) . On cherche alors la meilleure approximation de cet estimateur au sens de l'erreur quadratique moyenne qui n'est autre que l'espérance

conditionnelle par rapport à Y : $E[\hat{\theta}(X, Y)|y, \theta^{(q)}]$. Cette espérance dépend de θ et n'est pas calculable explicitement.

On utilise alors la procédure itérative suivante appelée ICE (pour «Espérance Conditionnelle Itérative») [MaP1991] [ABP 1992] :

- Initialisation de θ , soit $\theta^{(0)}$.
- $\theta^{(q+1)} = E\left[\hat{\theta}(X, Y)|y, \theta^{(q)}\right] = E\left[\arg \max_{\theta} \log (p(X, Y|\theta))|y, \theta^{(q)}\right] \dots\dots\dots (III.3)$
- Au sens de l'erreur quadratique moyenne, l'espérance conditionnelle $E\left[\hat{\theta}|Y = y^0\right]$ est la meilleure approximation. Mais celle-ci dépend elle-même de θ . L'algorithme adopte donc une démarche itérative pour approcher $E\left[\hat{\theta}|Y = y^0\right]$ et à chaque étape q , on recherche le paramètre $\theta^{(q+1)} = E\left[\hat{\theta}(X, Y)|Y = y^0, \theta^{(q)}\right]$. Notons que si l'espérance n'est pas explicitement calculable, on peut l'estimer par tirages aléatoires selon la loi conditionnelle à Y .

Alors nous allons dans la suite présenter l'algorithme ICE.

III.3.2.2. L'algorithme ICE

Dans le cas d'un champ de Markov caché, l'algorithme ICE crée itérativement des réalisations a posteriori et recalcule les paramètres de classes associés à $f_i^0 (\mu, \sigma)$ et de régularité spatiale (β). Dans le cadre d'estimation d'un mélange gaussien, les étapes de calcul suivantes sont effectuées pour chaque itération q et avec $1 \leq i \leq k$:

- Un certain nombre de réalisations a posteriori (avec l'indice r) sont simulées avec l'échantillonneur de Gibbs, en utilisant la formule (II.11) et les paramètres θ_i^{q-1} (défini par f_i^{q-1}) obtenus dans l'itération précédente.

- Les paramètres de classe $\theta_i^{q,r}$ (qui sont la valeur de la moyenne (μ_i^q) et la variance (σ_i^q) pour une classe de distribution gaussienne) sont estimés pour chaque réalisation, et puis moyennés pour avoir θ_i^q :

$$\mu_i^q = \frac{\sum_{n=1}^N y_n \delta(x_n - \omega_i)}{\sum_{n=1}^N \delta(x_n - \omega_i)} \dots\dots\dots (III.4)$$

$$(\sigma_i^q)^2 = \frac{\sum_{n=1}^N (y_n - \mu_i^q)^2 \delta(x_n - \omega_i)}{\sum_{n=1}^N \delta(x_n - \omega_i)} \dots\dots\dots (III.5)$$

- On a choisit d'estimer les paramètres de régularité β^q dans chaque itération de ICE avec une approche de gradient stochastique [Del, 1997], réalisée à travers une série de réalisations a priori calculées avec l'échantillonneur de Gibbs en utilisant la formule (II.11).

Dans le but de réduire le temps de calcul, généralement, on simule seulement une réalisation à posteriori pour chaque itération ICE et seulement une réalisation a priori pour chaque itération de gradient stochastique. Cette simplification ne cause aucune perte de performance significative.

On utilise une approximation de l'équation du gradient stochastique [Del 1997] [CoP 1995] [FjL 2000]. Soit $U_{\text{aposteriori}}^q$ l'énergie de la réalisation a posteriori courante et U_{apriori}^q l'énergie de la réalisation a priori dans l'itération r du gradient stochastique. Posant $\beta^{q,0}$ la valeur obtenue dans l'itération précédente β^{q-1} de ICE et on répète le calcul jusqu'à convergence :

$$\beta^{q,r} = \beta^{q,r-1} + \frac{1}{r} \frac{U_{\text{apriori}}^{q,r} - U_{\text{aposteriori}}^q}{U_{\text{aposteriori}}^q} \dots\dots\dots (III.6)$$

L'algorithme ICE nécessite une classe d'image initiale à partir de laquelle les paramètres de classe sont calculés. On a utilisé l'algorithme des K-means, qui subdivise itérativement les niveaux de gris en k classes distinctes. L'image est parcourue répétitivement jusqu'à la stabilité, attribuant chaque pixel à la classe ayant le centre le plus proche, et recalculer les centres de la classe à partir de tous les échantillons attribués à la fin de chaque itération. On note que cette méthode ne permette pas de distinguer facilement les différentes classes au sein de l'image. Son utilisation doit s'inscrire dans une procédure empirique afin de garantir un résultat robuste.

III.3.2. 3. Critère d'arrêt

Les algorithmes d'estimation de paramètres sont arrêtés lorsque les variations des paramètres de classes d'une itération à l'autre ne sont plus significatives. D'où le critère d'arrêt suivant :

$$\sqrt{\frac{1}{k} \cdot \sum_{i=1}^k (\mu_i^{(p+1)} - \mu_i^{(p)}) + \sum_{i=1}^k (\sigma_i^{(p+1)} - \sigma_i^{(p)})} < \varepsilon \dots\dots\dots (III.7)$$

Nous utiliserons $\varepsilon = 0.2$.

III.3.3. L'algorithme MPM pour la segmentation d'image

L'estimateur du MPM, pour les champs de Markov, a été introduit par Marroquin [Char 2000]. L'algorithme MPM consiste à estimer les lois marginales. Pour ce faire, il est nécessaire de générer m échantillons $\{x(1)...x(m)\}$ de la loi a posteriori. La loi des grands nombres devrait alors nous donner accès aux marginales :

$$P(X_s = x_s | Y = y) \approx \frac{1}{m - p} \sum_{q=p+1}^m \delta(x_s - x_s(q)) \dots\dots\dots (III.8)$$

où δ est le symbole de Kroneker. On rejette les p premiers échantillons car ils sont considérés dépendants de la configuration initiale. L'estimateur du MPM sera alors approché par le champ des étiquettes apparaissant le plus fréquemment dans la suite des (m-p) échantillons.

III.4. Critères d'évaluation

Il est nécessaire d'utiliser un critère numérique permettant une évaluation de la qualité des résultats. Pour évaluer nos résultats de segmentation, nous proposons d'exploiter deux critères utilisés généralement pour l'évaluation d'un résultat de segmentation dans un contexte non supervisé : l'un Le rapport signal sur bruit crête (PSNR). C'est un critère global, défini sur l'image complet ou sur une sous-image. L'autre la matrice de confusion. C'est une matrice ou un tableau affichant les statistiques de la précision de classification d'une image, notamment le degré de classification erronée parmi les diverses classes.

III.4.1. Le rapport signal sur bruit crête (PSNR)

L'opérateur PSNR mesure le rapport entre l'information et le bruit dans une image. Il est calculé à partir d'une image initiale img_E qui comporte l'image et le bruit et d'une image img_S qui est la version segmentée de l'image initiale img_E . Les images img_E et img_S doivent avoir la même dimension et le même type. Le PSNR s'agit de quantifier la performance des algorithmes en mesurant la qualité de réinsertion de l'image segmentée par rapport à l'image originale.

Le PSNR est défini par :

$$PSNR = 10 \cdot \log_{10} \left(\frac{d^2}{EQM} \right) \dots\dots\dots (III.9)$$

Où

- d est la dynamique du l'image. Dans le cas standard d'une image où les composantes d'un pixel sont codées sur 8 bits, $d = 255$
- EQM est l'erreur quadratique moyenne et est définie pour 2 images img_E et img_S de taille $m \times n$ comme :

$$EQM = \frac{1}{m \cdot n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|img_E(i, j) - img_S(i, j)\|^2 \dots\dots\dots (III.10)$$

Maximiser le PSNR revient à minimiser l'erreur quadratique. Les valeurs typiques de PSNR pour des images de bonne qualité varient entre 30 et 40 dB.

D'autres distances peuvent être utilisées, faisant par exemple intervenir la norme L^p au lieu de la norme L^2 . On peut ainsi mettre en oeuvre toute une famille d'indicateurs, certains étant plus sensibles aux grands écarts d'intensité, d'autres aux petits écarts. Cependant, le classement des images qui en résulte est tributaire du choix du critère : il est alors difficile de décider, dans l'absolu, quelle est la fonction de distance la plus pertinente. Ainsi, il semble impossible de mettre en oeuvre un critère en-dehors de tout contexte applicatif.

III.4.2. Matrice de confusion

La matrice de confusion est définie comme la répartition des pixels classés dans les différentes classes pour chacune des régions de référence [LiK 1979]. Nous définissons R_i , comme l'ensemble de pixels de référence (soit de la région d'apprentissage ou de la vérité terrain) appartenant à $i = 1, \dots, c$ et C_j , comme l'ensemble des pixels classés dans $j = 1, \dots, c$ où c est le nombre de classes.

Nous définissons la matrice $A = (A_{i,j})$. Soit :

$$A_{i,j} = \text{Card}(R_i \cap C_j) \dots\dots\dots (III.11)$$

Les coefficients $A_{i,j}$ correspondent au nombre de pixels appartenant à la région i classés dans j ; la matrice de confusion $M = (M_{i,j})$, est alors :

$$M_{i,j} = \frac{A_{i,j}}{\text{Card}(R_i)} \dots\dots\dots (III.12)$$

Enfin, la matrice de l'utilisateur est définie par :

$$U_{i,j} = \frac{A_{i,j}}{\text{Card}(C_j)} \dots\dots\dots (III.13)$$

Cette matrice nous permet d'observer la dispersion de la classification dans l'ensemble de référence pour toutes les classes.

A partir de la matrice de confusion, on tire différents indicateurs. En particulier, on peut lire cette matrice de deux manières différentes, selon que l'on s'intéresse aux lignes ou aux colonnes. Ces deux interprétations sont résumées par deux indicateurs, appelés précision utilisateur (« user's accuracy ») et précision producteur (« producer's accuracy »), qui sont intrinsèquement liés à chaque classe. Ils sont basés sur les valeurs de la diagonale et sur les marginales de la matrice. Par marginale, on entend la somme des coefficients sur chaque ligne i , notée A_{i+} , et la somme des coefficients sur chaque colonne j , notée A_{+j} (voir **Tab.III.1**).

Les expressions des indicateurs sont les suivantes :

- précision utilisateur : $P_{i+} = \frac{A_{ii}}{A_{i+}}$ correspond au pourcentage de pixels de l'ensemble de validation correctement classés (au sens de la vérité terrain) parmi tous ceux classés dans i par le classifieur;
- précision producteur : $P_{+i} = \frac{A_{ii}}{A_{+i}}$ correspond au pourcentage de pixels de l'ensemble de validation classés dans la classe i par la méthode de champ de Markov caché à la classe i par la vérité terrain.

		Classification Data				Total	P_{+j}
		C1	C2	C3	C4		
Reference Data	R1	A_{11}	A_{21}	A_{31}	A_{41}	$A_{+1} = \sum A_{1j}$	A_{11}/A_{+1}
	R2	A_{12}	A_{22}	A_{32}	A_{42}	$A_{+2} = \sum A_{2j}$	A_{22}/A_{+2}
	R3	A_{13}	A_{23}	A_{33}	A_{43}	$A_{+3} = \sum A_{3j}$	A_{33}/A_{+3}
	R4	A_{14}	A_{24}	A_{34}	A_{44}	$A_{+4} = \sum A_{4j}$	A_{44}/A_{+4}
Total		$A_{1+} = \sum_j A_{1j}$	$A_{2+} = \sum_j A_{2j}$	$A_{3+} = \sum_j A_{3j}$	$A_{4+} = \sum_j A_{4j}$	N	
P_{i+}		A_{11}/A_{1+}	A_{22}/A_{2+}	A_{33}/A_{3+}	A_{44}/A_{4+}		

Tab.III.1 - *Modèle de matrice de confusion à 4 entrées ($N = \sum_i a_{i+} = \sum_j a_{+j}$ est le nombre total de pixels dans les échantillons).*

L'interprétation des indicateurs par classe n'est pas simple. En effet, pour tirer une conclusion sur une classe, il faut considérer les deux indicateurs simultanément. Par exemple, il est possible que l'on ait un bon taux de classification producteur, mais que la précision utilisateur soit mauvaise ; cela signifie que la classe est très « contaminante », i.e., la carte de classification lui attribue plus de pixels qu'il ne faudrait. La précision utilisateur peut s'interpréter comme la probabilité qu'un pixel affecté à une classe représente bien le thème concerné sur le terrain.

Par ailleurs, on définit des indicateurs globaux de la qualité des classifications. Le plus répandu pour évaluer la qualité d'une classification est le taux global de classification (« overall accuracy ») qui calcule le pourcentage de pixels de référence ayant été bien classés (du moins au regard de la vérité terrain) :

$$\rho = \frac{1}{N} \sum_i A_{ii} \dots\dots\dots(III.14)$$

N est le nombre de pixels de référence classifiés dans l'image. Cependant, ρ ne prend pas en considération les coefficients non diagonaux de la matrice de confusion. C'est pour cette raison qu'un autre indicateur est aussi beaucoup utilisé : l'indicateur KAPPA, car celui-ci intègre les marginales de la matrice de confusion (et donc indirectement les coefficients non diagonaux) :

$$k = \frac{N \sum_{i=1}^c A_{i,i} - \sum_{k=1}^c (\sum_{i=1}^c A_{i,k} \cdot \sum_{j=1}^c A_{k,j})}{N^2 - \sum_{k=1}^c (\sum_{i=1}^c A_{i,k} \cdot \sum_{j=1}^c A_{k,j})} \dots\dots\dots(III.15)$$

Le coefficient Kappa k est un nombre réel, sans dimension et appartient à l'intervalle [0,1]. La valeur de kappa quantifie l'homogénéité globale du résultat de l'image segmentée. Une valeur de ce critère proche de 1 traduit un très bon résultat de segmentation.

III.5. Résultats sur d'images synthétiques

Nous avons travaillé sur deux types d'images synthétiques de taille 128×128 . La première **FIG.IV.2** et la seconde **FIG.IV.3** sont constituées de 4 classes et ont été bruitées en leur additionnant un bruit blanc gaussien de moyenne nul et d'écart type est la même pour toute l'image.

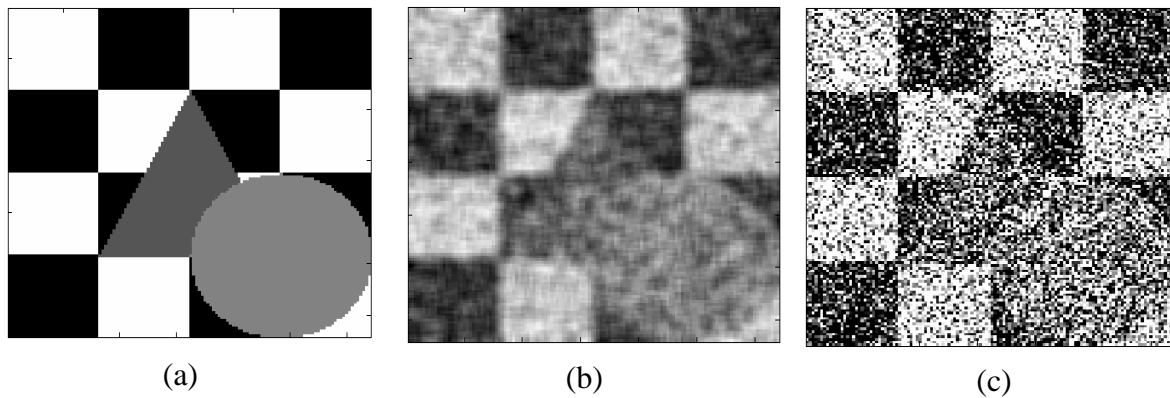


FIG.III.2 - (a) Image synthétique (1) originale; (b) Image bruitée avec flou et (c) Image bruitée sans flou.

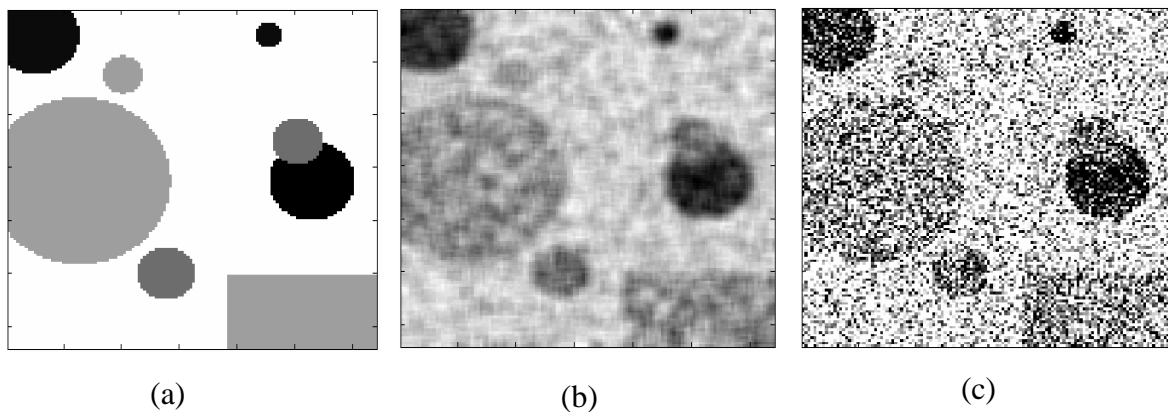


FIG.III.3 - (a) Image synthétique (2) originale; (b) Image bruitée avec flou et (c) Image bruitée sans flou.

III.5.1. Présentation des résultats

L'objectif de ces exemples étant d'obtenir des segmentations en composantes homogènes facilement exploitable et l'évaluation de résultat de segmentation des images synthétiques par des champs markoviens cachés.

Les segmentations de ces images obtenues à partir des champs de Markov cachés sont reportées sur les figures : FiG.III.4. (d), FiG.III.5. (d), FiG.III.6. (d) et FiG.III.7. (d). Nous avons utilisé comme modèle a priori un modèle de Potts simple avec un voisinage de huit connexions et un modèle gaussien pour le terme d'attache aux données de moyennes et de variances distinctes.

Les tableaux de Tab.III.2, Tab.III.6, Tab.III.10, Tab.III.14 résument les matrices de confusion pour chaque résultat de segmentation et les tableaux de Tab.III.3, Tab.III.7, Tab.III.11, Tab.III.15 condensent l'évaluation de résultat de segmentation.

Les tableaux de Tab.III.4, Tab.III.8, Tab.III.12, Tab.III.16 synthétisent les valeurs de précision globale ($\rho_{globale}$) et les tableaux de Tab.III.5, Tab.III.9, Tab.III.13, Tab.III.17 récapitulent les valeurs de précision moyenne ($\rho_{moyenne}$) pour chaque résultat de segmentation.

Le nombre d'itérations de l'algorithme d'ICE est mis à 30 itérations, avec 20 réalisations a posteriori par itération. La valeur initiale du paramètre de régularité du champ de Markov caché β est 0.5. Dans chaque itération d'ICE, le nombre maximal d'itérations pour le gradient stochastique est mis à 10, avec une réalisation a priori par itération. L'évaluation de gradient stochastique exige généralement très peu d'itérations. L'échantillonneur de Gibbs qu'on utilise a besoin de 100 itérations pour générer les réalisations a priori et a posteriori. La convergence des énergies globales est en fait tout à fait lente, particulièrement pour des réalisations selon la distribution a priori. La classification MPM basée sur les champs de Markov cachés compte sur 100 réalisations a posteriori.

Dans le but de réduire le temps de calcul, on a simulé seulement une réalisation a posteriori pour chaque itération d'ICE de la segmentation par champ de Markov cachés et seulement une réalisation a priori pour chaque itération du gradient stochastique pour

l'estimation du paramètre de régularité β de la segmentation par champ de Markov caché. Cette simplification ne cause aucune perte de performance significative.

III.5.2. Analyse et discussion

En générale, les résultats obtenus sont satisfaisants qui présentent des performances semblables en termes d'estimation des paramètres des modèles et en terme de segmentation. Si l'on analyse les variations des segmentations au point de vue visuel, on remarque qu'elles sont cohérentes. Mais on peut tirer quelques remarques :

- La première constatation est que les champs de Markov cachés fournit de bons résultats lorsque les images est sans flou .Cette impression est confirmée par les valeurs de précision (ρ), de kappa (k) et de PSNR associées à chaque résultat de segmentation présentés dans les tableaux : Tab.III.3, Tab.III.7, Tab.III.11, Tab.III.15.
- Afin d'assurer la convergence des algorithmes, nous avons refait les expériences pour toutes les images avec 100 itérations : aucune dégradation des segmentations n'est observée.
- La qualité complète est comparable pour le résultat de classification, mais le temps de calcul est la même. En particulier, la procédure d'échantillonneur de Gibbs se montre très coûteuse, puisqu'elle nécessite plus d'une 1heur pour finir un résultat de segmentation par les champs de Markov cachés.
- Dans le cas de cette segmentation, la valeur du paramètre de régularité a été estimée à $\beta = 0.8$.
- On constate que les deux méthodes (K-means et champs de Markov) ne capturent pas les mêmes détails de l'image réelle. Notamment d'un point de vue visuel, la segmentation par K-means (FiG.III.4. (c), FiG.III.5. (c), FiG.III.6. (c) et FiG.III.7. (c)) est clairement moins satisfaisante que la segmentation par champs de Markov.
- On remarque que le résultat de segmentation par les champs de Markov cachés d'images sans flou était plus satisfaisant que celui avec flou. Ainsi on peut dire que la méthode de segmentation fondue sur les champs de Markov cachées est adaptée aux images sans flou.

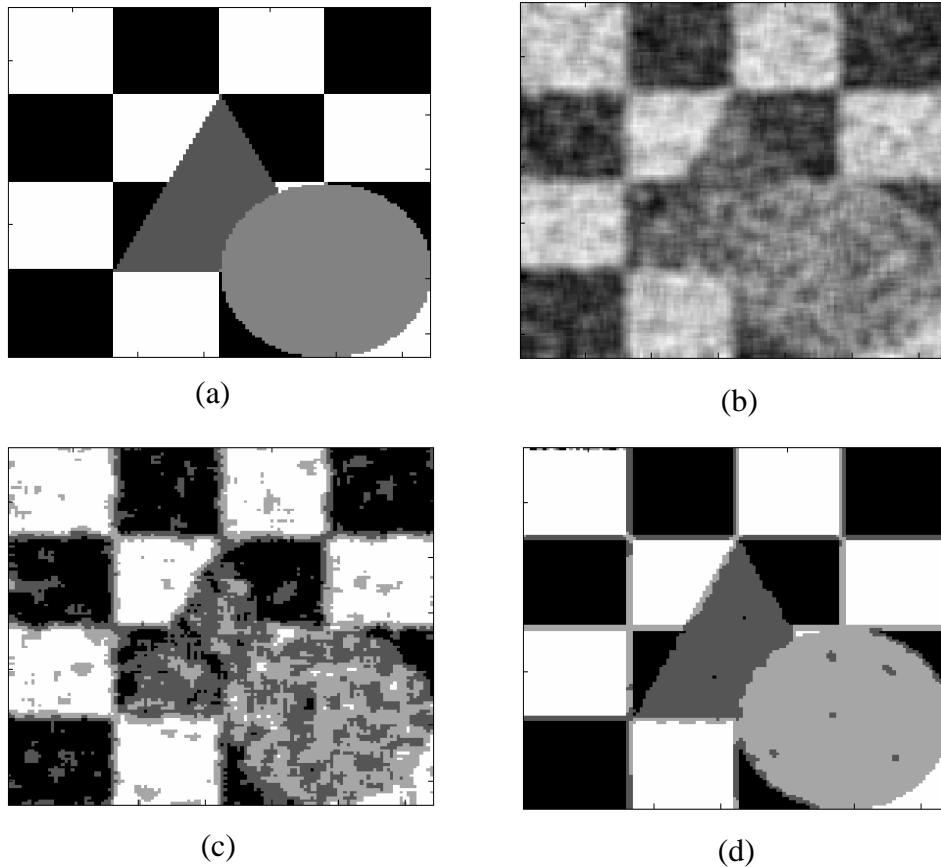


FIG.III.4 - Résultat de segmentation en niveau de gris de l'image synthétique (1) avec flou : (a) image originale, (b) image bruitée avec flou, (c) selon algorithme en Kmeans et (d) selon algorithme en champ de Markov caché.

0.9022	0.0652	0.0325	0
0.0151	0.9409	0.0440	0
0	0.0371	0.9629	0
0.0011	0.0365	0.0548	0.9076

Tab.III.2 - Matrice de confusion pour le résultat de la classification dans figure (d) de l'image synthétique (1) avec flou.

Précision globale	Précision moyenne	kAPPA	PSNR
91.9678	92.8425	0.8875	16.5565

Tab.III.3 - L'évaluation de résultat de segmentation de l'image synthétique (1) avec flou

Nombre of Pixels	Classification Data							
	R \ C	C1	C2	C3	C4	Row Total	Producer's Accuracy	Errors of Producer's
Reference Data	R1	5020	363	181	0	5564	5020/ 5564	544/ 5564
	R2	23	1433	67	0	1523	1433/ 1523	90/ 1523
	R3	0	119	3088	0	3207	3088/ 3207	119/ 3207
	R4	7	222	334	5527	6090	5527/ 6090	563/ 6090
	Column Total	5050	2137	3670	5527	16384		
	User's accuracy	5020/ 5050	1433/ 2137	3088/ 3670	5527/ 5527			
	Errors of User's	30/ 5050	704/ 2137	582/ 3670	0/ 5527			
	Tab.III.4 - Matrice de confusion correspondant à la classification faite par champ de Markov caché sur l'image synthétique (1) avec flou avec le nombre de Pixel							

$$\text{Précision globale} = \rho_{\text{globale}} = ((5020 + 1433 + 3088 + 5527) / 16384) * 100 = 91.9678 \%$$

Nombre of Pixels	Classification Data								
	Reference Data	R \ C	C1	C2	C3	C4	Row Total	Producer's Accuracy	Errors of Producer's
	R1	30.64	2.22	1.10	0	33.96	90.22	9.78	
	R2	0.14	8.74	0.41	0	9.29	94.10	5.9	
	R3	0	0.73	18.85	0	19.58	96.29	3.71	
	R4	0.04	1.35	2.04	33.74	37.17	90.76	9.24	
	Column Total	30.82	13.04	22.4	33.74	100			
	User's accuracy	99.41	67.06	84.14	100				
	Errors of User's	0.59	32.94	15.86	0				
Tab.III.5- Matrice de confusion du Tab-III.4 avec toutes les valeurs exprimées en pourcentage.									

Précision moyenne = $\rho_{moyenne} = ((90.22 + 94.10 + 96.29 + 90.76) / 4) = 92.8425\%$

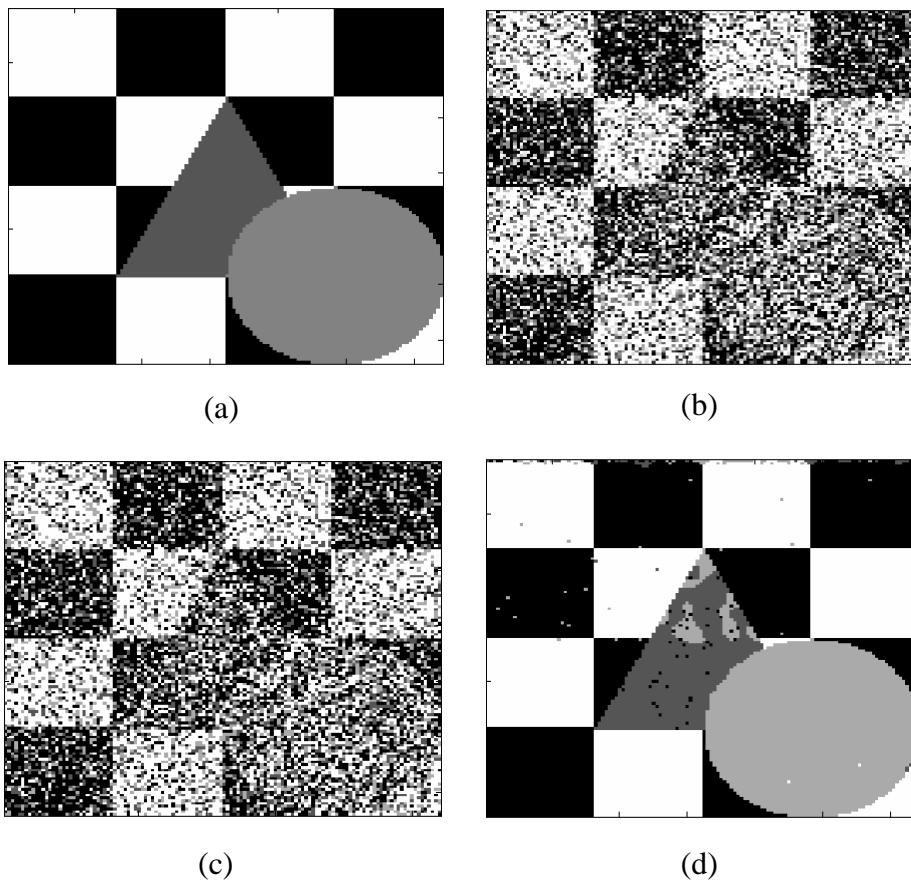


FIG.III.5 - Résultat de segmentation en niveau de gris de l'image synthétique (1) sans flou : (a) image originale, (b) image bruitée avec flou, (c) selon algorithme en Kmeans et (d) selon algorithme en champ de Markov caché.

0.9894	0.0058	0.0049	0
0.0295	0.8444	0.1261	0
0	0.0025	0.9969	0.0006
0.0013	0.0018	0.0069	0.9900

Tab.III.6- Matrice de confusion pour le résultat de la classification dans figure (d) de l'image synthétique (1) sans flou.

Précision globale	Précision moyenne	kAPPA	PSNR
97.7600	95.5175	0.9680	27.1147

Tab.IV.7 - L'évaluation de résultat de segmentation de l'image synthétique (1) sans flou

Nombre of Pixels	Classification Data							
	R \ C	C1	C2	C3	C4	Row Total	Producer's Accuracy	Errors of Omission
Reference Data	R1	5505	32	27	0	5564	5505/ 5564	59/ 5564
	R2	45	1286	192	0	1523	1286/ 1523	237/ 1523
	R3	0	8	3197	2	3207	3197/ 3207	10/ 3207
	R4	8	11	42	6029	6090	6029/ 6090	61/ 6090
	Column Total	5558	1337	3458	6031	16384		
	User's Accuracy	5505/ 5558	1286/ 1337	3197/ 3458	6029/ 6031			
	Errors of User's	53/ 5558	51/ 1337	261/ 3458	2/ 6031			
	Tab.III.8 - Matrice de confusion correspondant à la classification faite par champ de Markov caché sur l'image synthétique (1) sans flou avec le nombre de Pixel							

$$\text{Précision globale} = \rho_{\text{globale}} = ((5505 + 1286 + 3197 + 6029) / 16384) * 100 = 97.7600 \%$$

Nombre of Pixels	Classification Data								
	Reference Data	R \ C	C1	C2	C3	C4	Row Total	Producer's Accuracy	Errors of Producer's
	R1	33.60	0.20	0.16	0	33.96	98.94	1.06	
	R2	0.27	7.85	1.17	0	9.29	84.44	15.56	
	R3	0	0.05	19.53	0	19.58	99.69	0.31	
	R4	0.05	0.07	0.26	36.79	37.17	99.00	1.00	
	Column Total	33.92	8.16	21.11	36.81	100			
	User's Accuracy	99.05	96.18	92.45	99.97				
	Errors of User's	0.95	3.82	7.55	0.03				
Tab.III.9- Matrice de confusion du Tab-III.8 avec toutes les valeurs exprimées en pourcentage.									

Précision moyenne = $\rho_{moyenne} = ((98.94 + 84.44 + 99.69 + 99.00) / 4) = 95.5175\%$

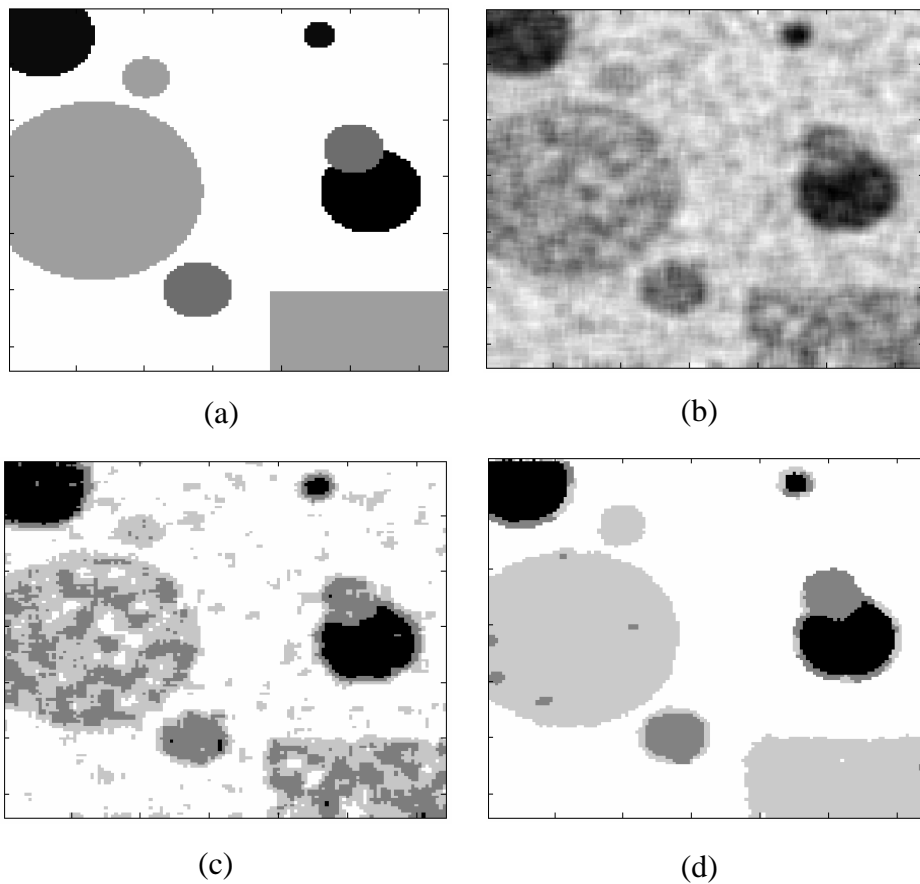


FIG.III.6 - Résultat de segmentation en niveau de gris de l'image synthétique (2) avec flou : (a) image originale, (b) image bruitée avec flou, (c) selon algorithme en Kmeans et (d) selon algorithme en champ de Markov caché.

0.8311	0.1046	0.0585	0.0058
0.0217	0.8463	0.1266	0.0054
0	0.0124	0.9697	0.0179
0	0.0087	0.0222	0.9691

Tab.III.10- Matrice de confusion pour le résultat de la classification dans figure (d) de l'image synthétique (2) avec flou.

Précision globale	Précision moyenne	kAPPA	PSNR
95.4895	90.405	0.9172	17.7763

Tab.IV.11 - L'évaluation de résultat de segmentation de l'image synthétique (2) avec flou

Nombre of Pixels	Classification Data								
	R \ C	C1	C2	C3	C4	Row Total	Producer's Accuracv	Errors of Producer's	
Reference Data	R1	1009	127	71	7	1214	1009 / 1214	205/ 1214	
	R2	12	468	70	3	553	468/ 553	85/ 553	
	R3	0	57	4444	82	4583	4444/ 4583	139/ 4583	
	R4	0	87	223	9724	10034	9724/ 10034	310/ 10034	
	Column Total	1021	739	4808	9816	16384			
	User's Accuracy	1009/ 1021	468/ 739	4444/ 4808	9724/ 9816				
	Errors of User's	12/ 1021	271/ 739	364/ 4808	92/ 9816				
	Tab.III.12 - Matrice de confusion correspondant à la classification faite par champ de Markov caché sur l'image synthétique (2) avec flou avec le nombre de Pixel								

$$\text{Précision globale} = \rho_{\text{globale}} = ((1009 + 468 + 4444 + 9724) / 16384) * 100 = 95,4895 \%$$

Nombre of Pixels	Classification Data							
Reference Data	R \ C	C1	C2	C3	C4	Row Total	Producer's Accuracy	Errors of Producer's
	R1	6.16	0.78	0.43	0.04	7.40	83.11	16.89
	R2	0.07	2.86	0.43	0.02	3.38	84.63	15.37
	R3	0	0.35	27.12	0.50	27.97	96.97	3.03
	R4	0	0.53	1.36	59.35	61.25	96.91	3.09
	Column Total	6.23	4.52	29.34	59.91	100		
	User's Accuracy	98.82	63.33	92.43	99.06			
	Errors of User's	1.18	36.37	7.57	0.94			
	Tab.III.13- Matrice de confusion du Tab-III.12 avec toutes les valeurs exprimées en pourcentage.							

Précision moyenne = $\rho_{moyenne} = ((83.11 + 84.63 + 96.97 + 96.91) / 4) = 90.405\%$

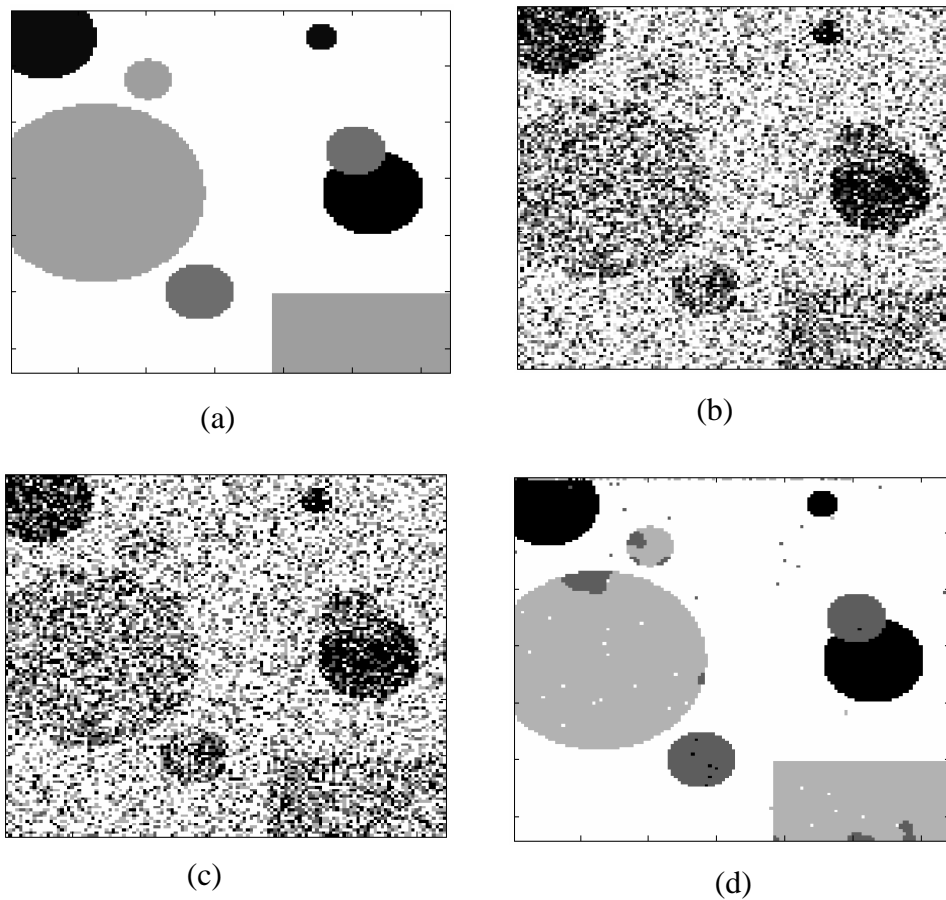


FIG.III.7 - Résultat de segmentation en niveau de gris de l'image synthétique (2) sans flou : (a) image originale, (b) image bruitée sans flou, (c) selon algorithme en Kmeans et (d) selon algorithme en champ de Markov caché.

0.9860	0.0016	0.0091	0.0033
0.0127	0.9873	0	0
0	0.0377	0.9572	0.0050
0	0.0018	0.0037	0.9945

Tab.III.14- Matrice de confusion pour le résultat de la classification dans figure (d) de l'image synthétique (2) sans flou.

Précision globale	Précision moyenne	kAPPA	PSNR
98.3215	98.1275	0.9691	30.4994

Tab.IV.15 - L'évaluation de résultat de segmentation de l'image synthétique (2) sans flou.

Nombre of Pixels	Classification Data							
	R \ C	C1	C2	C3	C4	Row Total	Producer's Accuracy	Errors of Producer's
Reference Data	R1	1197	2	11	4	1214	1197/1214	17/1214
	R2	7	546	0	0	553	546/553	7/553
	R3	0	173	4387	23	4583	4387/4583	196/4583
	R4	0	18	37	9979	10034	9979/10034	55/10034
	Column Total	1204	739	4435	10006	16384		
	User's accuracy	1197/1204	546/739	4387/4435	9979/10006			
	Errors of User's	7/1204	193/739	48/4435	27/10006			
	<p>Tab.III.16 - Matrice de confusion correspondant à la classification faite par champ de Markov caché sur l'image synthétique (2) sans flou avec le nombre de Pixel</p>							

$$\text{Précision globale} = \rho_{\text{globale}} = ((1197 + 546 + 4387 + 9979) / 16384) * 100 = 98.3215\%$$

Nombre of Pixels	Classification Data							
	R \ C	C1	C2	C3	C4	Row Total	Producer's Accuracy	Errors of Producer's
Reference Data	R1	7.31	0.01	0.07	0.02	7.41	98.61	1.39
	R2	0.04	3.33	0	0	3.38	98.73	1.27
	R3	0	1.06	26.78	0.14	27.97	95.72	4.28
	R4	0	0.11	0.23	60.91	61.24	99.45	0.55
	Column Total	7.35	4.51	27.08	61.07	100		
	User's accuracy	99.42	73.88	98.92	99.73			
	Errors of User's	0.58	26.12	1.08	0.27			
	Tab.III.17- Matrice de confusion du Tab-III.16 avec toutes les valeurs exprimées en pourcentage.							

$$\text{Précision moyenne} = \rho_{\text{moyenne}} = ((98.61 + 98.73 + 95.72 + 99.45) / 4) = 98.1275\%$$

III.6. Conclusion

Nous avons proposé dans ce chapitre un algorithme d'estimation des paramètres a priori et conditionnels, ICE, dans un cadre de mélanges de lois gaussiennes, d'un champ de Markov. La méthode de ICE donne de meilleurs résultats d'estimations, surtout lorsqu'il s'agit d'un bruitage important. Aussi l'algorithme ICE est très rapide. Nous avons également présenté l'algorithme MPM qui a été appliqué au problème de segmentation par des champs de Markov cachés. Nos différents algorithmes seront testés et évalués sur des problèmes de classification non supervisée d'images synthétiques. Ces algorithmes formeront un système de segmentation qui sera par la suite implémenté. Cette implémentation mettra en évidence l'intérêt et la performance de champ de Markov au niveau de la segmentation des images mammographies.

SIMULATION ET DISCUSSION DES RESULTATS

Dans les chapitres précédents, le modèle de champ de Markov caché et la méthode d'estimation des paramètres ont été présentés. Nous consacrons ce chapitre à l'application de cet modélisation et technique d'estimation des paramètres à la segmentation statique non supervisée d'images mammographies.

IV.1. Introduction

Nous proposons ici quelques résultats de simulation en intégrant une information sur la qualité de chaque résultat de segmentation pour mettre en évidence la performance de champs de Markov caché. Nous détaillons dans un premier temps le principe général des critères d'évaluation d'un résultat de segmentation défini [ZER 2003]. Nous présentons par la suite quelques résultats expérimentaux sur d'images mammographies et nous concluons.

IV.2. Critère d'évaluation d'un résultat de segmentation

Dans le cas général, aucune vérité terrain n'est disponible pour évaluer un résultat de segmentation. Nous proposons ici une méthode d'évaluation de la qualité d'un résultat de segmentation ne nécessitant aucune connaissance a priori sur les images à segmenter. On va présenter un critère utilisé généralement pour l'évaluation d'un résultat de segmentation dans un contexte non supervisé. La valeur de ce critère augmente avec la qualité du résultat de segmentation.

IV.2. 1. Définition du critère

L'objectif de l'évaluation est, étant donné une image et une méthode de segmentation, de déterminer la performance de cette méthode au sens d'un critère défini. Le principe utilisé consiste à estimer l'homogénéité des régions de l'image segmentée. Dans l'élaboration d'un critère d'évaluation d'un résultat de segmentation, on doit tenir compte des deux propriétés suivantes [MDP 1994] [Met 1983] :

- Une région de l'image segmentée ne doit contenir qu'une seule primitive (une texture ou un niveau de gris constant), ceci pour garantir qu'il n'y ait pas sous segmentation. Ainsi, une région est caractérisée par une des statistiques en son sein.
- Deux régions limitrophes doivent contenir deux primitives différentes pour garantir qu'il n'y ait pas sur segmentation. Ceci correspond à une disparité des statistiques entre ces deux régions.

En tenant compte de ces deux points, Nazif et Levine [LDZ1998] considèrent qu'une segmentation adéquate doit produire des images ayant une valeur d'uniformité intraclasse plus élevée, ceci est relié avec les propriétés statistiques (la variance) d'une classe.

L'uniformité d'une classe peut être calculée sur la base de la somme des variances des classes. Soient $I(x, y)$ une image en niveau de gris, R_i la $i^{\text{ème}}$ classe segmentée et A_i la surface de R_i , alors la mesure d'uniformité du niveau gris (GU) de l'image $I(x, y)$ est donnée par la formule suivante :

$$GU = \sum_i \sum_{(x,y) \in R_i} \left[I(x,y) - \frac{1}{A_i} \sum_{(x,y) \in R_i} I(x,y) \right]^2 \dots\dots\dots (IV.1)$$

IV.2.2. Valeur de l'uniformité intraclasse normalisée

Sahoo a proposé une mesure d'uniformité intraclasse normalisée (NU). Ce critère calcule la somme des variances normalisées des régions retranchée à 1. Donc la valeur du critère est donnée par la formule suivante [Fjr 2000] :

$$NU = 1 - GU / C \dots\dots\dots (IV.2)$$

Où

C est un facteur de normalisation. Généralement, d'autres caractéristiques peuvent être aussi utilisées.

La valeur de NU quantifie l'homogénéité globale du résultat de l'image segmentée. Une valeur de ce critère proche de 1 traduit un très bon résultat de segmentation.

IV.3. Résultats expérimentaux

Nous allons expérimenter la méthode de segmentation basée sur le champ de Markov caché, que nous avons étudiés aux chapitres précédents, sur une série d'images mammographies réelles. Les caractéristiques de ces images sont résumés dans le **Tab.IV.1**.

En effet, ne disposant pas d'une vérité terrain pour les images mammographies réelles à notre disposition, nous nous contenterons de montrer, à titre purement indicatif, les segmentations obtenues par le modèle étudié. En plus, nous allons exploiter la valeur de l'uniformité intraclasse normalisée (NU) afin d'évaluer les résultats de segmentation.

Figure	Taille	Type	Commentaires
FIG.IV.1	512×512	Mammographie	cas sain
FIG.IV.2	512×512	Mammographie (1)	présence d'anomalie
FIG.IV.3	512×512	Mammographie (2)	présence d'anomalie
FIG.IV.4	512×512	Mammographie (3)	présence d'anomalie

Tab.IV.1 - *Caractéristiques des images utilisées.*

IV.3.1. Présentation des résultats

Les exemples traités sont des images en niveau de gris de même taille. Cette série d'expérimentations concerne des images mammographies réelles, pour lesquelles il existe de valeur exacte du nombre de classes. Cependant il est possible d'obtenir un avis d'expert sur un nombre de classes approprié en fonction du contexte ou bien l'utilisation de l'histogramme.

L'objectif de ces exemples étant d'obtenir des segmentations en composantes homogènes facilement exploitable et d'attirer l'attention sur la qualité de la segmentation des images mammographies par les champs de Markov cachés.

Les segmentations de ces images obtenues à partir des champs de Markov cachés sont reportées sur les figures : FIG.IV.1. (c), FIG.IV.3. (c), FIG.IV.5. (c) et FIG.IV.7. (c). Nous avons utilisé comme modèle a priori un modèle de Potts simple avec un voisinage de huit connexions et un modèle gaussien pour le terme d'attache aux données de moyennes et de variances distinctes.

Le nombre d'itérations de l'algorithme d'ICE est mis à 30 itérations, avec seulement une réalisation a posteriori par itération. La valeur initiale du paramètre de régularité du champ de Markov caché β est 0.5. Dans chaque itération d'ICE, le nombre maximal d'itérations pour le gradient stochastique est mis à 10, avec une réalisation a priori par itération. L'évaluation de gradient stochastique exige généralement très peu d'itérations. L'échantillonneur de Gibbs qu'on utilise a besoin de 100 itérations pour générer les réalisations a priori et a posteriori. La

convergence des énergies globales est en fait tout à fait lente, particulièrement pour des réalisations selon la distribution *a priori*. La classification MPM compte sur 20 réalisations *a posteriori*.

Dans le but de réduire le temps de calcul, on a simulé seulement une réalisation *a posteriori* pour chaque itération d'ICE de la segmentation et seulement une réalisation *a priori* pour chaque itération du gradient stochastique pour l'estimation du paramètre de régularité β de la segmentation par champ de Markov caché. Cette simplification ne cause aucune perte de performance significative.

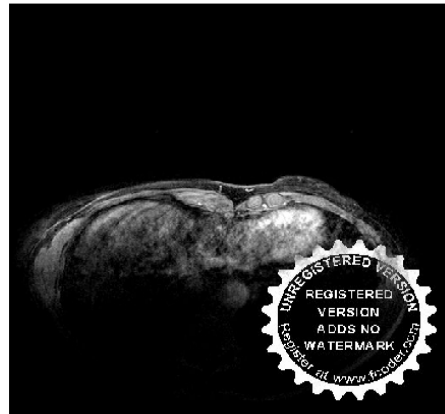
Les tableaux de Tab.IV.2, Tab.IV.4, Tab.IV.6, Tab.IV.8 récapitulent les valeurs de l'uniformité intraclasse normalisée (NU) pour chaque résultat de segmentation. En effet ce critère permet d'estimer l'homogénéité des régions de l'image segmentée.

IV.3.2. Analyse et discussion

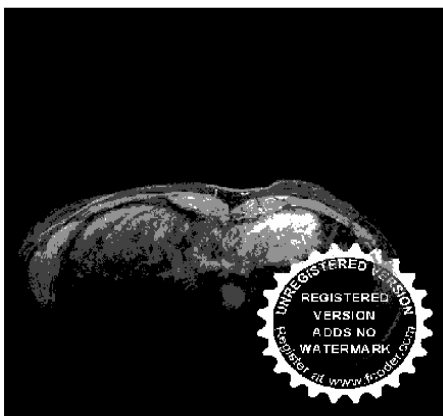
En générale, les résultats conquête sont convenables qui présentent la performance en terme d'estimation des paramètres de modèle et en terme de segmentation. Si l'on analyse les variations des segmentations au point de vue visuel et au sens du critère d'évaluation, on remarque qu'elles sont cohérentes. Mais on peut tirer quelques remarques :

- La première constatation est que les deux méthodes (K-means et champs de Markov) ne capturent pas les mêmes détails de l'image réelle. Notamment d'un point de vue visuel, la segmentation par K-means (FiG.IV.1 (b), FiG.IV.2 (b), FiG.IV.3 (b) et FiG.IV.4 (b)) est clairement moins satisfaisante que la segmentation par champs de Markov.
- Afin d'assurer la convergence des algorithmes, nous avons refait les expériences pour toutes les images avec 100 itérations : aucune dégradation des segmentations n'est observée.
- La qualité complète est comparable pour le résultat de classification, mais le temps de calcul est tout à fait différent. En particulier, la procédure d'échantillonneur de Gibbs se montre très coûteuse, puisqu'elle nécessite plus d'une 1heure et 30 minute pour finir un résultat de segmentation par les champs de Markov cachés.

- La distribution gaussienne est bien adaptée aux images mammographiques.
- Dans le cas de segmentation par les champs de Markov cachés, la valeur du paramètre de régularité a été estimée à $\beta = 0.8$.



(a)



(b)



(c)

FIG.IV.1 - Résultat de segmentation en niveau de gris de l'image mammographie sain : (a) image originale, (b) selon algorithme en K-means, (c) selon algorithme en champ de Markov caché.

Nombre Classe	Nombre Itération Gibbs	Nombre Itération ICE	Nombre Itération MPM
4	5	30	5

Tab.IV.2 – Nombre d'itération de chaque algorithme

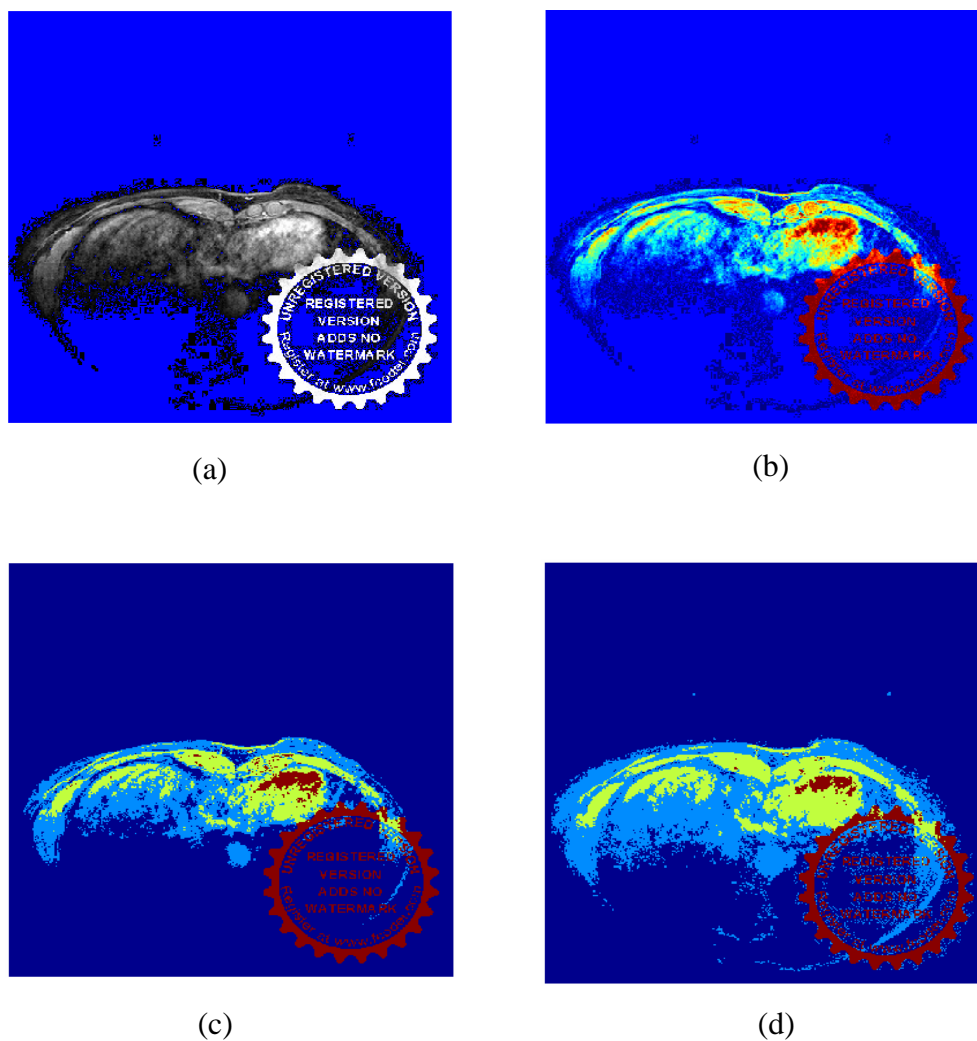
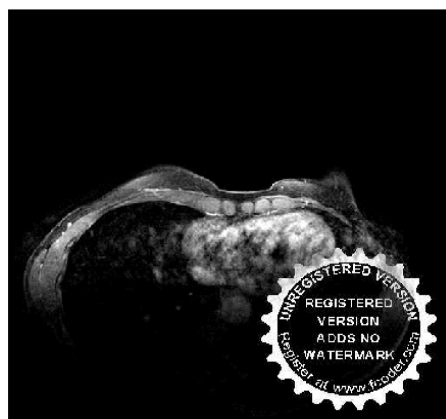


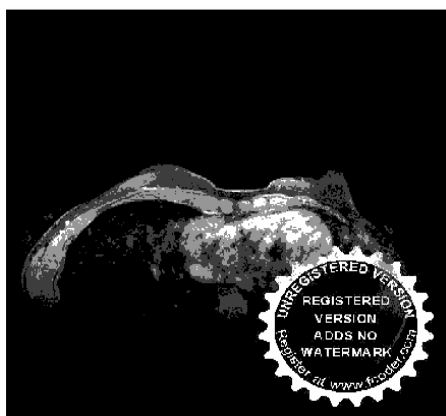
FIG.IV.2 - Résultat de segmentation en couleur de l'image mammographie sain : (a,b) image originale, (c) selon algorithme en K-means, (d) selon algorithme en champ de Markov caché.

NU	Temps
0.925	5 h20min

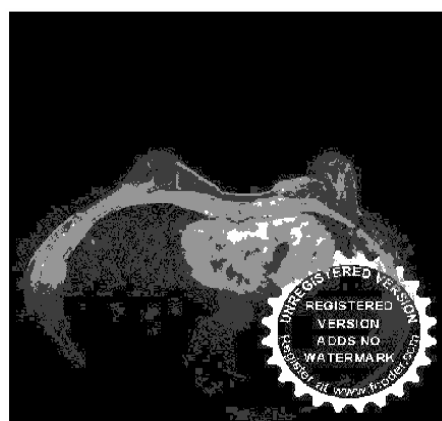
Tab.IV.3 - L'évaluation de résultat de segmentation de l'image mammographie sain



(a)



(b)



(c)

FIG.IV.3 - Résultat de segmentation en niveau de gris de l'image mammographie (2) 'présence d'anomalie' : (a) image originale, (b) selon algorithme en K-means, (c) selon algorithme en champ de Markov caché.

Nombre Classe	Nombre Itération Gibbs	Nombre Itération ICE	Nombre Itération MPM
4	6	30	5

Tab.IV.4 – Nombre d'itération de chaque algorithme

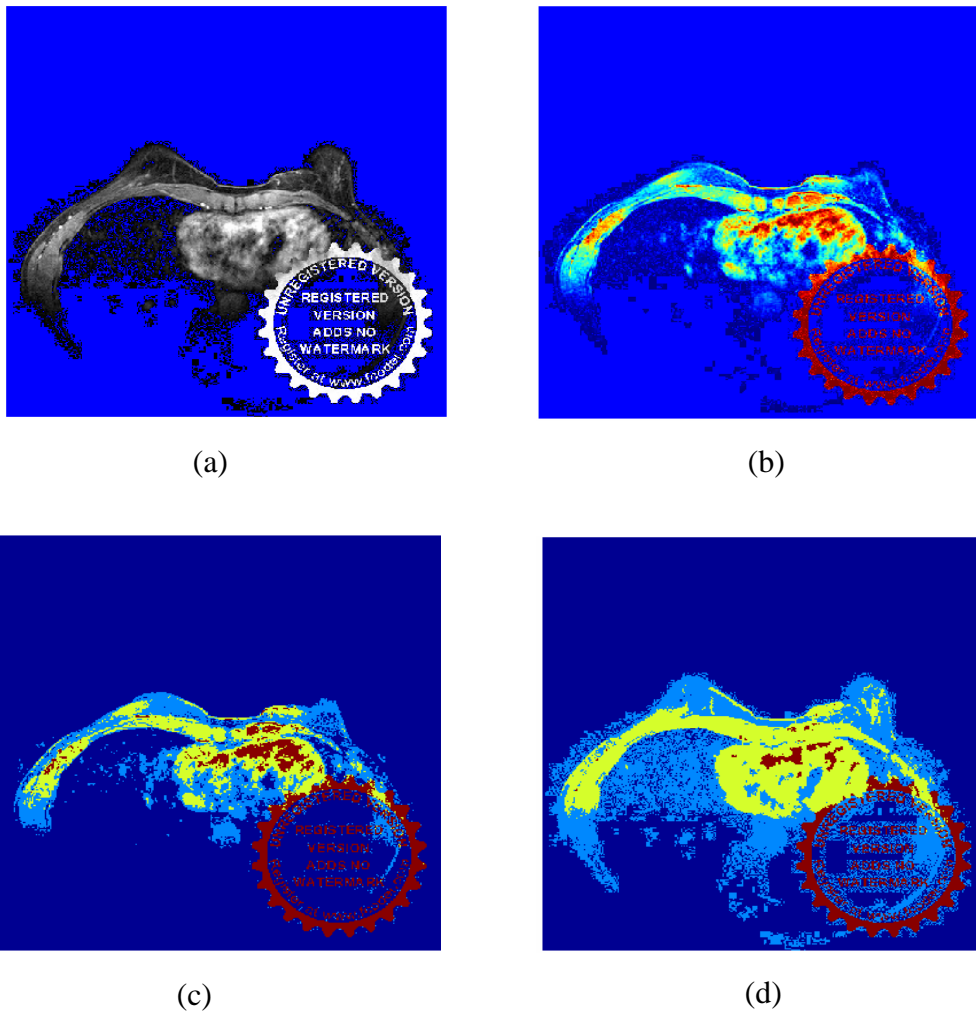
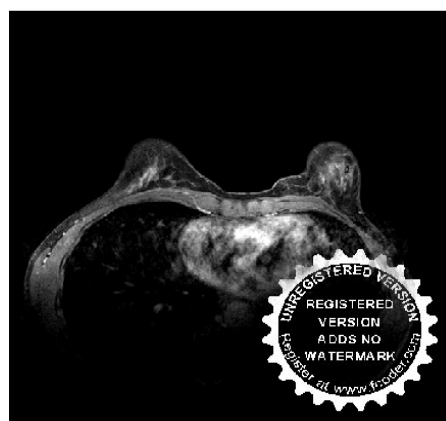


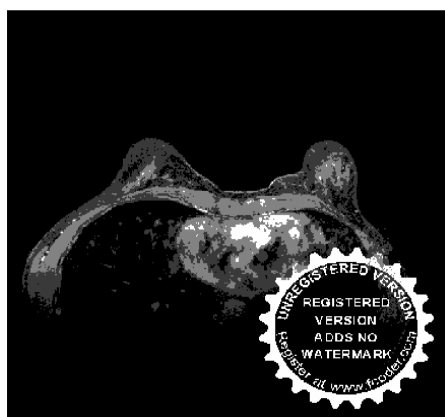
FIG.IV.4 - Résultat de segmentation en couleur de l'image mammographie (2) 'présence d'anomalie' : (a, b) image originale, (c) selon algorithme en K-means, (d) selon algorithme en champ de Markov caché.

NU	Temps
0.943	5h50min

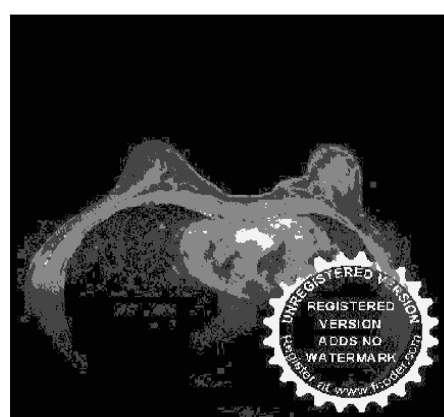
Tab.IV.5 - L'évaluation de résultat de segmentation de l'image mammographie (2)



(a)



(b)



(c)

FIG.IV.5 - Résultat de segmentation en niveau de gris de l'image mammographie (3) 'présence d'anomalie' : (a) image originale, (b) selon algorithme en K-means, (c) selon algorithme en champ de Markov caché.

Nombre Classe	Nombre Itération Gibbs	Nombre Itération ICE	Nombre Itération MPM
4	5	30	5

Tab.IV.6 – Nombre d'itération de chaque algorithme

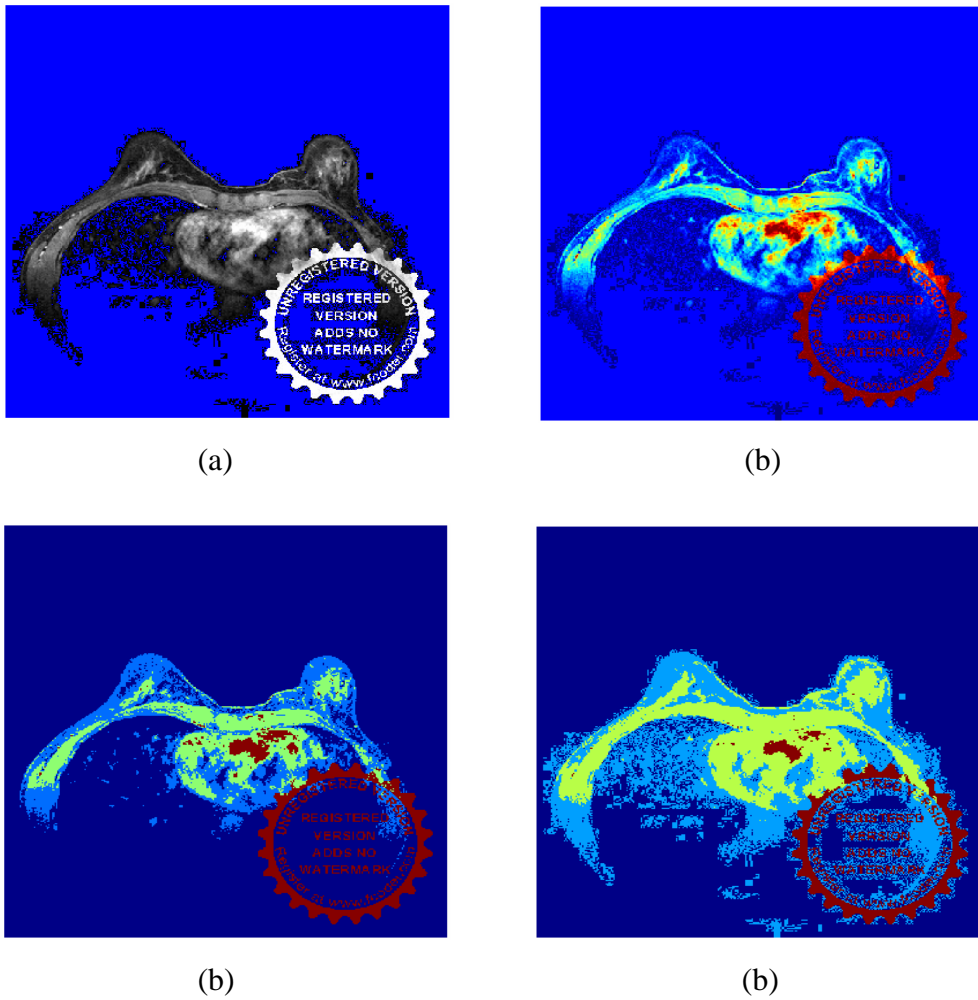
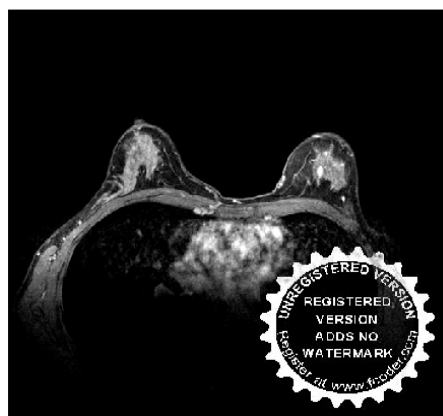


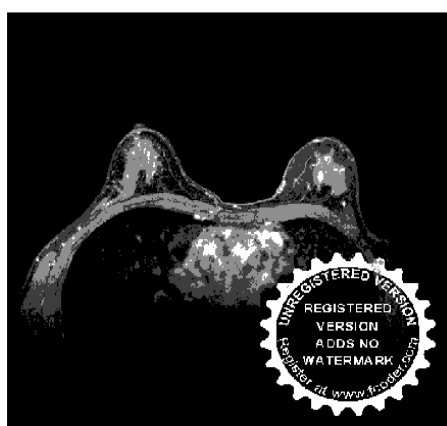
FIG.IV.6 - Résultat de segmentation en couleur de l'image mammographie (3) 'présence d'anomalie' : (a,b) image originale, (c) selon algorithme en K-means, (d) selon algorithme en champ de Markov caché.

NU	Temps
0.920	5h20min

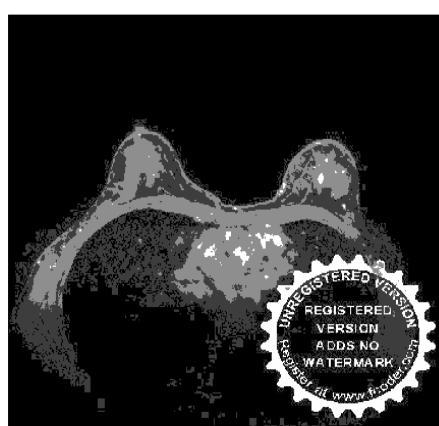
Tab.IV.7- L'évaluation de résultat de segmentation de l'image mammographie (3)



(a)



(b)



(c)

FIG.IV.7 - Résultat de segmentation en niveau de gris de l'image mammographique (4) 'présence d'anomalie' : (a) image originale, (b) selon algorithme en K-means, (c) selon algorithme en champ de Markov caché.

Nombre Classe	Nombre Itération Gibbs	Nombre Itération ICE	Nombre Itération MPM
4	5	30	20

Tab.IV.8 – Nombre d'itération de chaque algorithme

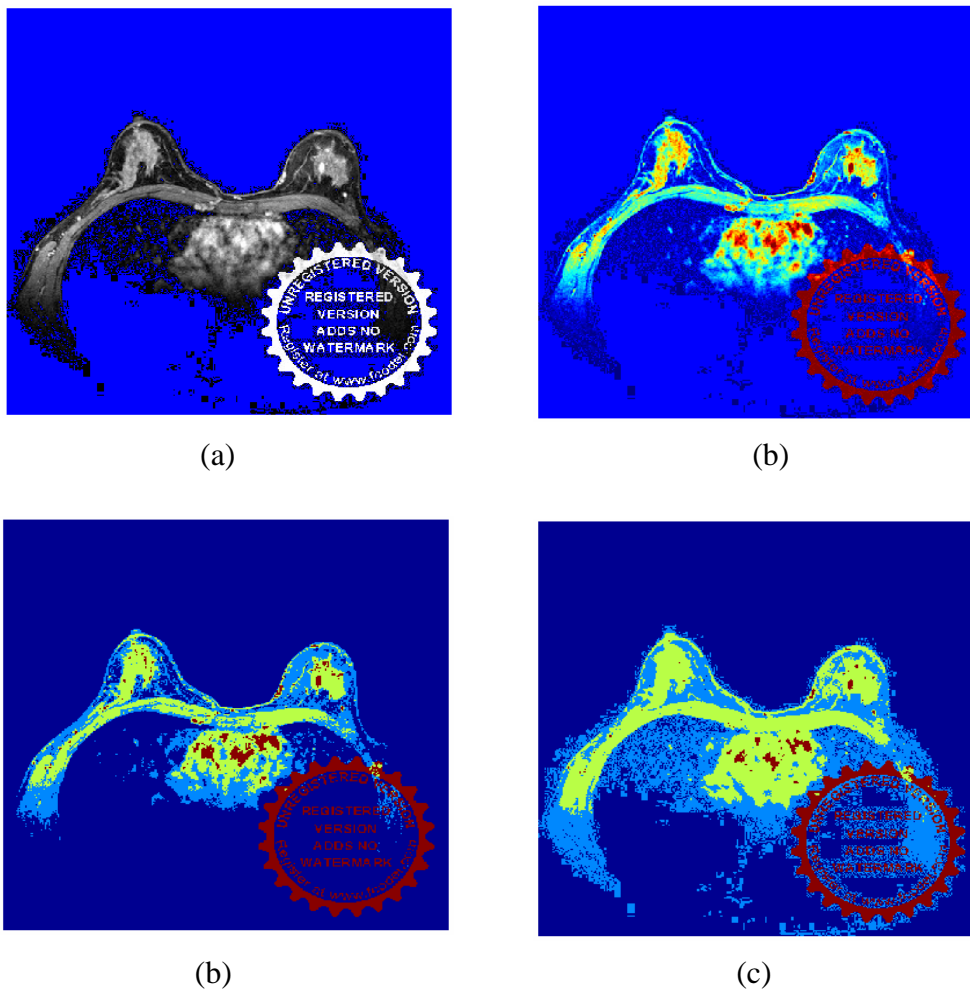


FIG.IV.8 - Résultat de segmentation en couleur de l'image mammographie (4) 'présence d'anomalie' : (a,b) image originale, (c) selon algorithme en K-means, (d) selon algorithme en champ de Markov caché.

NU	Temps
0.937	5h20min

Tab.IV.9 - L'évaluation de résultat de segmentation de l'image mammographie (3)

IV.4. Conclusion

Ce travail décrit la classification non supervisée des images mammographies dans le cadre de champ de Markov caché et l'estimation gaussien. Les champs de Markov cachés sont fréquemment employés pour imposer des contraintes de régularité spatiales dans l'estimation de paramètre et les étapes de classification. Cette approche produit des résultats excellents dans de nombreux cas, mais nos expériences indiquent que l'estimation du paramètre de régularité β est un problème délicat, particulièrement quand le taux du signal du bruit est bas. Le temps considérable de calcul est un autre inconvénient. Les frontières des régions des structures excellentes sont généralement mieux préservées. En plus de l'estimation robuste des paramètres de régularité.

Plusieurs faits ressortent de cette étude expérimentale. D'abord, l'intérêt de la prise en compte de l'information spatiale pour la classification des pixels d'une image apparaît de manière évidente dans les différents exemples traités. Comme il fallait s'y attendre, le choix d'un modèle gaussien conduit à des segmentations acceptables et l'on observe une nette amélioration au niveau de l'homogénéité de la segmentation avec un modèle spatial. Ce résultat est satisfaisant car cela confirme, que les approximations d'un champ de Markov caché utilisé préservent l'information markovienne.

CHAPITRE 5
CONCLUSION GÉNÉRALE

V.1. Récapitulatif

Dans cette thèse nous avons travaillé sur les axes suivants :

- Étude des images.
- Pour faire cette étude nous avons traité des images mammographies.
- Les modèles stochastiques que sont les champs de Markov s'avèrent être de bons détecteurs des régions (la région cancéreuse). C'est pour cela qu'ils ont été largement utilisés à cet effet comme pour la classification et la segmentation. Parmi les méthodes de segmentation contextuelles, nous nous sommes intéressés à la segmentation non supervisée d'image modélisées par les champs de Markov cachés. L'originalité de ce modèle markovien caché réside dans l'introduction des informations spatiales. Cette introduction s'est effectuée tout d'abord avec la notion de voisinage.

Pour aboutir à cela nous avons fait dans :

- La première partie a été consacrée à une présentation des images mammographies et aux différents modèles de segmentation. L'étude de ces modèles montre que la théorie la plus utilisée pour modéliser l'information extraite des images est sans conteste la théorie des probabilités basée sur les champs de Markov cachés. Ils introduisent des informations contextuelles avec une régularisation spatiale. Par contre, les autres méthodes de segmentation d'images, qu'on a vu, comportent certaines limites. Ces méthodes ne respectent pas la cohérence spatiale des classes dans l'image.
- La deuxième partie a traité la modélisation markovienne d'images. Elle permet la construction des champs de Markov cachés. Pour caractériser nos régions, nous avons utilisé trois attributs locaux (moyenne, variance et le paramètre de régularité β) . Après avoir précisé les différentes étapes de modélisation, nous avons proposé l'algorithme pour la réalisation de champ de Markov caché.
- La troisième partie s'est intéressée à l'estimation des paramètres d'attaches aux données et des probabilités conditionnelles a priori en vue d'une segmentation non supervisée. En premier lieu, pour l'estimation des paramètres, nous avons utilisé l'algorithme ICE qui présente l'avantage qu'il est peu coûteux en temps de calcul.

Ensuite pour réaliser la tâche de segmentation, on a opté pour une méthode de classification non supervisée qui minimise le nombre de pixels mal classés basé sur l'estimateur MPM. Dans cette partie, nos différents algorithmes seront testés et évalués sur des problèmes de classification non supervisée d'images synthétiques.

- La quatrième et dernière partie a été pour nous l'occasion de mettre à l'œuvre le champ caché utilisé dans les parties précédentes. On a implémenté cette segmentation markovienne non supervisée sur des images mammographies. En menant cette implémentation, nous sommes confrontés au problème de la croissance du temps de calcul des champs de Markov cachés. Ce problème s'accroît lorsque la taille des images devient très importante. La méthode champ devient meilleure lorsque le nombre des itérations de l'échantillonneur de Gibbs est faible. Compte tenu de cette étude, nous avons retenu que la puissance de champ de Markov cache donne des résultats satisfaisants dans les images mammographies au niveau de la prise de décision

Plusieurs améliorations pour la segmentation markovienne non supervisée sur des images mammographies sont néanmoins envisageables.

V.2. Perspectives

Ce manuscrit nous a ouvert la porte sur les méthodes stochastiques en analyse d'images tel que la segmentation et la classification ; nous nous sommes intéressés à la segmentation non supervisée d'image mammographie modélisées par les champs de Markov cachés. Mais pour la détection de l'anomalie il est souhaitable d'utiliser les chaînes de Markov cachés. Synoptique de la détection des anomalies est donnée comme suit :

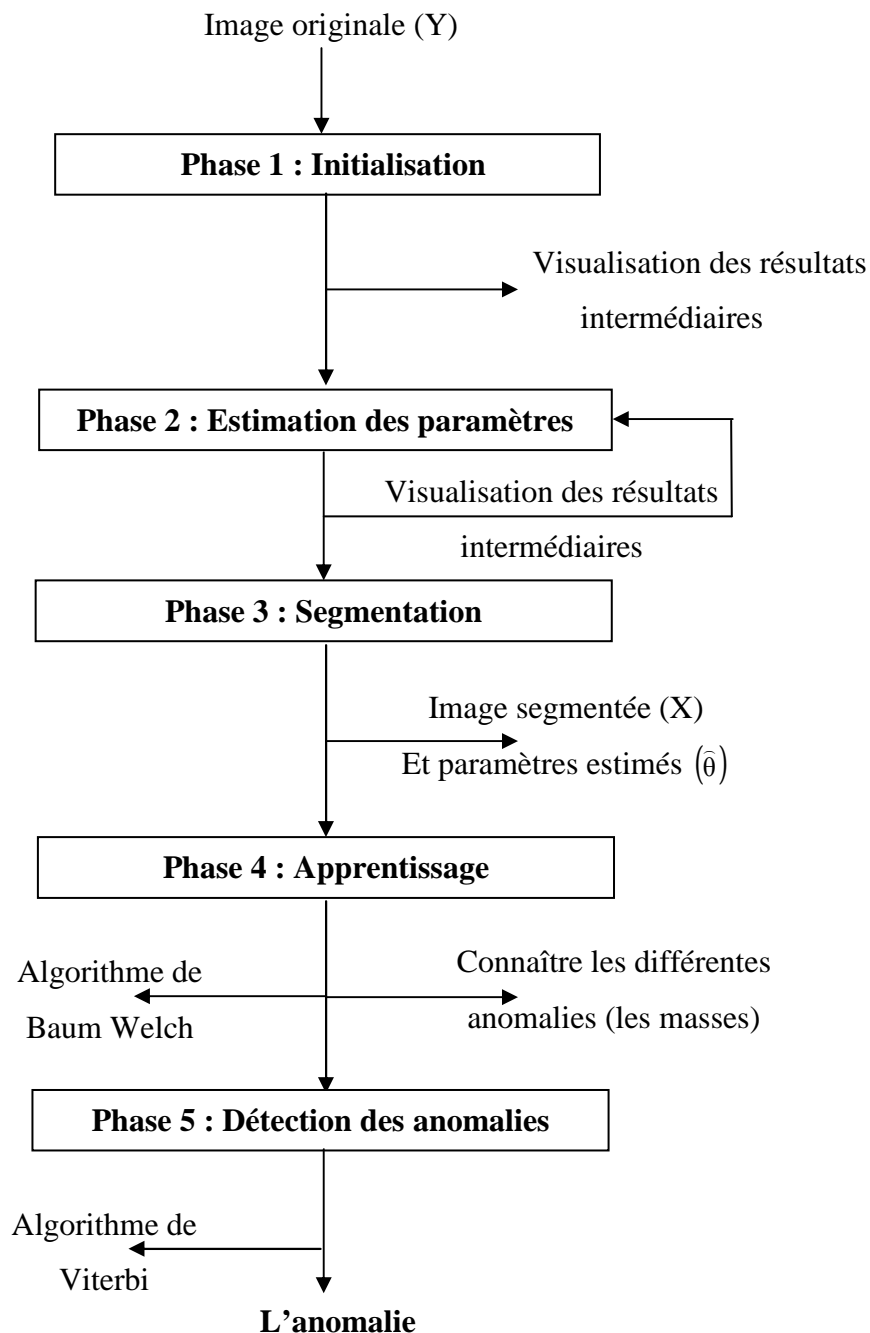


FIG.V.1 : Organigramme de la procédure de la détection des anomalies

Les perspectives de cette thèse sont nombreuses :

- Une recherche doit être faite sur l'étude du rôle effectif du paramètre spatiale de régularité β dans le modèle de Potts pour les champs de Markov cachés.
- Pour rendre cette segmentation plus robuste, il est souhaitable de fusionner plusieurs méthodes de segmentation, afin de réaliser un système de segmentation adaptatif pour les images mammographies.
- Une segmentation en 3D des zones fonctionnelles.

ANNEXE

**LES DIFFERENTS ALGORITHMES DE SEGMENTATION
NON SUPERVISEES**

Algorithme des K-means

arguments : Taille K (nombre de classes à déterminer).

initialisation : $C_0 = \{y_i, i = 1, 2, \dots, K\}$

tant que $C_m \neq C_{m-1}$ **faire**

/*Affectation des éléments de la séquence d'apprentissage */

pour tout vecteur z de la séquence d'apprentissage **faire**

si $\forall l \in \{1, 2, \dots, K\}, |z - y_l| \leq |z - y_1|$ **alors**
 $z \in R_l$;
 fin si

fin pour

/* Mise à jour des centres de classes */

pour $i = 1, \dots, K$ **faire**

$y_i = \text{centroïde}(\{z / z \in R_i\})$;

fin pour

fin tant que

Algorithme d'initialisation par K-means.

Echantillonneur de Gibbs

Soit x^0 la configuration initiale

Répéter indéfiniment :

Soit x^q la configuration courante

Pour tous les sites t de S :

- Poser $\tilde{x}_s = x_s^q, \forall s \neq t$ et $x_s^{q+1} = x_s^q, \forall s \neq t$
- Pour chaque ω_I de Ω :
 - Poser $\tilde{x}_s = \omega_I$
 - Calculer $P(\omega_I) = P\left(X_s = \tilde{x}_s \mid (X_t = x_t)_{t \in V_s}\right) = \frac{1}{Z_s} \exp(-U_s(\tilde{x}_s, (x_t)_{t \in V_s}))$
- Tirer u selon loi uniforme dans $[0,1]$
- Trouver le plus petit j tel que $\sum_{i=1}^j P(\omega_i) > u \cdot \sum_{i=1}^{|\Omega|} P(\omega_i)$
- Poser $x_s^{q+1} = j$

Algorithme d'échantillonneur de Gibbs.

Algorithme ICE

Initialisation des paramètres d'a priori $\lambda^{q,0}$ et des classes $f_i^0 (\mu_i^0, \sigma_i^0)$.

Répéter jusqu'à convergence

- Tirage des échantillons $x^q(1), \dots, x^q(t), \dots, x^q(m)$ selon

$$P(X_s = x_s | Y_s = y_s, (X_t = x_t)_{t \in V_s})$$

- Estimation des paramètres de classe pour chaque échantillon :

$$\hat{\mu}_i^{q,t} = \frac{\sum_{n=1}^N y_n \delta(x_n - \omega_i)}{\sum_{n=1}^N \delta(x_n - \omega_i)}.$$

$$(\hat{\sigma}_i^{q,t})^2 = \frac{\sum_{n=1}^N (y_n - \mu_i^q)^2 \delta(x_n - \omega_i)}{\sum_{n=1}^N \delta(x_n - \omega_i)}.$$

- Tirage des échantillons $x^q(1), \dots, x^q(r), \dots, x^q(m)$ selon

$$P(X_s = x_s | (X_t = x_t)_{t \in V_s})$$

- Estimation des paramètres de régularité pour chaque échantillon :

$$\hat{\lambda}^{q,r} = \lambda^{q,r-1} + \frac{1}{r} \frac{U_{\text{apriori}}^{q,r} - U_{\text{aposteriori}}^q}{U_{\text{aposteriori}}^q}$$

- Remise à jour des paramètres :

$$\mu_i^q = \frac{1}{m} \sum_{t=1}^m \hat{\mu}_i^{q,t}$$

$$(\sigma_i^q)^2 = \frac{1}{m} \sum_{t=1}^m (\hat{\sigma}_i^{q,t})^2.$$

$$\lambda^q = \frac{1}{m} \sum_{r=1}^m \hat{\lambda}^{q,r}.$$

Algorithme d'ICE sur un champ de Markov caché.

Algorithme MPM

- ▶ à partir d'une configuration initiale aléatoire, générer une suite de m réalisations $(x(1), \dots, x(m))$ selon la loi a posteriori $P(X|Y)$;
- ▶ éliminer les p premiers échantillons pour limiter l'impact de la configuration initiale ;
- ▶ Procédure de la classification
pour tous les sites s de S :
 - approcher la marginale a posteriori à l'aide de l'équation (3-13) ;
 - chercher l'étiquette qui maximise la marginale approchée, c'est-à-dire celle qui apparaît le plus fréquemment dans la suite $(x_s(p+1), \dots, x_s(m))$.

Algorithme de MPM sur un champ de Markov.

BIBLIOGRAPHIE

- [**ABP 1992**] O.Allagnat, J-M.Boucher, D-C.He et W. Pieczynski. hidden Markov fields and unsupervised segmentation of images. In International Conférence on Pattern Recognition, Vol.03, pp.96-100, The Hague, August 1992.
- [**BCD 1983**] M. Broniatowski, G. Celeux et J. Diebolt, “Reconnaissance de mélanges de densités par un algorithme d'apprentissage probabilistes”, In Data Analysis and Informatics 3, E. Diday (Ed.), North Holland, Amsterdam, pp. 359-373, 1983.
- [**Bes 1974**] J. Besag, “Spatial interaction and the statistical analysis of lattice systems”, Journal of Royal Statistical Society, B, 36, pp 192-236, 1974.
- [**Bes 1986**] J. Besag, “Statistical analysis of dirty picture”, Journal of Royal Statistical Society, B, 48, pp 259-302, 1986.
- [**BGP 2000**] C.Banga, F.Ghorbel and W. Pieczynski. Unsupervised Bayesian Classifier Applied to the Segmentation of Retina Image. IEEE. 2000.
- [**Bos 1994**] C. A. Bouman et M. Shapiro, “A Multiscale Random Field Model for Bayesian Image Segmentation”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 3, pp. 162-177, 1994.
- [**Bps 1970**] L. Baum, T. Petrie, G. Soules and N. Weiss, “A maximisation technique occurring in the statistical analysis of probabilistic functions of Markov chains”, Ann. Math. Statist., 41, pp. 164-171, 1970.
- [**BVDO 2004**] Y. Bengio, P. Vincent, J.F. Paiement, O. Delalleau, M. Ouimet, N. Le Roux. Spectral clustering and kernel pca are learning eigenfunctions. Technical Report 2004-3, Département d'informatique et recherche opérationnelle, centre de recherches mathématiques, Université de Montréal, Canada, 2004.
- [**Cha 2000**] B. Chalmond, “Eléments de modélisation pour l’analyse d’images”, Springer, 2000.
- [**Char 2000**] A. Chardin. Modèles énergétiques hiérarchiques pour la résolution des problèmes inverses en analyse d’images : Application à la télédétection. Thèse de Doctorat, Université de Rennes1– Janvier 2000.

- [CoM 2002] Comaniciu D. and Meer P. Mean shift: a robust approach toward feature space analysis. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, 24(5):603619, 2002.
- [CoP 1995] J-P.Cocquerez et S.Philipp. *Analyse d'images : filtrage et segmentation*. Masson. 1995.
- [CST 2000] N. Cristianini and J. Shawe-Taylor. *An introduction to support vector machines (and other kernel-based learning methods)*. Cambridge University Press, 2000.
- [DeE 1987] H. Derin and H. Elliott, "Modeling and Segmentation of Noisy and textured images using Gibbs random fields", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 9 No. 1, pp. 39-55, 1987.
- [DeE 1997] H. Derin and H. Elliott, "Modeling and Segmentation of Noisy and textured images using Gibbs random fields", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 9 No. 1, pp. 39-55, 1987.
- [Del 1997] J-P. Delmas. An Equivalence of the EM and ICE Algorithm for Exponential Family. *IEEE Trans. Signal Processing*. Vol.45, N°10, pp.2613-2615, 1997.
- [DHT 2001] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley and Sons, 2001.
- [DMP 1997] Y. Delignon, A. Marzouki and W. Pieczynski, "Estimation of generalized mixture and its application in image segmentation", *IEEE Trans. on Image Processing*, Vol. 6, No. 10, pp. 1364-1375, 1997.
- [DuJ 1989] R. C. Dubes et A. K. Jain. Random fields models in image analysis. *Journal of applied statistics*, 16(2):131–164, 1989.
- [Dur 2003] J-B. Durand. *Modèles à structure cachée : inférence, sélection de modèles et applications*. Thèse de Doctorat, Université de Joseph Fourier – Janvier 2003.
- [FUM 1981] K. FU et J. MUI, A survey on image segmentation, *Pattern Recognition*, Vol. 13, 1981, pp. 3–16.

- [GeG 1984] S. Geman and D. Geman, “Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 6, pp. 721-741, 1984.
- [Gol 1989] D.E. Goldberg. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, 1989.
- [GrC 1991] S. Grossberg G. carpenter. An adaptive resonance algorithm for rapid category learning and recognition. Neural Networks, (4): 493{504, 1991.
- [Gui 1996] H. Guillemet, thèse, Ecole Nationale Supérieure des Télécommunications, 1996.
- [Guy 1995] X. Guyon, “Random fields on network. modeling, statistics, and applications” Springer-Verlag, Probability and its Applications, New York, 1995.
- [Haa 1971] Haagensen 1971. Dieae of the breast.2nd edition, Saunders, Philadelphia.
- [HPB 1994] F. Heitz, P. Perez and P. Bouthemy, “Multiscale Minimization of Global Energy Functions in Some Visual Recovery Problems”, CVGIP: Image Understanding, Vol. 59, No, pp. 125-134, 1994.
- [Isi 1925] E. Ising. Beitrag zur theorie des ferromagnetims. Zeitschrift fur Physik, 31 :253 – 258, 1925.
- [KeH 1998] C. Kervrann and F. Heitz, “A hierarchical Markov modeling approach for the segmentation and tracking of deformable shapes”, Graphical Models and Image Processing, Vol. 60, No 3, pp. 173-195, 1998.
- [Koh 1982] T. Kohonen. Clustering, taxonomy and topological maps of patterns. IEEE Sixth international conference on pattern recognition, (8):114 122, 1982.
- [Li 1995] S. Z. Li, “Markov Random Field Modelling in Computer Vision”, Springer, 1995.
- [LiK 1979] T. M. Lillesand et R.W. Kiefer. Remote Sensing and Image Interpretation. JohnWiley and Sons, Inc., 1979.

- [LBG 1980] Y. Linde, A. Buzo, R. Gray. An algorithm for vector quantizer design. IEEE Transactions on Communications, 28(1) :84-95, January 1980.
- [Mai 2003] H. Maître, “Le traitement des images”, Hermès, 2003.
- [MDP 1994] A. Marzouki, Y. Delignon and W. Pieczynski. Adaptive segmentation of SAR images. IEEE. Vol.02, pp 449-454, 1994.
- [MaP 1991] N. Marthic, W. Pieczynski. Estimation of mixture and unsupervised segmentation of images. Proceedings of IGARSS, Helsinki, Finland, pp.1083-1086, June 1991.
- [MCP 2000] M. Mignotte, C. Collet, P. Pérez, P. Bouthemy, “Markov random field and fuzzy logic modeling in sonar imagery: application to the classification of underwater floor”, Computer Vision and Image Understanding, Vol. 79, No. 1, pp. 4-24, 2000.
- [MMP 1987] J. Marroquin, S. Mitter, T. Poggio, “Probabilistic solution of ill-posed problems in computational vision”, Journal of the American Statistical Association, 82, pp. 76-89, 1987.
- [MMS 2002] M. Mignotte, J. Meunier, J.-P. Soucy, C. Janicki, “Comparison of deconvolution techniques using a distribution mixture parameter estimation: application in SPECT imagery”, Journal of Electronic Imaging, Vol. 11, No. 1, pp. 11-25, 2002.
- [MRR 1953] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, “Equations of state calculations by fast computing machines”, J. Chem. Phys., 21, pp. 1087-1092, 1953.
- [MVR 2001] J. L. Marroquin, F.A. Velasco, M. Rivera and M. Nakamura, “Gauss-Markov Measure Field Models for Low-Level Vision”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 23, pp. 337-348, 2001.
- [PaP 1993] N. PAL et S. PAL, A review on image segmentation techniques, Pattern Recognition, Vol. 26, No. 9, 1993, pp. 1277–1294.

- [Per 1998] P. Perez, “Markov random fields and images”. « CWI Quarterly », Vol. 11, No. 4, pp. 413-437, 1998.
- [PDZ 2005] G. Perrin, X. Descombes, and J. Zerubia. Adaptive simulated annealing for energy minimization problem in a marked point process application. In Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR), St Augustine, Florida, USA, novembre 2005.
- [Pey 2001] N. Peyrard. Approximation de type champ moyen des modèles de champ de Markov pour la segmentation de données spatiales, Thèse de Doctorat, Université de Joseph Fourier – Octobre 2001.
- [Pie 2003] W. Pieczynski, “Modèles de Markov en traitement d’images”, Traitement du Signal, Vol.20, No. 3, pp. 255-278, 2003.
- [Pie 1992] W. Pieczynski, “Statistical image segmentation”, Machine Graphics and Vision, Vol.1, No. 1/2, pp. 261-268, 1992.
- [PSZ 2005] G. Poggi, G. Scarpa, and J. Zerubia, “Supervised segmentation of remote sensing images based on a tree-structured MRF model”, IEEE Trans. On Geoscience and Remote Sensing, Vol. 43, No. 8, pp. 1901–1911, 2005.
- [Ric 2005] J. A. Richards, “Analysis of remotely sensed data : the formative decades and the future”, IEEE Trans. on Geoscience and Remote Sensing, Vol. 43, No. 3, pp. 422-432, 2005.
- [Rob 1991] C. Robert, “L’analyse statistique Bayésienne”, Economica, 1991.
- [Roe 1986] Roebuck 1986. mammography and screening for breast cancer. Brit. Med. .J.292, 223-226
- [Ros 1999] C. Rosenberger. Mise en œuvre d’un système adaptatif de segmentation d’images. Thèse de Doctorat, Université de Rennes1 – Décembre 1999.

- [**RPB 2003**] S. Reed, Y. Petillot, and J. Bell, “An automatic approach to the detection and extraction of mine features in sidescan sonar”, *IEEE Journal of Oceanic Engineering*, Vol. 28, No. 1, pp. 90-105, 2003.
- [**SaC 2006**] F. Salzenstein and C. Collet. Fuzzy markov random fields versus chains for multispectral image segmentation. A paraître : *IEEE Trans. On PAMI*, Jun 2006.
- [**Soi 2004**] P. Soille. *Morphological image analysis*. Springer, 2004.
- [**TNT 2004**] C. Tison, J.M. Nicolas, F. Tupin and H. Maître, “ A new statistical model of urban areas in high resolution SAR images for Markovian segmentation”, *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 42, No. 10, pp. 2046-2057, 2004.
- [**TuR 2005**] F. Tupin and M. Roux, “Markov random field on region adjacency graphs for the fusion of SAR and optical data in radar grammetric applications”, *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 43, No. 8, pp. 1920-1928, 2005.
- [**Win 1995**] G. Winkler. *Image analysis, random fields and dynamic Monte Carlo methods: a mathematical introduction*. Springer Ed., 1995.
- [**Win 2003**] G. Winkler, “Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A mathematical introduction”, Springer, 2003.
- [**You 1988**] L. Younes, “Estimation and annealing for Gibbsian fields”, *Annales de l’institut Henri Poincaré*, 2, 1988.