

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des sciences et de la technologie Houari BOUMEDIENE
Faculté d'Electronique et d'Informatique



MEMOIRE

Présenté pour l'obtention du diplôme de MAGISTER

EN : ELECTRONIQUE

Spécialité : Traitement du Signal et des Images

Par : ZEMMOURI Farid

SUJET :

**Segmentation des images IRM par Chaîne de
Markov Cachée**

Soutenu le 07 décembre 2008, devant le jury composé de :

Mr Y. SMARA	Professeur,	USTHB Président
Mme L. HAMAMI	Maître de Conférences,	ENP Directrice de Thèse
Mme A. SERIR	Maître de Conférences,	USTHB Examineur
Mr M. DEBYECHE	Maître de Conférences,	USTHB Examineur
Mme H. BENTOUMI	Chargée de Cours,	USTHB Invité

Remerciements

J'adresse mes plus vifs remerciements à Madame L.Hamami, Directrice du Laboratoire Signal et Communications de l'Ecole Nationale polytechnique, qui a encadrée ce mémoire. Elle m'a fait connaître le domaine du traitement d'images et a su diriger ce travail avec efficacité et rigueur tout en me laissant libre dans mes choix. Sans parler de ses conseils avisés et de l'aide nécessaire qu'elle a bien voulu m'apporter.

Je remercie Monsieur Y. Smara qui m'a fait l'honneur de présider le jury de soutenance. J'adresse mes remerciements également à Mesdames A. Serir et H. Bentoumi et Monsieur M.DEBYECHE pour l'honneur qu'ils me font en jugeant ce travail.

Je remercie chaleureusement Dr. Mohamed.DEKMOUN, radiologue au service d'IRM du Centre Hôpitalo- Universitaire de Bab El Oued à Alger, qui m'a orienté dans le domaine de la radiologie. Je lui suis très reconnaissant pour ses éclaircissements et pour le temps qu'il a consacré en sélectionnant les images de la base de données.

Je tiens aussi à remercier ma famille pour leur soutien moral à savoir mes très chers parents pour leurs encouragements et leur patience durant toutes mes années d'étude, mes frères et sœurs ainsi que toute la famille ZEMMOURI.

Table des matières

Remerciements

Table des matières

Liste des figures

Liste des tableaux

Abréviations

Introduction générale.....	1
Chapitre 1 : Imagerie par résonance magnétique IRM	
1. Introduction.....	6
2. Principe physique des images IRM.....	6
2.1 Origine quantique.....	6
2.2 Effet du champ magnétique statique sur un échantillon.....	7
2.3 Récupération d'un signal radiofréquence (RF)	8
3. La formation des images IRM.....	11
4. Défauts des images acquises par IRM.....	12
4.1 Les facteurs liés au mouvement.....	13
4.2 Le déplacement chimique.....	13
4.3 Le champ statique.....	13
4.4 Le recouvrement de spectre.....	13
4.5 Le bruit.....	13
4.6 L'effet de volume partiel.....	14
5. Domaines d'application de l'IRM.....	14
5.1 L'IRM anatomique.....	15
5.2 L'IRM fonctionnelle.....	16
6. Conclusion.....	17
Chapitre 2 : Méthodes de segmentation des images	
1. Introduction.....	19
2. Synthèse des approches de segmentation des images.....	19
2.1 Approches statistiques.....	20
2.1.1 Paramètres du premier ordre.....	20
2.1.2 Paramètres du second ordre.....	21
2.2 Les méthodes basées sur un modèle.....	21

2.2.1 Modèle fractal.....	21
2.2.2 L'approche multifractale.....	21
2.2.3 Modèles stochastiques.....	24
2.3 Les méthodes de morphologie mathématique.....	24
2.3.1. Ligne de partage des eaux hiérarchique : algorithme des cascades.....	24
2.3.2. Connexions par sauts.....	25
2.3.3. Zones plates et quasi-plates.....	26
2.3.4. Fusion de région dans un treillis de partitions.....	26
3. Méthodes du traitement de l'incertitude et de l'imprécision.....	27
4. Les méthodes de segmentation non supervisé.....	28
4.1 Segmentation par les K-moyennes.....	28
4.1.1 Nuées dynamiques.....	28
4.1.2 Mise en pratique des K-moyennes.....	30
4.1.2.1 Initialisation.....	31
4.1.2.2 Classification.....	31
4.1.2.3 Mise à jours des moyennes.....	31
4.1.2.4 Conclusion et remarques.....	32
4.2 Modèles flous en analyse d'images.....	32
5. Conclusion.....	35
Chapitre 3 : Chaîne de Markov Cachée CMC	
1. Introduction.....	37
2. Modélisation d'image.....	38
3. Parcours de Hilbert-Peano.....	39
3.1. Construction de Parcours de Hilbert.....	39
3.2. Contraintes	41
4. Décision bayésienne.....	41
4.1 Maximum A Posteriori : MAP.....	42
4.2 Marginal Posterior Mode : MPM.....	42
5. Chaînes de Markov Caché CMC.....	43
5.1 Modèle de chaîne de Markov cachée.....	43
5.1.1. Loi <i>a priori</i> de X.....	43
5.1.2. Matrice d'observation.....	44
5.1.3. Matrice de la loi initiale.....	44
5.1.4. Matrice de Transition.....	44

5.1.5 Loi du couple (X, Y)	45
5.2 Estimation des paramètres.....	47
5.2.1 Algorithme d'estimation: ICE.....	47
5.2.2 Procédure Forward-Backward.....	48
5.2.3 Ré-estimation des paramètres de la chaîne de Markov.....	50
5.2.4 Ré-estimation des paramètres des lois d'attache aux données.....	51
5.2.5 Classification bayésiens.....	53
6. Résultats et évaluation.....	54
6.1 Introduction.....	54
6.2 Images utilisées.....	55
6.2.1 Images simulées : les fantômes du MNI.....	55
6.2.2 Images réelles.....	55
6.3 Critère d'évaluation quantitative.....	55
6.4 Segmentation des images IRM.....	58
6.4.1 Segmentation des fantômes.....	59
6.4.1.1 Comparaison avec la méthode de croissance de région.....	61
6.4.1.2 Comparaison avec la méthode d'ondelletes.....	64
6.4.2 Segmentation des images réels.....	65
7. Conclusion.....	67
Chapitre 4 : Chaîne de Markov Caché floue CMC-f	
1. Introduction.....	69
2. Modélisation d'une chaîne de Markov floue.....	69
3. Estimation des paramètres de la chaîne de Markov floue.....	71
3.1 Les probabilités Forward-Backward.....	71
3.2 Ré-estimation des paramètres de la chaîne de Markov floue.....	72
3.3 Ré-estimation des paramètres des lois d'attache aux données.....	73
3.3.1 Cas des classes dures.....	73
3.3.2 Cas de la classe floue.....	73
3.4 Classification bayésienne.....	74
4. Approximation numérique du modèle.....	74
5. Segmentation par Chaîne de Markov Floue.....	75
6. Conclusion.....	79

Chapitre 5 : Présentation de logiciel CMC_IRM _Segmentation

1. Introduction.....	81
2. Logiciel utilisé Builder C++.....	81
3. Logiciels développés.....	82
3.1 Les menus.....	82
3.1.1 Menu Fichier.....	82
3.1.2 Menu Pré-traitement.....	83
3.1.3 Menu Segmentation.....	84
Conclusion générale.....	86

Références bibliographiques

Listes des figures

FIG 1.1 : Les deux orientations de spin d'un proton.....	7
FIG 1.2 : Précession de l'aimantation dans un référentiel tournant.....	8
FIG 1.3 : Relaxation transversale.....	10
FIG 1.4 : Relaxation longitudinale.....	10
FIG 1.5 : Signal FID recueilli par les bobines réceptrices.....	11
FIG 1.6 : Représentation d'une image volumique, constituée de trois coupes perpendiculaires.....	12
FIG 1.7 : L'effet de volume partiel en IRM.....	14
FIG 1.8 : Exemples des images réelles pondérées T1, (a) Moëlle cervicale, (b) Genou, (c) Cerveau, (d) Bassin.....	15
FIG 1.9 : Exemples d'application : reconnaissance des sillons, classification.....	16
FIG 1.10 : Exemple d'IRM fonctionnelles.....	17
FIG 2.1 : segmentation d'une image IRM pondéré T1 d'un tissu cérébral à 5 classes.....	31
FIG 3.1 : Parcours de Hilbert, Ordre 1, 2, 3.....	40
FIG 3.2 : Construction du parcours de Hilbert pour une image 8×8 (initialisation, étape intermédiaire et parcours final).....	40
FIG 3.3 : Construction du parcours de Hilbert pour une image 3 D.....	41
FIG 3.4 : Modèle Gauche Droite (Bakis) de transition entre les états.....	45
FIG 3.5 : Schéma présentant les liens entre les processus X et Y dans un modèle de chaîne de Markov cachée.....	46
FIG 3.6 : Organigramme de l'algorithme d'estimation de paramètres par la méthode ICE dans le contexte des chaînes de Markov cachées.....	53
FIG 3.7 : Image de référence du BrainWeb.....	55
FIG 3.8 : Définition des valeurs de Vrais Positifs (VP), Faux Positifs (FP) et Faux Négatifs (FN) par comparaison de l'image de référence et de l'image segmentée.....	56
FIG 3.9 : les trois tissus séparés proposés dans le site BrainWeb, (a) LCR, (b) MG, (c) MB.....	58
FIG 3.10 : Etapes de la segmentation par Chaîne de Markov cachée.....	58
FIG 3.11 : a) image coronale 0% de bruit et 0% de non-uniformité, b) image segmentée en 3 tissus.....	59

FIG 3.12 : a) image transversale 40% de bruit et 9% de non-uniformité, b) image segmentée en 5 classes.....	60
FIG 3.13 : convergence du modèle CMC.....	60
FIG 3.13 : Résultats de CMC sur des images fantômes et comparaison avec les résultats de la segmentation par croissance de régions.....	63
FIG 3.14 : résultats de CMC sur des images fantômes et comparaison avec les résultats de la segmentation par ondelettes.....	65
FIG 3.15 : (a) image réelle sagittale, (b) image segmentée en 6 classes par CMC.....	66
FIG 3.16 : convergence du modèle CMC.....	66
FIG 3.17 : Loi d'attache aux données.....	66
FIG 3.18 : (a) image réelle sagittale du genou, (b) image segmentée en 5 classes par CMC...	67
FIG 4.1 : a) image fantôme 0% de bruit et 5% de non-uniformité, b) image segmenter en 3 tissus, c) image 0% de bruit et de 3% de non-uniformité segmenter en 3 tissus.....	75
FIG 4.2 : Etapes de la segmentation par chaîne de Markov floue.....	76
FIG 4.3 : a) image brute 0% de bruit, 5% de no-uniformité édité, b) image segmenter CMC-f ou $f(\Omega_1) = 1$, c) CMC-f ou $f(\Omega_1) = f(\Omega_2) = 1$, b) CMC-f ou $f(\Omega_1) = 2$, $f(\Omega_2) = 1$	77
FIG 4.4 : a) image brute 0% de bruit, 3% de non-uniformité, b) CMC classique, b) CMC-f ou $f(\Omega_1) = 1$, d) CMC-f ou $f(\Omega_1) = 2$ et $f(\Omega_2) = 1$	77
FIG 4.5 : a) image brute 40% de bruit et 9% de non-uniformité coupe 3mm, b) CMC-f ou $f(\Omega_1) = 1$, c) CMC-f ou $f(\Omega_1) = f(\Omega_2) = 1$	77
FIG 4.6 : a) image segmentée par CMC-f où $f(\Omega_1) = 1$, b) CMC-f où $f(\Omega_1) = f(\Omega_2) = 1$...	78
FIG 4.7 : a) CMC-f où $f(\Omega_1) = 1$, b) CMC-f où $f(\Omega_1) = 2$ et $f(\Omega_2) = 1$	78
FIG 5.1 : Fenêtre principale.....	82
FIG 5.2 : Menu Fichier.....	82
FIG 5.3 : Boîte de sélection de l'image à traiter.....	83
FIG 5.4 : Menu Pré-traitement.....	84
FIG 5.5 : Menu Segmentation.....	84
FIG 5.6 : Fenêtre d'apprentissage.....	84
FIG 5.7 : Segmentation d'image 0% de bruit et 5% de non-uniformité en 4 classes et filtré par un filtre médian.....	85

Liste des tableaux

Tableau 1.1 : Ordre de grandeur des temps de relaxation à 1.5 Tesla.....	11
Tableau 3.1 : Valeurs du critère de Jaccard pour des images simulées de différents niveaux de bruit et de non-uniformité.....	57
Tableau 3.2 : Valeurs d'évaluation quantitative par rapport à différents paramètres de bruit et de non-uniformité.....	61

Abréviations

BOLD	Blood Oxygenation Level-Dependent.
CMC	Chaînes de Markov Cachée.
CMC-f	Chaîne de Markov Cachée floue.
EM	Estimation Maximization.
FID	Free Induction Decay.
FCM	<i>Fuzzy C -Mean</i>
HMRF	Champs de Markov caché (Hidden Markov Random Fields).
ICE	Iterative Conditional Estimation.
LPE	Ligne de Partage des Eaux.
MAP	Maximum A Posteriori.
MPM	Marginal Posterior Mode.
RMN	Résonance Magnétique Nucléaire.
SEM	Stochastic Estimation Maximization.

Introduction générale

Les outils technologiques dans le domaine de la médecine sont en plein développement, principalement ceux qui touchent le diagnostic, le contrôle et le suivi des maladies qui concernent les organes internes du corps humain. Il existe plusieurs techniques d'imagerie médicale permettant l'observation des organes internes, nous en citons la radiologie par Rayons X, les Ultrasons, l'IRM, ...etc. Dans ce présent travail, nous nous intéressons à l'Imagerie par Résonance Magnétique IRM.

L'IRM est une technique d'imagerie médicale non-invasive et extrêmement précise spatialement, qui peut permettre l'observation anatomique des organes internes ainsi que la détection d'éventuelles anomalies.

Le scanner IRM permet l'observation des différentes parties du corps humain, le cerveau, la colonne vertébrale, le bassin et le genou, ...etc.

Les chercheurs s'intéressent plus aux images médicales du cerveau car c'est l'organe le plus sensible et probablement le plus complexe dans le corps humain.

Pour faire le suivi de certaines maladies liées à des lésions cérébrales internes, le médecin doit analyser des images médicales. Pour étudier l'évolution d'une tumeur, il est nécessaire de connaître avec exactitude les changements survenus. Or l'interprétation visuelle des images médicales, n'est pas toujours fiable. De plus, il est parfois nécessaire d'analyser plusieurs images avant d'aboutir à la décision finale. Par conséquent, il est souhaitable et même nécessaire de disposer d'outils d'aide au diagnostic. Ces systèmes d'aide permettent aux médecins d'avoir des informations exactes de manière rapide quant aux caractéristiques des éléments contenus dans l'image. Ils permettent aussi aux médecins de naviguer dans un organe dans l'espace tridimensionnel ou même en 4D pour des images animées.

Pour un diagnostic fiable, dans le domaine médical, la précision est importante. En termes d'analyse d'images, il est nécessaire que le traitement soit précis. Les possibilités de traitement automatique de ces images s'avèrent pourtant délicates, car des capacités aussi banales pour l'œil humain que la reconnaissance d'un objet pose de réelles difficultés pour l'outil informatique. De plus, la plupart des images médicales contiennent plusieurs artefacts qui rendent difficile le traitement. Ces artefacts sont principalement dus aux bruits, au patient lui-même à cause de ses mouvements pendant l'acquisition de l'image,... Les Images par Résonance Magnétique (IRM) sont particulièrement caractérisées par *l'effet de volume partiel* que l'on retrouve lorsqu'un pixel ayant un certain niveau de gris correspond en réalité à un mélange de

deux ou plusieurs tissus ; les pixels sont ici appelés *mixels*. Cet artefact existe principalement aux frontières entre tissus.

L'objectif de notre travail est de concevoir un système de segmentation d'images médicales. Nous nous concentrons sur la qualité de celles-ci, car tout traitement ultérieur dépend étroitement du résultat de la segmentation. L'objectif de ce traitement pouvant être l'étude de l'évolution d'une tumeur ou même d'un changement d'état non pathologique.

La segmentation constitue un des outils majeurs du traitement d'images. L'objectif consiste à restituer, à partir d'une ou plusieurs observations bruitées, une image thématique de l'organe observé. Par exemple pour l'IRM du cerveau, la segmentation extrait les trois principales matières composant le cerveau : la matière blanche (MB), la matière grise (MG) et le Liquide Céphalo-Rachidien (LCR). Parmi les nombreuses techniques développées depuis une quarantaine d'années, les méthodes statistiques, qui s'appuient sur des modèles probabilistes, ont été utilisées avec succès. Parmi ces méthodes, on distingue classiquement deux grandes familles d'algorithmes de segmentation : les méthodes locales/contextuelles et les méthodes globales. Les méthodes locales/contextuelles traitent les pixels de l'image de façon indépendante ou en prenant en considération un voisinage immédiat du pixel considéré, alors que les méthodes globales, dont font partie le modèle de Markov CMC sur lesquels s'appuient notre travail, traite simultanément toute l'information contenue dans l'image.

L'essor des modélisations markoviennes en traitement d'images, a été introduit dans les années 70-80 par les champs de Markov cachés (« Hidden Markov Random Fields », HMRF). Le succès des modélisations markoviennes dans le contexte du traitement d'images est principalement dû au fait que, lorsque l'image thématique recherchée peut être représentée par un processus de Markov fini et lorsque la structure du bruit n'est pas trop complexe, alors cette image thématique peut être reconstruite à partir de l'image bruitée à l'aide de différents critères de classification bayésiens.

Dans les années 90, une autre solution, proposée par Benmiloud [7] dans le domaine de traitement d'image de télédétection, a consisté à modéliser le processus caché par une chaîne de Markov CMC (processus 1D). Cette alternative nécessite alors l'utilisation d'un parcours (par exemple le parcours de Hilbert-Peano) pour transformer l'image 2D en un signal 1D. Dans ce contexte, l'estimation des paramètres du modèle avec différents algorithmes (EM, SEM, ICE) permet d'exprimer la loi de Chaîne, ainsi que les solutions bayésiennes, de façon explicite, sans mise en oeuvre de méthodes d'approximation. De façon générale, le modèle de CMC a permis de traiter de nombreux problèmes de traitement de signal, par exemple en traitement de la parole, en génétique, en reconnaissance de l'écriture manuscrite, en robotique, en reconnaissance de visage

[16], en traitement de signal biomédical. En ce qui concerne la segmentation d'images, les CMC ont aussi été utilisées avec succès dans divers contextes, comme en segmentation d'images de télédétection [13], d'images sonar, ou même d'images vidéo [31]. Dans le contexte de la segmentation d'images, les modélisations basées sur la chaîne de Markov constitue une alternative intéressante au modèle de champ de Markov caché, même si ces derniers offrent une modélisation précise et plus intuitive des relations spatiales dans l'image. Le gain en temps de calcul est généralement très significatif, principalement en raison de la complexité de l'étape d'estimation de paramètres des HMRF, ce qui donne aux CMC un réel intérêt pour les problèmes de segmentation et d'analyse d'image.

Lors de l'acquisition d'IRM, on doit choisir l'épaisseur des coupes, c'est-à-dire la distance entre les différentes coupes ; plus cette distance est petite, plus le nombre de coupes nécessaire pour représenter tout volume est grand. L'effet de volume partiel est étroitement lié à l'épaisseur des coupes, plus elles diminuent et plus l'effet de volume partiel diminue. Cependant, un nombre de coupes élevé peut devenir une difficulté pour le traiteur d'images. L'utilisation des « Chaîne de Markov Cachée floue » permet de pallier le problème de l'effet de volume partiel tout en utilisant des images avec une assez grande épaisseur de coupes.

Dans ce nouveau contexte, il a été proposé dans la littérature d'utiliser, non pas une mesure de l'incertitude sur les observations (caractéristique de l'approche probabiliste des CMC classiques), mais une mesure de l'imprécision (caractéristique de l'approche floue). Ces modèles markoviens s'appuient sur la théorie des ensembles flous introduits par Zadeh en 1965, en intégrant une mesure de l'imprécision dans le modèle classique des Chaînes de Markov Cachées (CMC). Le modèle proposé dans [13] permet de conserver les propriétés de robustesse au bruit des CMC (liées à la mesure d'incertitude) et d'enrichir ce dernier avec une caractéristique floue (mesure d'imprécision).

Dans notre travail, nous utilisons les deux approches CMC et CMC floue. Les CMC permettent de segmenter les images de 1mm d'épaisseur et d'économiser le temps de calcul. Les CMC floues offrent la possibilité de segmenter des images en tenant compte de l'effet de flou ; car par expérience dans le cas des images présentant un flou, on peut augmenter le nombre de classes à détecter et ainsi en plus des classes caractérisant les tissus homogènes on obtient des classes représentant l'effet de flou.

Ce mémoire s'articule autour de cinq chapitres qui nous permettrons de présenter les différents aspects de notre travail.

Le premier chapitre est dédié au cadre d'application de ce mémoire. Il présente quelques notions et principes fondamentaux de la technique de l'imagerie par résonance magnétique et il fournit une description générale de l'objectif de notre travail.

Le second chapitre présente un état de l'art des méthodes de segmentation d'images, plus particulièrement des images IRM.

Le troisième chapitre est consacré à la chaîne de Markov cachée et leur utilisation en segmentation des images IRM au cours de notre travail avec les différentes étapes par lesquelles nous sommes passés. Dans le même chapitre on essaye de présenter les résultats que nous avons obtenus et leur évaluation à travers des images simulées et un critère d'évaluation quantitative. On présente également nos résultats sur des images réelles.

Le quatrième chapitre décrit l'approche floue dans le contexte de chaîne de Markov cachée floue et on présente les résultats obtenus sur des images IRM fantômes.

Enfin, le cinquième chapitre présente brièvement l'application développée par le logiciel Builder C++.

Chapitre 1

Principes de l'Imagerie par Résonance Magnétique

1. Introduction

La RMN est une technique d'imagerie très utilisée dans le domaine médical. Elle permet d'observer de manière non invasive la distribution spatiale de diverses propriétés intrinsèques des tissus et délivre directement des images numérisées fournissant des informations médicales au praticien.

Pour réaliser des images par résonance magnétique, le patient est placé à l'intérieur d'une machine produisant un champ magnétique statique intense (environ de 1.5T). L'analyse des signaux émis par le corps du sujet permet la construction de trois types d'image selon trois angles de vue. Les IRM offrent la possibilité d'observer l'anatomie des organes, de faire la localisation et le diagnostic de certaines lésions ou même de faire le suivi d'un changement d'état non pathologique [1, 2, 11].

L'objectif de ce chapitre est de présenter de façon détaillée les principes physiques de l'imagerie par résonance magnétique. Nous allons voir les caractéristiques et les défauts des IRM ce qui nous aidera dans son analyse.

2. Principe physique de l'imagerie par résonance magnétique

2.1 Origine quantique

Les particules élémentaires comme les protons possèdent un moment angulaire intrinsèque, appelé spin, qui caractérise leur rotation sur elles-mêmes. A ce spin est associé un moment magnétique intrinsèque $\vec{m}$ traduisant le fait qu'une particule puisse se comporter comme une aiguille aimantée sous l'influence d'un champ magnétique. En pratique, certains noyaux atomiques, tel celui de l'hydrogène (en abondance dans les tissus humains) présentent un moment cinétique de spin et un moment magnétique non nuls qui leur confèrent certaines propriétés sous l'influence d'un champ magnétique.

En présence d'un champ statique $\vec{B}_0$, la mécanique quantique attribue à un proton deux états possibles correspondant à un état stable en énergie (état de spin parallèle, c'est le plus probable selon la physique statistique de Boltzmann) et un état instable (état de spin antiparallèle) comme indiqué sur la figure 1.1 (a) et (b).

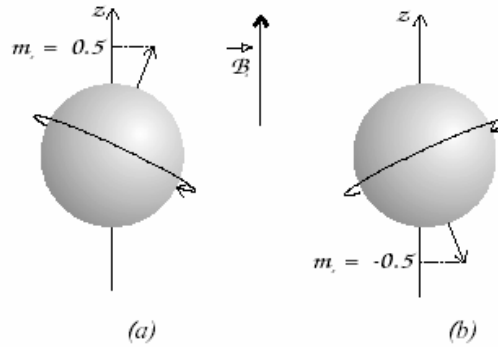


Figure 1.1 : Les deux orientations de spin d'un proton.

A gauche est représenté l'état de basse énergie d'une particule pour lequel la projection de son spin sur l'axe z est parallèle au champ magnétique. A droite est représenté l'état de haute énergie, moins stable que le précédent. L'axe z est parallèle au vecteur champ magnétique statique.

Par convention, on dit que ce proton est de spin $\frac{1}{2}$. De plus, sous l'influence de $\vec{B}_0$, le moment magnétique du proton effectue un mouvement de précession. Dans ce mouvement, il balaye un cône dont l'axe est parallèle au vecteur champ magnétique à une fréquence caractéristique appelée fréquence de Larmor :

$$f_0 = \frac{\gamma}{2\pi} \cdot B_0 \quad (1.1)$$

Où γ , le rapport gyromagnétique, est spécifique au corps étudié.

2.2 Effet du champ magnétique statique $\vec{B}_0$ sur un échantillon

Les appareils d'imagerie par résonance magnétique ont une résolution qui ne leur permet d'observer que l'aimantation résultante (ou moment macroscopique du corps d'origine nucléaire) d'un très grand nombre de protons. L'approche quantique du problème peut être alors remplacée par une approche classique où l'échantillon présente une aimantation nucléaire résultante $\vec{M}(x, y, z)$ au point (x, y, z) .

En présence d'un champ statique uniforme, l'aimantation résultante d'un échantillon homogène sera alors alignée avec ce champ statique en régime d'équilibre (figure 1.2).

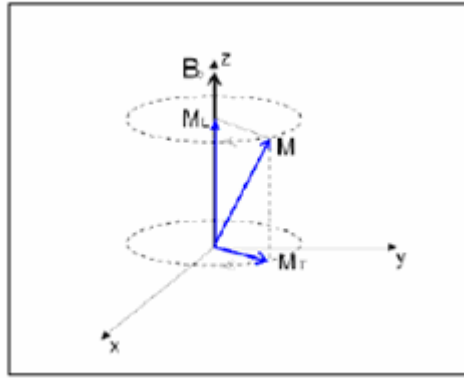


Figure 1.2 : Précession de l'aimantation dans un référentiel tournant.

2.3 Récupération d'un signal radiofréquence (RF)

Une brève impulsion RF modifie la direction de l'aimantation de l'échantillon. Quand l'onde RF d'excitation est supprimée, ou, en d'autres termes, lorsque le champ tournant n'est plus appliqué, l'aimantation va progressivement revenir à sa situation d'équilibre en raison de phénomènes de relaxation (non pris en compte dans la présentation simple précédente). Ce retour à l'équilibre est accompagné d'une faible émission d'énergie à la fréquence propre de précession, sous forme d'une onde électromagnétique. Cette onde induit un courant électrique dans les bobines radiofréquences réceptrices (antennes) placées à proximité de l'organe imagé. L'évolution dans le temps du courant électrique dans ces bobines traduit celle du moment magnétique macroscopique. Mais toute l'énergie absorbée par la matière lors du processus d'excitation n'est pas restituée en une seule fois, c'est-à-dire que tous les noyaux ne retournent pas à leur position initiale de la même façon. Ceci est dû à deux causes :

1. La première cause est que quand un noyau bascule, l'énergie qu'il émet peut être absorbée immédiatement par un noyau voisin qui va lui aussi basculer. C'est la relaxation spin-spin caractérisée par une constante de temps T_2 ou temps de relaxation transversale. L'hétérogénéité du champ statique réduit cette constante, on parle alors de temps de relaxation effectif T_2^* .
2. La deuxième cause est que l'énergie libérée par un noyau revenant dans sa position initiale peut être absorbée par le réseau dans lequel le noyau se trouve lié, provoquant une excitation de ce réseau qui retarde l'émission d'énergie. C'est la relaxation spin-réseau caractérisée par la constante de temps T_1 ou temps de relaxation longitudinale.

Les conséquences de ces relaxations se retrouvent dans le processus de décroissance des composantes longitudinale et transversale de l'aimantation résultante lors du retour dans sa position initiale. Considérons le cas particulier où $\vec{M}$ est initialement colinéaire au champ magnétique statique. L'application d'une impulsion RF fait basculer l'aimantation d'un angle θ , rendant ainsi sa composante transversale non nulle et réduisant l'amplitude de sa composante longitudinale. Il est alors possible d'observer comment s'effectue la relaxation de chacune des composantes de l'aimantation résultante.

La composante M_x de l'aimantation sur un axe x quelconque dans le plan transversal s'écrit:

$$M_x(t) = M_x(t_0) e^{-\frac{t}{T_2}} \cos \omega_0 t \quad (1.2)$$

Où $M_x(t_0)$ est la composante de $\vec{M}$ sur l'axe x juste après l'application de l'impulsion RF, à l'instant t_0 . Ce signal, représenté dans la figure 1.3, est appelé signal de précession libre (Free Induction Decay : FID voir Figure 1.5). Il subit une décroissance exponentielle décrite par l'équation de la composante transversale M_{xy} :

$$M_{xy}(t) = M_{xy}(t_0) e^{-\frac{t}{T_2}} \quad (1.3)$$

Où $M_{xy}(t_0)$ est la composante transversale juste après l'application de l'impulsion RF.

De la même manière, la composante longitudinale M_z de l'aimantation résultante sur l'axe z (voir figure 1.4) s'écrit:

$$M_z(t) = M_z(t_0) \left(1 - e^{-\frac{t}{T_1}} \right) \quad (1.4)$$

En réalité, l'échantillon étudié est en général une combinaison de différents noyaux et de différents environnements, le signal de précession libre est donc la superposition de signaux issus de chaque entité. Même s'il est vrai que l'allure du signal est modifiée lorsque l'on considère l'ensemble des noyaux, il n'en demeure pas moins qu'elle se rapproche de celle

représentée en figure 1.3. Une étude plus détaillée des phénomènes de relaxation permet de remonter aux liaisons chimiques des noyaux : c'est l'objet de la spectroscopie par RMN.

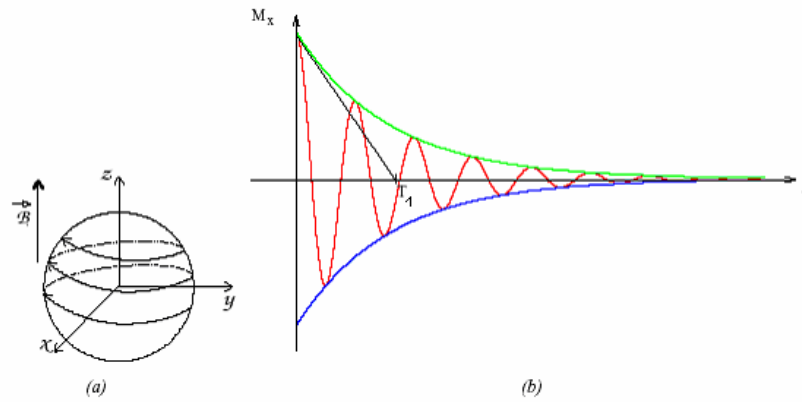


Figure 1.3 : Relaxation transversale

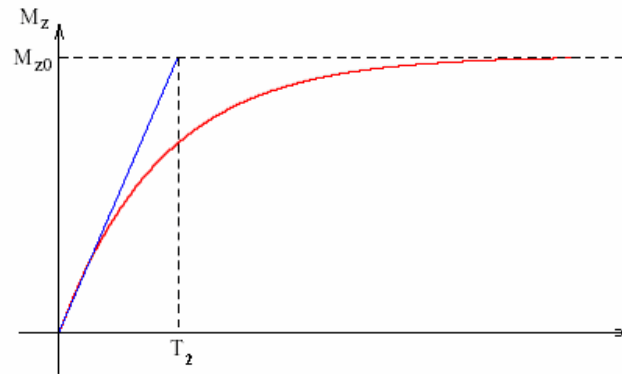


Figure 1.4 : Relaxation longitudinale

La figure 1.4 présente le retour à l'équilibre de la composante longitudinale de l'aimantation résultante selon la constante de temps T_2 .

Le tableau 1.1 porte des valeurs des deux temps de relaxation, ces derniers dépendent des tissus biologiques rencontrés et du champ magnétique statique.

	T1	T2
Liquide céphalo-rachidien	2500 ms	2000 ms
Matière grise	900 ms	90 ms
Matière blanche	750 ms	80 ms
Graisse	300 ms	40 ms

Tableau 1.1 : Ordre de grandeur des temps de relaxation à 1.5 Tesla

3. La formation des images IRM:

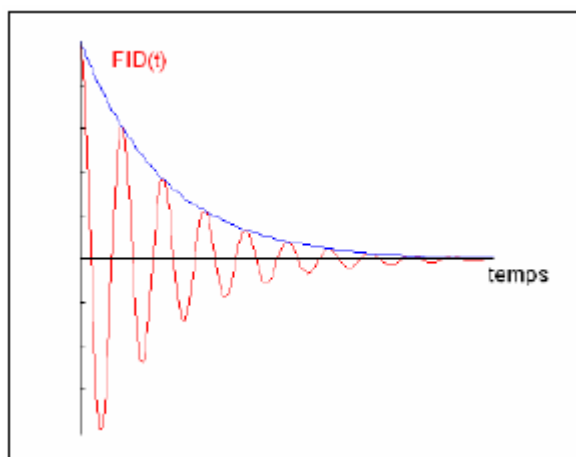


Figure 1.5: Signal FID recueilli par les bobines réceptrices

Comme nous l'avons évoqué plus haut, le retour à l'équilibre du proton induit un champ magnétique FID (Free Induction Decay), ce signal sera transformé en signal électrique puis en une image numérique.

L'image numérique IRM représente ainsi la répartition spatiale d'une grandeur physique. Chaque élément d'une coupe correspond à la représentation de l'aimantation nucléaire d'un volume de l'espace exploré. Ces éléments sont appelés pixels quand on considère individuellement une coupe et voxels quand on considère le volume constitué de l'empilement des coupes.

La problématique liée au codage de l'image est alors de savoir différencier les signaux issus des différentes zones examinées. Ce codage est réalisé grâce au couplage de trois gradients différents :

- Le gradient de coupe sélective qui sélectionne le plan de coupe, celui-ci pouvant être quelconque dans l'espace.
- le gradient de codage de phase qui sélectionne les lignes dans le plan de coupe sélectionné.
- le gradient de fréquence qui sélectionne les colonnes dans le plan de coupe sélectionné.

L'orientation des coupes, qui dépend en particulier du gradient de coupe sélective, varie en fonction de la pathologie recherchée. On retrouve cependant trois orientations principales : axiale, coronale et sagittale (figure 1.6).

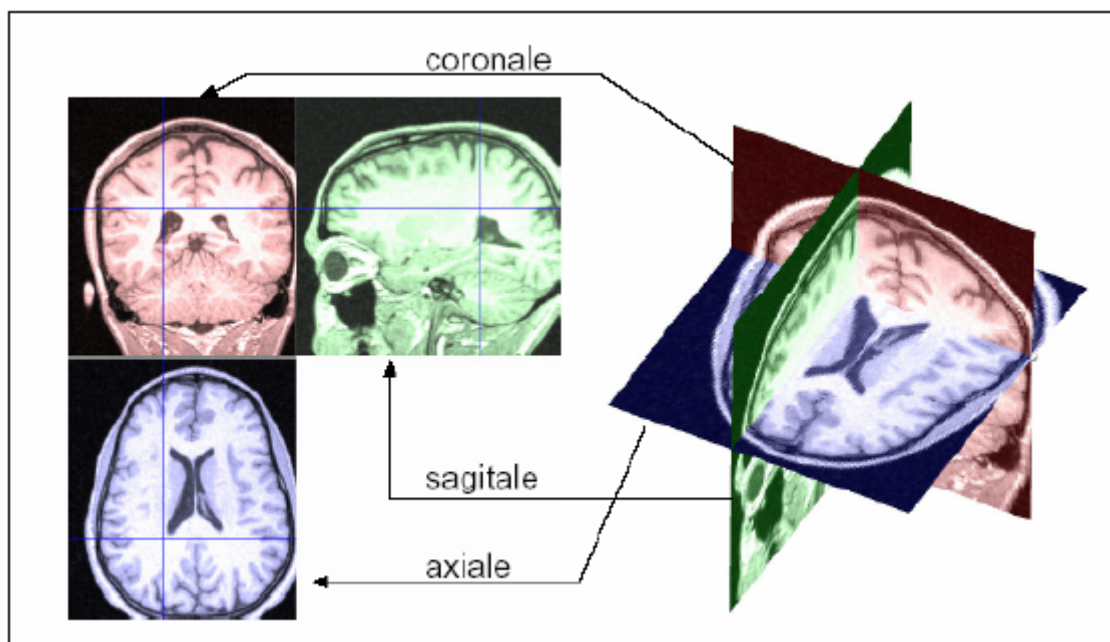


Figure 1.6 : Représentation d'une image volumique, constituée de trois coupes perpendiculaires.

4. Défauts des images acquises par IRM

Trois paramètres entrent en jeu dans la formation d'une image : la densité protonique ρ , le temps de relaxation spin-spin T_1 et le temps de relaxation spin-réseau T_2 . Mais d'autres facteurs que l'on ne maîtrise pas toujours affectent la qualité des images et provoquent des artefacts, compliquant la segmentation des tissus cérébraux ou la reconnaissance de structures cérébrales.

4.1 Les facteurs liés au mouvement

Tout mouvement de protons lors de la formation de l'image a des conséquences sur la qualité de l'image acquise. De même, comme en photographie, les mouvements du sujet nuisent à la qualité de l'image.

Dans les vaisseaux, les protons possèdent un mouvement plus ou moins rapide qui va perturber les moments magnétiques longitudinaux et transversaux des régions analysées et donc modifier des signaux de résonance magnétique.

4.2 Le déplacement chimique

Dans les milieux biologiques, les électrons ne sont pas libres mais liés à une molécule. Dans ces conditions, le champ magnétique effectif qui agit sur un proton n'est pas véritablement celui escompté car il est modifié par son environnement chimique. Il en résulte une modification de la fréquence de résonance des protons, ou déplacement chimique.

4.3 Le champ statique

Théoriquement, on considère que le champ statique est homogène. En pratique, il ne l'est pas tout à fait. On parle alors de l'hétérogénéité du champ qui modifie localement la fréquence de résonance des protons. De plus, les variations locales de susceptibilité magnétique vont provoquer des distorsions du champ magnétique.

4.4 Le recouvrement de spectre

Il y a recouvrement de spectre lorsque le champ d'exploration est plus petit que l'échantillon étudié. Ce phénomène se traduit par une superposition d'informations, situées en dehors du champ d'exploration, sur l'image.

4.5 Le bruit

Le bruit en RMN a un spectre de fréquence très large, assimilable à celui d'un bruit blanc. Ses origines sont multiples. L'ensemble des fréquences communes au bruit et au signal dépend de la bande passante englobant toutes les fréquences de résonance du signal électromagnétique. Plus celle-ci est large, moins le rapport signal sur bruit est bon.

4.6 L'effet de volume partiel

L'effet de volume partiel apparaît lorsque plusieurs tissus occupent le même élément de volume ou voxel. La conséquence en est que le niveau de gris associé à de tels voxels, appelés mixels, est obtenu à partir d'un signal hybride. Cet artefact se produit dans les cas suivants :

- Le voxel imagé chevauche la frontière entre deux ou plusieurs tissus différents (figure 1.7). Ce cas de figure est prépondérant en IRM.
- La résolution spatiale dépasse la taille des structures présentes dans l'image.

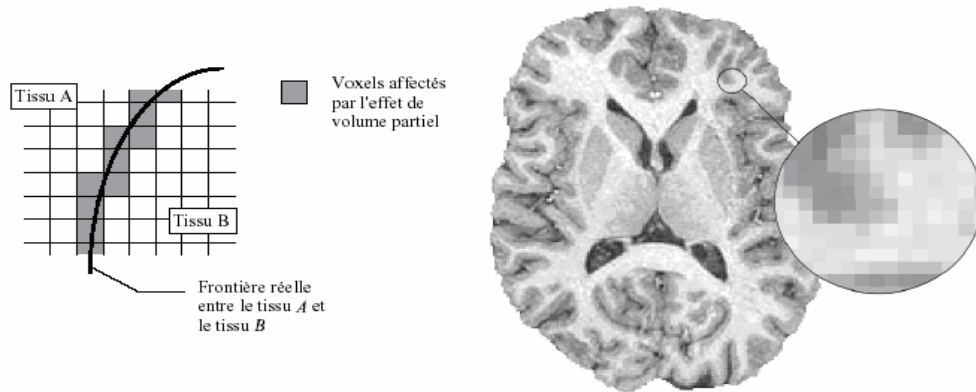


Figure 1.7: L'effet de volume partiel en IRM.

5. Domaines d'application de l'IRM

L'apport d'informations sur la nature des tissus explorés (liquide, solide, dense,...) par une mesure non invasive de certaines de leurs caractéristiques intrinsèques, telles que les temps de relaxation ou la densité de matière, est sans doute à l'origine du succès de l'IRM dans le domaine de l'imagerie biomédicale. C'est encore à l'heure actuelle la technique d'imagerie la plus exploitée dans le diagnostic de lésions internes. Les images anatomiques sont aussi souvent utilisées comme référence dans les domaines de recherche mettant en jeu des modalités d'acquisition différentes. Toutefois, en plus des informations concernant la structure d'un organe, les appareils d'IRM permettent d'obtenir d'autres informations et ce par le simple choix de séquences d'acquisition spécifiques.

Dans ce qui suit, nous allons présenter deux de ces domaines d'application : l'IRM anatomique et l'IRM fonctionnelle.

5.1 L'IRM anatomique

L'imagerie anatomique a pour objectif de représenter certaines structures anatomiques internes en réponse aux diverses excitations appliquées. En effet, la plupart des tissus mous se distinguent en IRM par des caractéristiques statiques et dynamiques propres, représentées dans l'image par des variations de niveaux de gris. Cette imagerie utilise classiquement les séquences d'acquisition 3D à base d'écho de gradient pour obtenir rapidement une pondération de l'image en T_1 . Bien que l'on puisse aussi les utiliser à la production d'images pondérées en T_1 , les séquences à base d'écho d'impulsion RF sont plutôt destinées à fournir les pondérations en T_2 ou en densité de protons (Figure 1.8).

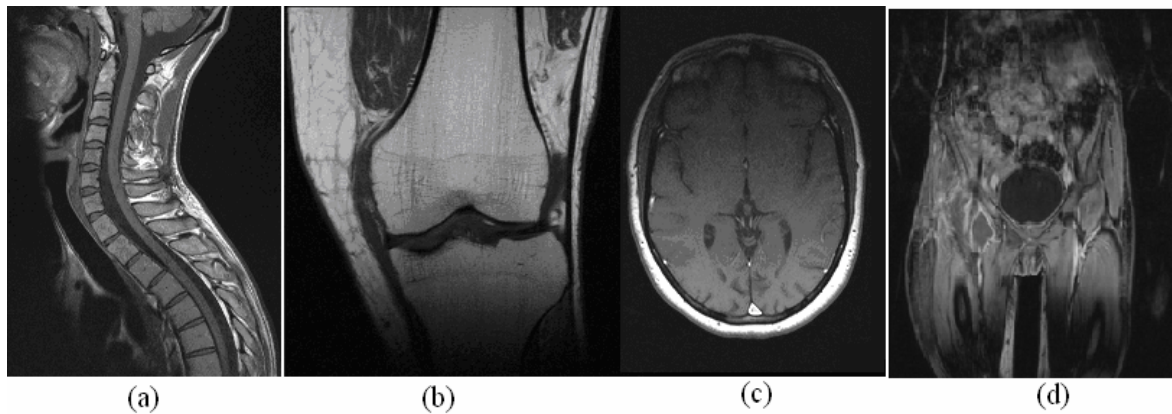


Figure 1.8 : Exemples des images réelles pondérées T_1 , (a) Cervical, (b) Genou, (c) Cerveau, (d) Bassin.

En plus de l'intérêt clinique que peuvent présenter ces images lors d'examens radiologiques, il existe de nombreuses applications à la représentation des structures anatomiques internes. Nous montrons à titre d'exemple sur la Figure 1.9 deux applications réalisées grâce aux outils de traitement d'image. La première application concerne la reconnaissance automatique des sillons du cortex cérébral humain, destinée à faciliter l'interprétation des images anatomiques et l'exploitation des images fonctionnelles. La seconde concerne l'élaboration d'une méthode automatique de segmentation et de classification des principales structures cérébrales internes, et notamment la détermination des matières blanche, grise et du liquide céphalo-rachidien, dont le but principal était de faciliter l'analyse morphométrique des différentes régions du cerveau.

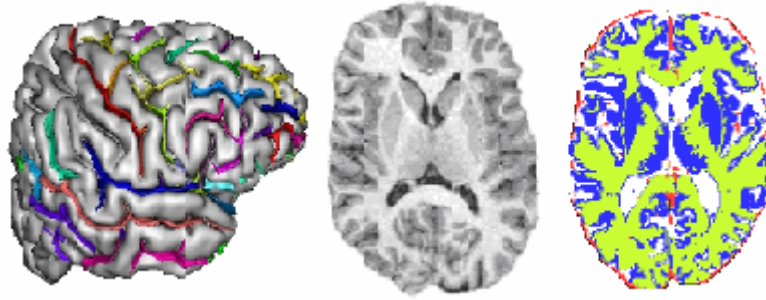


Figure 1.9 : Exemples d'application : reconnaissance des sillons, classification

5.2 L'IRM fonctionnelle

L'IRM fonctionnelle tente d'établir la nature des relations qui existent entre l'activation cérébrale issue d'une tâche spécifique et l'augmentation du signal de résonance magnétique correspondant. Le phénomène d'entrée de fluide, mis en évidence par les séquences d'IRM (Angiographie par Résonance Magnétique), permet de justifier cette augmentation par l'accroissement du flux sanguin localisé dans les régions d'activité cérébrale. Un modèle complémentaire, dénommé effet BOLD ('Blood Oxygenation Level-Dependent'), s'appuie plutôt sur les variations de la réponse magnétique des tissus entre un état de repos et un état d'activation. Ce modèle suppose qu'une activation cérébrale provoque une augmentation de la concentration de molécules fortement oxygénées (l'oxyhémoglobine) dans les régions capillaires.

Dans un état de repos, certains métabolites perturbent localement le champ magnétique à cause de leurs propriétés paramagnétiques. Dans un état d'activation, la variation de susceptibilité magnétique des cellules sanguines, due à l'apport en oxyhémoglobine, induit une réponse magnétique différente de celle de l'état de repos. Il est donc possible, par simple différenciation ou par analyse statistique, de déterminer des zones d'activité fonctionnelle en comparant des images acquises au repos et des images acquises lors de l'exercice d'une tâche spécifique. La localisation de ces zones se fait par superposition avec les images de l'anatomie du même sujet acquise par une séquence classique (Figure 1.10).

En pratique la réponse hémodynamique des tissus est rapide, elle se situe aux alentours de 2 à 3 secondes. Elle nécessite donc l'emploi de séquences d'imagerie ultra-rapide de type écho-planar. La faible différence de signal enregistrée entre deux images oblige de plus à multiplier les acquisitions. Les paradigmes expérimentaux choisis consistent la plupart du temps à accumuler plusieurs séries d'images, successivement acquises au repos et en période

d'activation. Les tâches accomplies par le sujet peuvent être motrices, sensorielles ou même cognitives.

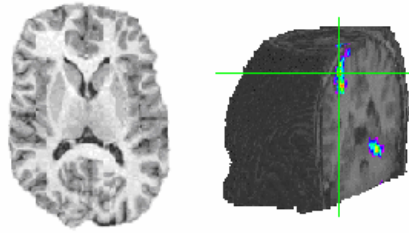


Figure 1.10 : Exemple d'IRM fonctionnelle

6. Conclusion

Les images acquises par résonance magnétique sont altérées par des artefacts comme le bruit et l'effet de volume partiel. Elles sont néanmoins de bonne qualité par rapport à d'autres types d'imagerie médicale, du point de vue de la résolution spatiale et du rapport signal sur bruit. Le rendu anatomique de ces images laisse présager de nombreuses applications dans le domaine de l'imagerie médicale. Pourtant, au vu des informations fournies par l'IRM, l'extraction ou la visualisation des tissus s'avère compliquée sans l'utilisation des méthodes de traitement.

Chapitre 2

Méthodes de segmentation d'images

1. Introduction

La segmentation d'images est le traitement qui consiste à rassembler les pixels d'une image en zones homogènes en utilisant le contenu de celle-ci. Le résultat d'une segmentation est alors une carte qui constitue un résumé de l'image initiale et qui en facilite l'analyse, bien qu'elle ne corresponde pas toujours à une réalité physique. Cette notion d'homogénéité dépend bien sûr des critères qui ont été choisis lors de la construction de la carte. Il existe essentiellement deux familles de méthodes de segmentation : les méthodes par recherche de frontières et les méthodes par recherche de régions.

Les méthodes par recherche de frontières déterminent les contours des zones homogènes en se basant sur les fluctuations de l'intensité de l'image. Par exemple, l'utilisation d'équations aux dérivées partielles permet de construire ces contours en déterminant les changements brusques du gradient. De manière duale, l'approche par recherche de régions consiste à regrouper des pixels connexes présentant des caractéristiques proches, et à leur attribuer une étiquette. L'idée fondamentale de cette approche est de considérer la segmentation d'une image comme l'estimation bayésienne d'un processus aléatoire (à valeurs discrètes) à partir d'un processus observé tout en tenant compte de la dépendance spatiale entre pixels.

2. Synthèse des approches de segmentation d'images

La mise en œuvre des attributs de l'analyse de texture est très utilisée dans les techniques de segmentation dédiées à l'imagerie médicale. En effet, l'IRM permet de visualiser un organe particulier et les méthodes de segmentation qui s'y rattachent concernent la caractérisation du tissu de l'organe en question. Cette caractérisation du tissu est d'autant plus importante qu'elle peut aider le médecin dans son diagnostic et qu'elle permet de quantifier les changements pathologiques de certaines maladies au cours du temps. Le lien entre texture et tissu est envisageable.

Pourtant, en traitement d'images, la définition de la texture est sujette à polémique. Par exemple, l'une est formulée de la manière suivante : *est texture tout ce qui n'est pas contour* [1]. Cette définition permet effectivement de distinguer les deux types d'information contenue dans l'image, l'homogénéité des caractéristiques de la texture au sein d'une même région et les discontinuités qui les séparent. Cependant, elle est incomplète car elle ne tient pas compte du facteur d'échelle d'observation qui modifie l'aspect de la texture. On différencie alors les textures

structurées, caractérisées par une primitive qui se répète spatialement à intervalle régulier et les textures aléatoires dont l'infime détail (le pixel) voit sa luminance évoluer selon une fonction stochastique. Il s'agit respectivement des macro-textures et des micro-textures.

En imagerie médicale, les textures observées dans les images sont de type aléatoire et les techniques de segmentation associées peuvent être répertoriées dans trois groupes : les méthodes statistiques, les méthodes basées sur un modèle et les méthodes de morphologie mathématique.

2.1 Approches statistiques

Ces méthodes utilisent un ou plusieurs paramètres statistiques dont la valeur, propre à chaque texture, permet de discriminer les différentes régions dans l'image. On définit alors l'ordre d'une statistique selon le nombre de pixels mis en jeu dans le calcul de la valeur de ces paramètres.

2.1.1 Paramètres du premier ordre

Les valeurs de ces paramètres sont calculées à partir des niveaux de gris dans une région d'intérêt. Les plus courants sont les moments statistiques de degré 1 et 2, appelés moyenne et variance. Si y_i est le niveau de gris des pixels $i \in \{1 \dots N\}$, où N est le nombre de pixels dans la région, alors la moyenne $\bar{y}$ et la variance σ^2 s'écrivent :

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \quad (2.1)$$

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2 \quad (2.2)$$

Certains moments d'ordre supérieur, comme la dissymétrie (ou skewness) définie par l'expression (2.3) et l'aplatissement (ou kurtosis) défini par l'expression (2.4), sont plus spécifiques à la forme de l'histogramme des niveaux de gris. Ils évaluent la déformation d'un histogramme par rapport à une référence de type fonction gaussienne.

$$s = \frac{1}{N} \frac{\sum_{i=1}^N (y_i - \bar{y})^3}{\sigma^3} \quad (2.3)$$

$$k = \frac{1}{N} \frac{\sum_{i=1}^N (y_i - \bar{y})^4}{\sigma^4} - 3 \quad (2.4)$$

2.1.2 Paramètres du second ordre

Les moments statistiques du premier ordre ne permettent qu'une description partielle de la distribution des niveaux de gris dans une région. Ils ne prennent pas en compte l'arrangement des niveaux de gris dans la région, critère pourtant complémentaire du critère de luminance dans l'analyse de texture aléatoire. Les moments statistiques du second ordre apportent une information supplémentaire.

2.2 Les méthodes basées sur un modèle

2.2.1 Modèle fractal

L'approche fractale a été introduite en analyse d'images. De nombreuses applications sur la caractérisation des textures en découlent [1]. Les fractales sont des fonctions mathématiques qui permettent d'obtenir un modèle de surface ou de contour. Certains processus physiques comme la mesure de la longueur d'une côte peuvent être décrits par des fonctions fractales. Les objets de type fractal ont la propriété fondamentale d'autosimilarité, c'est-à-dire que leurs propriétés statistiques sont préservées quelle que soit l'échelle d'observation.

L'approche fractale se réfère à un attribut de texture qui caractérise un objet par une dimension topologique non entière, appelée dimension fractale. En traitement d'images, la dimension fractale est une mesure qui permet de caractériser le degré d'irrégularité d'une surface ou d'une courbe. Elle est généralement associée à la rugosité d'une texture. D'un point de vue théorique, la dimension fractale d'un objet est définie par une valeur fractionnaire supérieure à sa dimension topologique. Pratiquement, plus la dimension fractale d'une surface est grande, plus cette surface est rugueuse.

2.2.2 L'approche multifractale

L'approche multifractale est intéressante car elle fournit un paramètre local, caractéristique de la dimension fractale d'un objet. Il est alors possible de caractériser une image par l'ensemble des dimensions fractales mesuré sur une partie du signal (ensemble iso- α).

La notion d'approche multifractale peut être décrite de la manière suivante, Soient μ une mesure de probabilité de Borel définie sur l'intervalle $[0,1]$, v_n une suite croissante d'entiers positifs et $I_{i,n}$, une suite d'intervalles de $[0,1]$:

$$I_{i,n} = \left[\frac{i}{v_n}, \frac{i+1}{v_n} \right] \quad (2.5)$$

Avec $i \in \{0, \dots, v_n - 1\}$

La caractérisation de la géométrie d'un ensemble et de la répartition des mesures définie sur cet ensemble nécessite la formulation de la quantité suivante :

$$\tau_n(q) = -\frac{1}{\ln(v_n)} \ln \left(\sum_{i/\mu(I_{i,n})} \mu(I_{i,n})^q \right) \quad (2.6)$$

On dit que μ possède un comportement multifractal si :

$$\lim_{n \rightarrow \infty} \tau_n(q) = \tau(q) \quad (2.7)$$

Pour q appartenant à un intervalle non vide de $\mathbb{R}$, $\tau(q)$ caractérise le comportement global de la mesure quand la taille des intervalles tend vers 0. Elle est reliée à la notion de dimensions généralisées. Soit $I_n(x)$ l'intervalle $I_{i,n}$, contenant le point x . On définit alors E_α comme le sous ensemble des points ayant le même comportement d'échelle, décrit par α :

$$E_\alpha = \left\{ x \in [0,1] / \lim_{n \rightarrow \infty} \frac{\ln(I_n(x))}{\ln(v_n)} = -\alpha \right\} \quad (2.8)$$

α est appelé coefficient de Hölder au point x . Il caractérise l'invariance d'échelle locale de la mesure.

Si α existe au point x alors, lorsque $n \rightarrow \infty$,

$$\mu(I_n(x)) \approx v_n^{-\alpha(x)} \quad (2.9)$$

v_n étant la taille linéaire de l'intervalle centré sur x .

La description multifractale de μ est obtenue en calculant l'ensemble des exposants α possibles et en évaluant la taille du sous ensemble E_α à l'aide de sa dimension de Hausdorff notée $f(\alpha)$.

Cette description $(\alpha, f(\alpha))$ est à la fois locale et globale. Elle est désignée sous le nom de spectre multifractal de μ .

D'un point de vue algorithmique, le calcul du coefficient de Hölder est obtenu en estimant la limite point à point du rapport :

$$\frac{\ln(\mu(I_n(x)))}{\ln\left(\frac{1}{v_n}\right)} \quad (2.10)$$

L'expression (2.9) peut également s'écrire sous la forme :

$$\ln(\mu(I_n(x))) \propto \alpha(x) \ln\left(\frac{1}{v_n}\right) \quad (2.11)$$

En deux dimensions, on recouvre l'image de taille $L \times L$ par un quadrillage délimitant des boîtes de taille $r \times r$. Chaque boîte de ce quadrillage constitue un intervalle dans lequel on effectue la mesure, c'est-à-dire la somme des niveaux de gris μ_{sum} normalisée par la somme y_{sum} des niveaux de gris sur toute l'image. Le coefficient de Hölder peut alors être estimé en chaque point par :

$$\lim_{r \rightarrow 0} \ln\left(\frac{\mu_{\text{sum}} / y_{\text{sum}}}{r / L}\right) \quad (2.12)$$

Toutefois, dans la pratique, on approche cette limite en calculant μ_{sum} pour différentes valeurs de r . Le coefficient de Hölder est alors estimé par régression linéaire. D'autres mesures sont envisageables comme le minimum de singularité ou le maximum de singularité.

2.2.3 Modèles stochastiques

Posé le problème de la segmentation dans un cadre stochastique permet à la fois de modéliser les textures, de modéliser les déformations engendrées par l'acquisition et de disposer d'algorithmes efficaces.

Des représentations stochastiques par champ aléatoire de Markov sont de plus en plus fréquemment utilisées dans le processus de modélisation d'images. La caractéristique principale des champs de Markov est de définir des modèles statistiques globaux à partir de la spécification des propriétés locales d'interactions spatiales, temporelles ou spatio-temporelles entre primitives d'images. Ces interactions locales permettent de décrire statistiquement des ensembles de primitives comme des niveaux de gris, des éléments de contour ou des vecteurs représentant un ensemble de caractéristiques particulières. De par leur généralité, ces modèles trouvent leurs applications dans tous les problèmes nécessitant une connaissance a priori sur les données de l'image [1].

2.3 Les méthodes de morphologie mathématique

2.3.1. Ligne de partage des eaux hiérarchique : algorithme des cascades

La ligne de partage des eaux (LPE) est une connexion par cheminement. Le critère « tous les points de A sont inondés à partir d'un même minimum » étant connectif, la LPE segmente l'espace de définition E en bassins versants connexes par arc et ainsi en un ensemble de composantes connexes ponctuelles qui constituent les lignes de partage. La LPE constitue un opérateur morphologique pour la segmentation.

Typiquement, le processus de segmentation morphologique consiste à appliquer la LPE sur l'image gradient, dérivée de l'image originale. Un tel procédé conduit en pratique à des images très sursegmentées : présence d'un grand nombre de minima dans l'image, et donc d'un grand nombre de bassins versants.

Plusieurs méthodes ont été proposées, fondées sur des graphes qui ont abouti à des algorithmes de segmentation multi-échelle très rapides et efficaces [22].

Un autre algorithme de LPE hiérarchique non paramétrique : l'algorithme de cascades de LPE's, a été proposé par Beucher. Si la LPE est construite sur l'image gradient, cette méthode repose sur l'hypothèse que les contours les plus forts correspondent aux objets à segmenter. Concrètement à chaque étape de cette segmentation hiérarchique, on cherche à supprimer les contours les plus faibles entourés de contours de valeurs plus élevées, en fusionnant les bassins

versants des contours supprimés. Un tel processus peut être itéré de façon à supprimer, dans une nouvelle passe, l'ensemble des contours devenus à leur tour les plus faibles.

L'implémentation la plus simple de l'algorithme de cascades est basée sur la reconstruction de la fonction gradient de l'image mosaïque avec sa LPE. L'image à niveaux de gris à segmenter et son gradient sont respectivement f et g :

$$E \rightarrow \tau = \{t_{min}, t_{min} + 1, \dots, t_{max}\}$$

Soit $LPE(g)$ la LPE du gradient g . L'image mosaïque m est obtenue comme la valeur moyenne de la fonction f dans chaque bassin versant de $LPE(g)$. Puis, on calcule le gradient de la mosaïque, qu'on notera g_1 . L'algorithme de cascades est comme suit :

1. Calculer la LPE de la fonction g_1 : $LPE(g_1)$.
2. Définir une nouvelle fonction numérique, h :

$$h(x) = \begin{cases} g_1(x) & \text{si } x \in LPE(g_1) \\ t_{max} & \text{si } x \in LPE^c(g_1) \end{cases}$$

Avec $h \geq g_1$.

3. Reconstruire g_1 à partir de h par érosion géodésique $g_1^* = \varepsilon_{g_1}^{rec}(h)$. La reconstruction va remplir chaque bassin versant avec un plateau de hauteur égale à la profondeur minimale des lignes de la LPE qui entourent le bassin versant. Par conséquent, s'il y a un bassin versant adjacent avec une hauteur inférieure à celle définie, alors le plateau généré ne sera pas un minimum.
4. Calculer la LPE de la fonction g_1^* : les bassins versants sont associés aux minima significatifs.
5. Construire une nouvelle mosaïque à partir de cette dernière LPE, m_1 , sur laquelle on calcule le gradient g_2 et ensuite on peut itérer l'algorithme pour obtenir la hiérarchie.

2.3.2. Connexions par sauts

La LPE peut être considérée comme un algorithme de croissance de régions : à partir des minima (ou des marqueurs sélectionnés), une propagation pixel à pixel est faite pour définir les bassins versants jusqu'à ce que l'espace soit totalement partitionné. Il existe d'autres méthodes

de croissance de régions qui peuvent être considérées comme appartenant à la morphologie mathématique. Le principe de base est que le processus d'agrégation des germes est un critère connectif, et pour cela, dans la partition maximale finale, les points $\mathbf{x} \in E$ doivent être associés à un seul germe (car les régions dont les points sont agrégés à plus d'un germe seront pulvérisées en singletons).

La connexion par saut est une méthode de segmentation par croissance de régions proposée par Serra [22] et fondée sur des variations bornées d'amplitude. Elle est réalisée, en fixant les régions des variations aux minima ou maxima de la fonction f (intérêt des extrema pour marquer les structures distinctes de l'image).

La segmentation selon d'un critère connectif, engendre la connexion par sauts. La procédure alternative à partir des maxima se construit par dualité, et l'on peut combiner les deux procédures dans une démarche symétrique, où la remontée s'arrête quand elle rencontre une descente, et où les zones à cheval entre montée et descente sont subdivisées.

Etant donné qu'elle dépend d'un paramètre croissant, la connexion par sauts engendre une pyramide. Par ailleurs, elle est aussi une granulométrie lorsqu'on restreint les valeurs du paramètre aux progressions géométriques.

En pratique la segmentation par sauts s'avère une bonne méthode de partition : peu de zones ponctuelles et de classes visuellement significatives.

2.3.3. Zones plates et quasi-plates

Les zones plates d'une fonction f sont les composantes connexes avec une valeur constante. Si on obtient les zones plates d'une image brute, la partition obtenue n'est guère intéressante à cause de la sur-segmentation. Il existe quelques alternatives à suivre; par exemple, filtrer préalablement l'image f pour obtenir une simplification des détails. On doit utiliser un filtre connexe, typiquement un nivellement dont l'une des propriétés principales est l'élargissement des zones plates. Si on considère maintenant une séquence de nivellements avec des marqueurs de plus en plus forts (une pyramide de nivellement), les zones plates associées à la famille de fonctions successives constituent une hiérarchie ou pyramide.

2.3.4. Fusion de région dans un treillis de partitions

Lorsqu'on utilise des critères connectifs de segmentation par sauts et de segmentation par zones quasi-plates, les régions homogènes s'élargissent. Quelques hypothèses ont été proposées, elles considèrent que les contours des objets se trouvent quelque part dans les zones de transition et chaque objet est marqué par une grande zone plate. En établissant un seuil sur la taille des

zones plates et en ne gardant que celles de plus grande surface nous obtenons une image de marqueurs indiquant la position des régions homogènes, qui peuvent être utilisées pour la LPE.

Si on ne vise pas à définir des marqueurs, mais plutôt à obtenir une partition significative de la fonction, il faut envisager d'autres méthodes pour pallier la sur-segmentation. Un algorithme de fusion de régions peut être utilisé pour cela. Le principe consiste en la fusion ordonnée des régions initiales (qui peuvent être les pixels de l'image) selon un certain critère de similarité. La principale difficulté dans la fusion de régions est de trouver un bon critère de similarité ainsi qu'une mise en œuvre efficace de l'algorithme.

3. Méthodes du traitement de l'incertitude et de l'imprécision

Comme décrit dans [11], le modèle de traitement de l'incertitude le plus connu est celui basé sur la théorie des probabilités. Il permet de traiter des incertitudes liées à des fréquences d'apparition d'événements et a été pendant longtemps le seul moyen de modéliser l'incertitude. Son avantage est sa rigueur mathématique et la grande exploration de la méthode qui en a été faite par les statisticiens. Néanmoins, la notion de probabilité n'est pas très bien adaptée à traiter d'autres aspects du raisonnement humain, comme l'imprécision et la confiance. La notion de degré de confiance a été introduite dans la théorie de l'évidence de Dempster et Shafer (Srivastava et Shafer se référant à l'ouvrage de G.Shafer). Ils n'imposent plus l'additivité de la confiance comme pour les probabilités et peuvent traiter des données sans attribuer une valeur de probabilité à chaque élément de l'univers. En ce qui concerne la modélisation de l'imprécision, la notion d'intervalle d'erreur (utilisée par exemple en physique pour décrire l'erreur de mesure) est la plus ancienne. C'est par l'introduction des sous ensembles flous (nous reparlerons des sous-ensembles flous plus loin) en 1965 que Zadeh rapproche la notion d'intervalle et le raisonnement humain en considérant plutôt des frontières graduelles que des frontières abruptes.

En imagerie par résonance magnétique, nous ne pouvons pas traduire, « chaque individu appartient à une classe » par « chaque pixel appartient à une structure », puisque, comme nous l'avons vu dans le premier chapitre, les pixels de mélange (région intermédiaire) sont constitués de plusieurs structures.

Les données sont alors imprécises. Une question que nous pouvons nous poser pour chaque pixel de mélange est de savoir quelle est la structure la plus représentée dans le mélange. La généralisation aux pixels purs est naturelle; classifier un pixel peut donc signifier trouver la structure majoritaire au sein d'un mélange de pixels.

4. Les méthodes de segmentation non supervisée

4.1 Segmentation par les K-moyennes

L'algorithme des K-moyennes, est un cas particulier d'une méthode de classification automatique beaucoup plus générale : les nuées dynamiques de Diday.

4.1.1 Nuées dynamiques

La méthode des nuées dynamiques consiste à considérer que les classes formées d'individus ont un comportement de nuées groupées autour de noyaux.

Considérons sur $\mathbf{E}$ la mesure $f(x, w_i)$ de dissimilarité (ou dissemblance) entre l'individu x et la classe w_i cette mesure, à valeurs dans $\mathbf{R}^+$, est calculée à partir du noyau de w_i . On se propose de classer $\mathbf{X}$ de telle façon que la somme des mesures entre les individus et leur classe soit la plus faible possible. La fonction objectif à minimiser s'écrit alors :

$$g_f(\Omega) = \sum_{w_i \in \Omega} \sum_{x_k \in w_i} f(x_k, w_i) \quad (2.13)$$

Afin de minimiser la fonction *objectif*, l'algorithme le plus fréquemment utilisé est le suivant :

1. $it = 0$
2. choisir une classification initiale de $\mathbf{X}$: $\left\{ w_I^{(it)} \dots w_c^{(it)} \right\}$
3. $it = it + 1$
4. actualiser chaque classe suivant :

$$w_i^{(it)} = \left\{ x_k \in X, f(x_k, w_i^{(it-1)}) \leq f(x_k, w_j^{(it-1)}) \forall j \neq i \right\}$$

5. Si $\exists i, w_i^{(it)} \neq w_i^{(it-1)}$ alors retourner à l'étape 3.

Lors d'une itération, on affecte à chaque individu la classe dont la caractérisation est la plus proche des caractéristiques de cet individu. Les classes ainsi redessinées voient leur forme évoluer progressivement et leurs caractéristiques changer dynamiquement.

A l'issue de la classification, les points de $\mathbf{E}$ appartenant à la frontière entre les classes d'indices respectifs i et j vérifient :

$$f(x, w_i) = f(x, w_j) \quad (2.14)$$

Dans la méthode des nuées dynamiques, les caractéristiques de chaque classe dépendent du noyau qui décrit cette classe ; f est alors une fonction de dissimilarité entre un individu et le noyau d'une classe. Diday propose alors plusieurs exemples de mesure de f , chacun garantissant la convergence de l'algorithme de descente que nous avons présenté précédemment vers un minimum (local).

L'exemple le plus simple réduit le noyau d'une classe à un point de E , centre de gravité de la classe :

$$v_i = \frac{1}{\text{card } w_i} \sum_{x_k \in w_i} x_k \quad (2.15)$$

La mesure peut se construire à l'aide d'une fonction distance d suivant $f(x, w_i) = d(x, v_i)$. Cette version des nuées dynamiques s'appelle méthode des k-moyennes.

Par extension, une classe w_i peut être représentée par un noyau constitué d'un ensemble de plusieurs points $\{v_i^1, \dots, v_i^L\}$. Une façon de les déterminer est de minimiser :

$$\sum_{x_k \in w_i} \inf_{l=1 \dots L} d(x_k, v_i^l)$$

La nouvelle mesure est alors par exemple :

$$f(x, w_i) = \frac{1}{L} \sum_{l=1 \dots L} c_l d(x_k, v_i^l)$$

Avec c_l le coût associé au $i^{\text{ème}}$ point du noyau ; ou encore :

$$f(x, w_i) = \frac{1}{L} \sum_{l=1 \dots L} d^2(x_k, v_i^l)$$

La description par noyau réparti sur plusieurs points permet de mieux s'adapter à des classes de forme plus complexe. Un noyau constitué autour de l'axe principal d'inertie Δ_i de la classe w_i pourrait donner comme mesure :

$$f(x, w_i) = \inf_{x' \in \Delta_i} d(x, x') \quad (2.16)$$

4.1.2 Mise en pratique des k-moyennes

Les individus à classer étant décrits par des vecteurs de l'espace des caractéristiques, on munit généralement ce dernier de distances métriques. En choisissant une métrique pour chacune des classes, la distance d'un individu x_k au centre v_i de la classe w_i s'écrit alors naturellement :

$$d(x, w_i) = d_i(x, v_i) \quad (2.17)$$

En pratique, on utilise fréquemment la distance euclidienne. Lorsque k augmente, on constate empiriquement que la partition de E a tendance à donner aux classes des formes de cercles et des surfaces comparables.

La méthode des K-moyennes donne une généralisation efficace du seuillage optimal. L'image brute Y est à niveaux de gris appartenant à l'intervalle $[0, 255]$. On choisit arbitrairement ou empiriquement un nombre de classes K :

Principe : classer chaque pixel de l'image Y_s selon son niveau de gris dans l'une des K classes et l'image segmentée X est à valeur dans $(1, \dots, K)$.

On peut résumer la méthode des K-moyennes sur trois étapes essentielles comme suit :

4.1.2.1 Initialisation

On divise l'intervalle $I = \{0, \dots, 255\}$ en K intervalles selon divers critères :

- intervalles de même longueur, notés $(I_k)_{k \in \dots, K}$
- intervalles déduits selon l'analyse de l'histogramme.

Chaque classe k est associée à une moyenne, notée m_k . L'image n'est pas encore segmentée, on initialise :

$$M_k = \text{milieu}(I_k) \quad (2.18)$$

4.1.2.2 Classification

Pour chaque élément s de X on a :

$$X_s = \underset{k \in \{1, \dots, K\}}{\text{Arg min}} (\text{dist}(Y_s, m_k)) \quad (2.19)$$

4.1.2.3 Mise à jour des moyennes

Pour chaque classe K , on calcule

$$m_k = \frac{1}{|\{s | X_s = m_k\}|} \sum_{s | X_s = m_k} Y_s \quad (2.20)$$

Pour faire une segmentation par la méthode des K-moyennes on choisit un nombre K de classes, la figure (2.1) ci-dessous montre un exemple de segmentation non supervisée d'une image IRM à $K = 5$ classes.

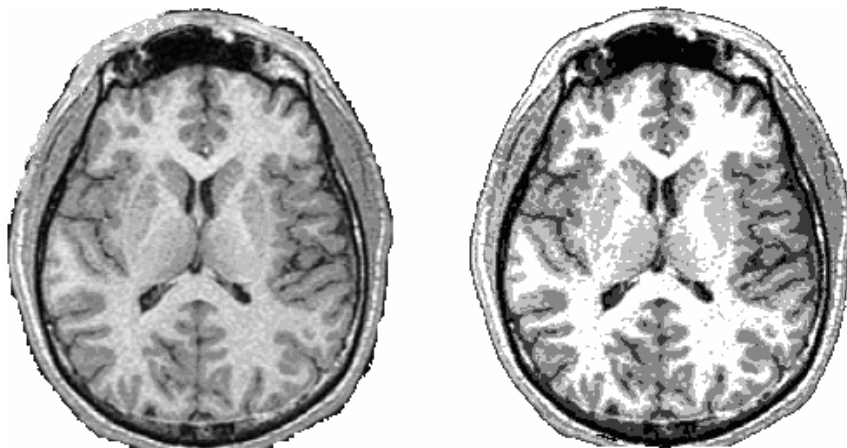


Figure 2.1 : segmentation d'une image IRM cérébrale pondérée T1 à 5 classes

4.1.2.4. Discussion

- On arrête les itérations lorsque les moyennes n'évoluent plus.
- Algorithme simple et efficace : bonne initialisation pour des algorithmes plus sophistiqués (modèles markoviens par exemple).
- Algorithme non supervisé et robuste.
- Fonctionne uniquement sur la distribution fréquentielle des valeurs de niveaux de gris : aucune information spatiale.
- Classes peu régulières, non connexes.

4.2 Modèles flous en analyse d'images

En traitement d'images, les méthodes de classification floue modélisent le degré d'appartenance d'un pixel à chaque classe, sous l'hypothèse que ce pixel puisse ne pas appartenir à une classe unique. Cette affectation multiple caractérise les imprécisions de mesure sur les observations. Une approche particulièrement sollicitée en analyse d'images repose sur la théorie des ensembles flous qui représente un cadre théorique rigoureux dans lequel le traiteur d'images peut exploiter toutes les richesses de l'information pour une classification automatique. Une classification floue basée sur l'algorithme des *c*-moyennes floues (*fuzzy c -means*) a été appliquée à la segmentation des images IRM. Une telle classification permet de caractériser l'effet de volume partiel dans les images IRM (voir définition dans chapitre1).

Les algorithmes de classification ont la tâche d'associer un pixel à une classe unique. Un algorithme de classification floue a pour objectif de déterminer les degrés d'appartenance u_{ik} de l'individu x_k dans chacune des classes w_i . Si c classes sont recherchées pour n individus, l'ensemble des degrés d'appartenance sont stockés dans la matrice de la *c*-partition floue U :

$$U = [u_{ik}]_{i=1 \dots c, k=1 \dots n} \quad (2.21)$$

Trois contraintes sont définies pour la *c*-partition floue. La première impose que les valeurs des u_{ik} soient comprises entre 0 et 1. La deuxième impose que l'appartenance d'un individu soit répartie sur l'ensemble des classes, sans exclure l'existence de degrés d'appartenance nuls. Enfin, la troisième exprime que la classification ne produise pas de classe vide afin que tous les pixels soient caractérisés. Ces contraintes s'écrivent :

$$0 \leq u_{ik} \leq 1 \quad \forall i \in [1; c], \forall k \in [1, n]$$

$$\sum_{i=1}^c u_{ik} = 1 \quad \forall k \in [1; n] \quad (2.22)$$

$$0 < \sum_{k=1}^n u_{ik} < n \quad \forall i \in [1; c]$$

Les algorithmes des c -moyennes floues, adaptés au formalisme flou, ont le rôle de minimiser une fonction objectif J , au sens des moindres carrés, définie comme :

$$J(U, V) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \cdot d^2(x_k, v_i) \quad (2.23)$$

Où m , plus grand que 1, est le facteur de flou et $d^2(x_k, v_i)$ est relative à une métrique définie entre le vecteur x_k et un centroïde v_i . L'algorithme des c -moyennes floues cherche alors à déterminer l'ensemble $V = (v_1, \dots, v_c)$ ainsi que la matrice U de c -partition floue qui minimisent J . La principale difficulté de la recherche des variables optimales provient de la non-convexité de la fonction objectif J qui possède des minima locaux dans lesquels s'échouent les procédés classiques de descente.

Pratiquement, m est choisi dans l'intervalle $[1.5, 2.5]$. Sa présence assure le réglage de la force du flou. Plus m est élevé et moins les individus ayant une forte appartenance à une classe particulière sont discriminés par rapport à ceux qui sont partagés entre plusieurs classes.

La minimisation de J sous les contraintes 2.14 est obtenue par un procédé itératif. L'initialisation de l'algorithme est effectuée en remplissant la matrice U par des valeurs aléatoires vérifiant les conditions 2.14. A chaque itération, l'algorithme met à jour la matrice U et les centroïdes v_i en calculant les valeurs actualisées des degrés d'appartenance par :

$$u_{ik}^{new} = \frac{(d(x_k; v_i))^{\frac{2}{1-m}}}{\sum_{j=1}^c (d(x_k; v_j))^{\frac{2}{1-m}}} \quad (2.24)$$

Et les valeurs actualisées des centres de classe par :

$$v_i^{new} = \frac{\sum_{k=1}^n u_{ik}^m \cdot x_k}{\sum_{k=1}^n u_{ik}^m} \quad (2.25)$$

La condition d'arrêt est vérifiée lorsque la comparaison entre les centres de classe v_i et v_i^{new} montre que l'évolution de ces valeurs à chaque itération n'est plus significative. On remarquera qu'aucune contrainte sur le voisinage n'est imposée à la minimisation du critère défini par le terme J .

Enfin, une ultime étape est nécessaire lorsque le résultat souhaité est une classification non floue. On parle alors de « défuzzification ». Lorsque l'on ne souhaite ni mettre en évidence les voxels où les degrés d'appartenance sont approximativement les mêmes pour chaque classe, ni faire intervenir d'autres informations avant de réaliser la défuzzification, une façon naturelle de procéder est de considérer que la classe finale d'un pixel est celle pour laquelle le degré d'appartenance est maximal :

$$\forall k \in [1; n] \quad x_k \in w_i \Leftrightarrow u_{jk} < u_{ik} \quad \forall i \neq j \in [1; c] \quad (2.26)$$

L'algorithme des c -moyennes floues repousse l'étape de décision de la défuzzification à la fin de la convergence.

Une critique souvent formulée à la classification floue provient de la forme des fonctions d'appartenance traçant l'évolution du degré d'appartenance en fonction du niveau de gris. On s'aperçoit, dans les schémas classiques de convergence, que la forme des fonctions

d'appartenance, associée au facteur de flou m , n'est pas toujours satisfaisante car la propriété de décroissance du degré d'appartenance lorsque le niveau de gris s'éloigne de la moyenne des intensités dans la classe n'est pas vérifiée. Des propositions ont été faites pour remédier à ce défaut en imposant un modèle de fonction d'appartenance ou en changeant les contraintes de minimisation.

L'algorithme des c -moyennes flous fait également l'hypothèse que la décision sur l'affectation d'un pixel à une classe peut être prise uniquement à partir du niveau de gris d'un pixel, indépendamment de son voisinage. Contrairement aux modèles markoviens (champs de Markov Cachée) incluant un terme de régularisation, l'homogénéité spatiale des régions n'est pas assurée et la présence de bruit et de volumes partiels nuisent à la qualité de la segmentation générée lors de la défuzzification.

5. Conclusion

Avant toute conception de système d'analyse d'image, il est nécessaire d'étudier le pourquoi et le comment, c'est-à-dire le contexte de travail, les informations disponibles (les connaissances a priori), les restrictions, etc.

Nous avons présenté dans ce chapitre une synthèse des méthodes sur la segmentation d'images, tout en nous focalisant sur celles utilisées en imagerie médicale. Nous avons commencé notre étude par la méthode de classification dure et celle floue. Les deux étant les plus appropriées à nos exigences. Nous avons rencontré la méthode des K-moyennes et la FCM dans plusieurs travaux relatifs à la segmentation d'images médicales ; parfois elle a été utilisée comme une étape de pré segmentation.

Chapitre 3

Chaîne de Markov Cachée CMC

1. Introduction

Les méthodes statistiques de classification basées sur les modélisations markoviennes (champ, chaîne ou arbre) se sont montrées très robustes [30]. Ce succès est principalement dû au fait que, lorsque le processus caché X (représentant l'image segmentée) peut être représenté par un processus de Markov, il est alors possible de reconstruire X à partir du processus observé Y (représentant l'image bruitée) à l'aide de différents critères bayésiens de classification tels que le MAP (« Maximum A Posteriori ») ou le MPM (« Marginal Posterior Mode »).

L'essor des modélisations markoviennes en traitement d'images date de l'introduction dans les années 80 des champs de Markov cachés (« Hidden Markov Random Fields », HMRF) [13]. Les HMRF se sont distingués des méthodes statistiques classiquement utilisées (méthodes aveugles et contextuelles) par la prise en compte de toute l'information disponible de l'image bruitée dans le processus de segmentation. Le modèle des HMRF, qui suppose que le processus caché X peut être représenté par un champ de Markov (processus 2D), nécessite le recours à des méthodes itératives du type de l'échantillonneur de Gibbs pour estimer la loi du champ ; de plus, le calcul des solutions bayésiennes n'étant pas explicite, ces dernières sont approchées grâce à la mise en oeuvre des méthodes de Monte-Carlo (MPM) et d'algorithmes itératifs type recuit simulé (MAP). À titre d'exemple, le modèle des HMRF a permis de répondre à de nombreux problèmes de classification d'images de modalités variées, segmentation d'images de télédétection, d'images sonar [13] (et les références qu'elle contient), d'images médicales. Les HMRF ont même trouvé des applications en vision par ordinateur et en imagerie industrielle.

Dans les années 90, une autre solution, proposée par Ben Miloud [7], a consisté à modéliser le processus caché X par une chaîne de Markov (processus 1D). Cette alternative nécessite alors l'utilisation d'un parcours (par exemple le parcours de Hilbert-Peano) pour transformer l'image 2D en un signal 1D. Dans ce contexte, l'estimation des paramètres du modèle avec différents algorithmes (EM, SEM, ICE) permet d'exprimer la loi de X , ainsi que les solutions bayésiennes, de façon explicite, sans mise en oeuvre de méthodes d'approximation. De façon générale, le modèle des CMC a permis de traiter de nombreux problèmes de traitement de signal (cf. [31] et les références qu'elle contient) ; par exemple en traitement de la parole [18], en génétique, en reconnaissance de l'écriture manuscrite, en robotique, en reconnaissance de gestes, en traitement de signal bio-médical [32]. En ce qui concerne la segmentation d'images, les CMC ont aussi été utilisées avec succès dans divers contextes [33], comme en segmentation d'images

de télédétection [4, 6, 7, 35], d'images sonar, ou même d'images vidéo [7]. Dans le contexte de la segmentation d'images, les modélisations basées sur les chaînes de Markov constituent une alternative intéressante au modèle des champs de Markov cachés [35], même si ces derniers offrent une modélisation précise et plus intuitive des relations spatiales dans l'image. Le gain en temps de calcul est généralement très significatif, principalement en raison de la complexité de l'étape d'estimation de paramètres des HMRF, ce qui confère aux CMC un réel intérêt pour les problèmes de segmentation et d'analyse d'image.

L'objectif de ce chapitre est de présenter de façon détaillée le principe de la segmentation non supervisée d'images basée sur le modèle des chaînes de Markov cachées. Les principes fondamentaux nécessaires à la compréhension des nouveaux modèles présentés dans les chapitres suivants y sont entièrement exposés. Nous débutons ce chapitre par la présentation du parcours de Hilbert-Peano utilisé pour transformer l'image 2D en un vecteur 1D, ainsi que la modélisation statistique utilisée pour représenter les images. La section suivante est consacrée au principe de la classification bayésienne dans un cadre général, et deux critères de classification bayésienne couramment utilisés, le MAP et le MPM y sont rappelés. Par la suite, nous présentons le principe général des chaînes de Markov cachées, la représentation statistique d'une chaîne de Markov et l'expression de la loi du couple (X, Y) dans le contexte des CMC. La question de l'estimation des paramètres, permettant de rendre le traitement non supervisé, y est ensuite abordée pour laquelle nous proposons d'utiliser l'algorithme ICE dont nous expliquerons le principe. A la fin de ce chapitre, nous présentons les solutions, dans le contexte des CMC, des deux critères bayésiens présentés.

2. Modélisation d'image

En mathématique une image traditionnelle peut être vue comme une application. Par exemple une image noir et blanc (sur 8 bits) de taille 256*256 serait une fonction :

$$f : \{0, \dots, 255\}^2 \rightarrow \{0, \dots, 255\} \quad (3.1)$$

Ainsi la valeur $f(x, y)$ représente la valeur du pixel de coordonnées (x, y) . Cette modélisation peut se généraliser pour représenter des images de dimension 3, des images en couleurs ou des vidéos. Dans cette section l'image est représentée par une grille $N \times M$ (où N et M sont la largeur et la hauteur de l'image). L'objet de la segmentation est alors de trouver une

partition de cette grille, c'est à dire de pouvoir associer un nom (ou étiquette) à chaque pixel. Dans un souci de généralité, par la suite nous nommons S la grille et s un site (pixel) de cette grille. Ainsi $s \in S_k$ signifie que le pixel s appartient à la région k . En imagerie médicale nous choisissons l'ensemble des étiquettes : ce sont les structures cérébrales par exemple.

$$S = \bigcup_{0 \leq i < n} S_i \text{ (avec } n \in \mathbb{N}) \quad (3.2)$$

Pour travailler avec ces sites nous introduisons deux variables aléatoires :

$Y : S \rightarrow \{0, \dots, 255\}$ et $X : S \rightarrow N$ où N est l'ensemble des nombres naturel. Y représente grossièrement les valeurs des pixels de l'image et X les étiquettes associées à ces pixels. L'image proprement dite n'est alors qu'une réalisation de Y et le résultat de la segmentation est la réalisation de X la plus probable. On appelle alors une *carte des étiquettes* l'ensemble $\{x_s, s \in S\}$ qui associe à chaque pixel s une région k . Cela se traduit par :

$$x_s = k \Leftrightarrow s \in S_k \text{ (avec } s \in S) \quad (3.3)$$

3. Parcours de Hilbert-Peano

3.1. Construction du Parcours de Hilbert-Peano

Le parcours de Hilbert-Peano a été découvert par le grand mathématicien allemand David Hilbert en 1891. Ce parcours appartient à la famille des « Space Filling Curves » (SFCs), qui incluent les parcours en zigzag, le balayage en Z et les parcours standards (ligne par ligne, colonne par colonne). Les SFCs ont été employées dans une grande variété d'applications de traitement d'image telles que la compression, le tramage demi-tons (« half-toning »), la reconnaissance de formes et l'analyse de texture. Tous ces parcours peuvent être employés pour convertir un ensemble 2D de pixels en séquence 1D.

Néanmoins, le parcours de Hilbert-Peano fournit une bonne exploitation du voisinage 2D, essentiellement grâce à la contrainte qu'il respecte : « tout pixel de la chaîne possède comme voisins, deux pixels voisins sur l'image ». Cette propriété, ainsi que l'aspect pseudo-aléatoire des changements de directions, impliquent que le parcours de Hilbert-Peano fonctionne bien (du point de vue statistique) pour de nombreuses familles d'images. Le parcours de Hilbert est la

limite d'un ordre H_0, H_1, H_2 des courbes, dont certaines sont montrées dans la figure 3.1. Elles sont définies par ce qui suit:

- H_0 est un seul point.
- H_1 se compose de quatre copies (de point) de H_0 liées aux trois segments droits de la longueur h perpendiculairement entre eux. Quatre orientations de cette courbe 1, 2, 3, et 4 marquées, sont montrées dans la figure 3.1a.
- La prochaine courbe, H_2 dans l'ordre est construite en reliant quatre copies de différentes orientations de H_1 avec trois segments droits de la longueur $h/2$ (montré dans la figure 3.1b). Il y a quatre orientations possibles de H_2 ; celle montrée sur la figure représente l'orientation #2 dont le parcours est construit avec des orientations 1223 de H_1 reliées par des segments.

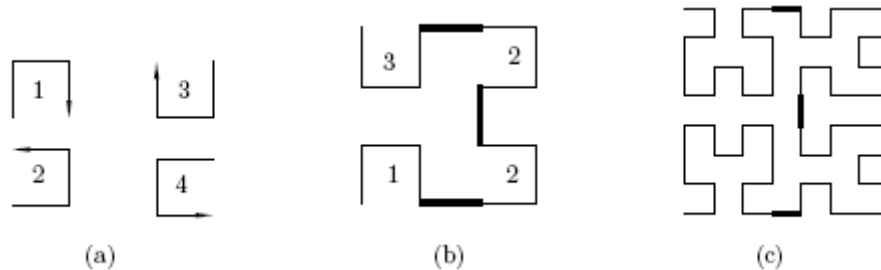


Figure 3.1 : Parcours de Hilbert, Ordre 1 (a), 2 (b) et 3 (c)

La figure 3.2, présente la construction du parcours de Hilbert-Peano pour une image 8×8 .

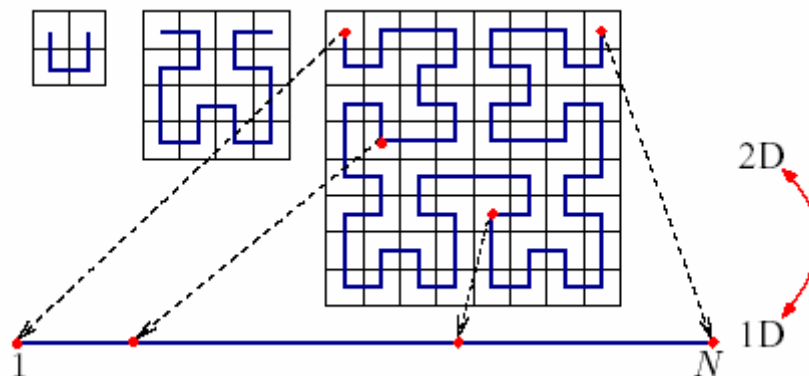


Figure 3.2 : Construction du parcours de Hilbert pour une image 8×8 (initialisation, étape intermédiaire et parcours final).

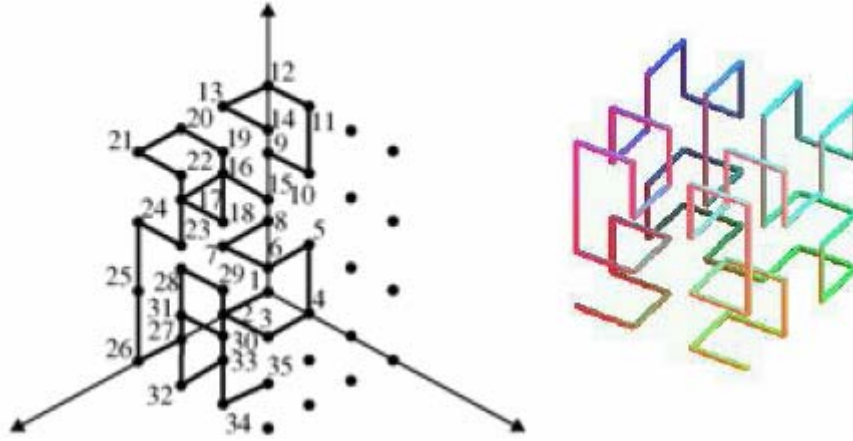


Figure 3.3 : Construction du parcours de Hilbert pour une image 3 D.

3.2. Contraintes

On remarque que le parcours de Hilbert ne peut être utilisé que pour une image carrée de taille $2^n \times 2^n$ où n est l'ordre du parcours de Hilbert, par exemple pour $n = 8$ on a une image de taille 256×256 . Dans notre travail on propose de résoudre le problème de la taille de l'image par une normalisation qui consiste à rendre l'image carrée (par exemple une image de taille 256×200 devient de taille 256×256).

La première étape de tous les algorithmes de segmentation d'images basés sur le modèle des CMC consiste à transformer l'image bruitée en un vecteur. Tous les traitements s'effectuent sur ce vecteur, puis, une fois la segmentation de la chaîne achevée, l'image thématique est finalement reconstruite à l'aide du parcours de Hilbert-Peano inverse. Dans la suite, nous utiliserons donc une structure 1D pour représenter une image dont la séquence est issue de sa transformation par le parcours de Hilbert-Peano.

4. Décision bayésienne

Nous allons évoquer différents critères bayésiens de classification permettant de déduire la réalisation inobservable $\hat{x}$ du processus $\mathbf{X}$ connaissant $\mathbf{y}$. Chaque estimateur est associé à une fonction de coût $L(x, \hat{x})$, mesurant la perte entre $\mathbf{x}$ et son estimé $\hat{x}$. Dans le cadre de la classification bayésienne, on recherche la règle de décision optimale $\hat{x}_{\text{opt}}$ qui minimise la perte moyenne, c'est-à-dire l'espérance conditionnelle de la fonction de coût, telle que :

$$\hat{x}_{\text{opt}}(y) = \arg \min E[L(x, \hat{x}) | Y = y] \quad \hat{x} \in \Omega^N \quad (3.4)$$

Plusieurs estimateurs sont possibles en fonction du type d'approche adoptée : locale, contextuelle ou globale. Nous avons retenu les approches globales, qui sont théoriquement optimales puisqu'elles exploitent toute l'information contenue dans l'image. Dans les deux paragraphes suivants, nous décrivons brièvement les deux critères bayésiens de la classification utilisés dans notre travail, à savoir le MAP et le MPM. Leur mise en œuvre dans la modélisation des CMC sera détaillée par la suite.

4.1 Maximum A Posteriori : MAP

Le principe du MAP consiste à pénaliser de la même façon toutes les configurations estimées qui sont différentes de x . La fonction de coût correspondante est donnée par :

$$L(x, \hat{x}) = 1 - \delta(x, \hat{x}) \quad \text{avec } \delta(x, \hat{x}) = \begin{cases} 1 & \text{si } x = \hat{x} \\ 0 & \text{sin on} \end{cases}$$

On peut montrer que l'estimateur basé sur cette fonction de coût est :

$$\hat{x}_{\text{MAP}}(y) = \arg \max_{x \in \Omega^N} (P(X = x | Y = y)) \quad (3.5)$$

4.2 Marginal Posterior Mode : MPM

La fonction de coût de l'estimateur du MPM est proportionnelle au nombre de pixels mal classés :

$$L(x, \hat{x}) = \sum_{n=1}^N (1 - \delta(x_n, \hat{x}_n))$$

On en déduit donc, pour chaque n , l'estimateur suivant :

$$\hat{x}_{n\text{MPM}}(y) = \arg \max_{x_n \in \Omega} (P(X = x | Y = y)) \quad (3.6)$$

Dans le cadre d'une approche globale, nous verrons que les chaînes de Markov permettent le calcul explicite des solutions bayésiennes de classification, contrairement au modèle des HMRF Champs de Markov cachés (Hidden Markov Random Fields) qui fait appel à des méthodes d'approximation de type Monte-Carlo (MPM) ou recuit simulé (MAP).

Dans le cadre non supervisé, nous cherchons donc à estimer la distribution *a posteriori* de X connaissant $Y = y$. La règle de Bayès permet d'exprimer cette distribution en fonction de la distribution conjointe de X et Y selon :

$$P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \quad (3.7)$$

5. Chaînes de Markov Cachées CMC

5.1 Modèle de chaîne de Markov cachée

Dans le contexte des chaînes de Markov cachées, on considère que le processus aléatoire X peut être représenté par une chaîne de Markov. L'image thématique inobservable recherchée x , qui apparaît donc comme une réalisation d'une chaîne de Markov $X = X_{1 \rightarrow N}$, est par conséquent cachée derrière les données bruitées y , et la modélisation globale retenue pour $(X; Y)$ est celle d'une chaîne de Markov cachée. Le modèle Chaîne de Markov caché CMC est caractérisé par une matrice d'observation $\mathbf{O}$, la matrice de la loi initiale $\boldsymbol{\pi}$ et la matrice de transition $\mathbf{T}$, donc on peut noter le modèle par $M = \{\mathbf{O}, \boldsymbol{\pi}, \mathbf{T}\}$.

5.1.1. Loi *a priori* de X

Soit X un processus stochastique où chaque X_n est à valeur dans $\Omega = \{1, \dots, K\}$. Le processus $X_{1 \rightarrow N}$ est une chaîne de Markov finie d'ordre 1, de longueur N , à K états si et seulement si :

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = P(X_{n+1} = j | X_n = i) \quad (3.8)$$

Pour tout $n \in \mathbb{N}$, pour tout état j et pour toute suite d'états $i_0, i_1, \dots, i_n, i$ pour lesquels la probabilité conditionnelle a un sens, c.-à-d. tels que

$$P(X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) > 0 \quad (3.9)$$

Si en plus la probabilité $P(X_{n+1} = j | X_n = i)$ ne dépend pas de n , c.-à-d. si

$$\forall n \in \mathbb{N}, \quad P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i) \quad (3.10)$$

On dit que la chaîne de Markov est *homogène*.

Nous considérons ici uniquement le cas d'une chaîne homogène, pour laquelle la probabilité $P(X_n = x_n | X_{n-1} = x_{n-1})$ ne dépend pas de la position n dans la chaîne. Les chaînes de Markov sont classées dans les systèmes ou les algorithmes sans mémoire. La propriété de Markov est exprimée par l'expression suivante:

- si la valeur de X_n est connue à l'instant n , la loi des variables futures (X_{n+1} , X_{n+2} etc.) ne dépend pas du passé (les valeurs de X_{n-1} , X_{n-2} etc.).

5.1.2. Matrice d'observation

On choisit un alphabet fini ($V = \{v_1, v_2, \dots, v_n\}$). La matrice des observations $\mathbf{O}$ [$k \times ng$] représente la probabilité d'observer le symbole v_{ng} (un niveau de gris où $ng = 256$ niveaux) relativement à un état k donné (la classe).

$$\mathbf{O} [k \times ng] = P(o_i = v_i / X_i = k) \quad (3.11)$$

5.1.3. Matrice de la loi initiale

La loi initiale de la chaîne, qui correspond à la distribution de X_1 , est définie par :

$$\pi_i = P(X_1 = i), \quad \forall i \in \Omega \quad (3.12)$$

5.1.4. Matrice de Transition

On peut ainsi définir la matrice de transition $\mathbf{T}$, qui regroupe l'ensemble des probabilités de transition $\{t_{i,j}\} \forall i, j \in \Omega$ définies par :

$$t_{i,j} = P(X_n = j | X_{n-1} = i), \quad \forall i, j \in \Omega \quad (3.13)$$

Avec $t_{i,j} \geq 0$ et $\sum_{j \in \Omega} t_{i,j} = 1$.

C'est-à-dire si π est un vecteur propre à gauche de $\mathbf{T}$ associé à la valeur propre 1.

On peut alors exprimer la loi *a priori* de X selon :

$$P(X = x) = \pi_{x_1} \prod_{n=2}^N t_{x_{n-1}, x_n} \quad (3.14)$$

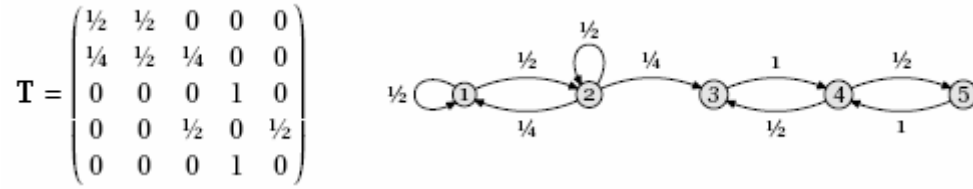


Figure 3.4 : Matrice de transition entre les états.

5.1.5 Loi du couple (X, Y)

De façon classique, les méthodes de restauration de signaux bruités basées sur les CMC utilisent les deux hypothèses suivantes :

- **H₁** : les variables aléatoires Y_n sont indépendantes conditionnellement à X (le processus des classes).
- **H₂** : la distribution de Y_n conditionnellement à X est égale à sa distribution conditionnellement à X_n . D'une autre manière nous pouvons dire : la loi de Y_n conditionnellement $(X_n)_{1 \leq n \leq N}$ ne dépend que X_n .

Ces deux hypothèses permettent d'écrire $P(Y/X)$ comme suit :

$$P(Y = y|X = x) = \prod_{n=1}^N P(Y_n|X_n) \quad (3.15)$$

Les relations entre les processus X et Y dans un modèle de CMC peuvent alors être représentées par le schéma de la figure 3.5 où chaque flèche représente la dépendance (conditionnement) d'une variable aléatoire à une autre. La probabilité de chaque X_n est conditionnée par son prédécesseur (X_{n+1} relié à X_n) ; la probabilité de chaque Y_n est conditionnée par la variable aléatoire X_n correspondante (Y_n relié à X_n).

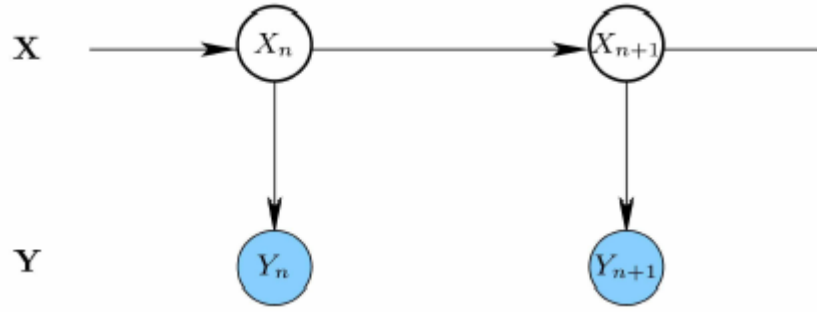


Figure 3.5 : Schéma présentant les liens entre les processus X et Y dans un modèle de chaîne de Markov cachée.

Remarque 1 :

Notons l'introduction de modèles récents qui généralisent les CMC classiques et qui ont permis d'affaiblir les hypothèses H_1 et H_2 . Le modèle des chaînes de Markov Couple [20, 21], strictement plus général que le modèle classique, et qui considère directement la markovianité du processus $Z = (X, Y)$, permet sous certaines conditions, de s'affranchir de l'hypothèse H_2 . Les chaînes de Markov Triplet [23], obtenues en introduisant un processus latent U , considèrent alors la markovianité du triplet (X, U, Y) . De façon similaire, on peut montrer que les chaînes Triplet sont plus générales que les chaînes Couple. La distribution de Y_n conditionnellement à X_n caractérise alors la répartition des observations y_n conditionnellement à l'état x_n considéré. Le processus observé étant à valeur dans R , à chaque état x_n de l'espace des états peut être associée une densité de probabilité $f_{x_n}(y_n)$. Les distributions $f_{x_n}(y_n)$, qui relient le processus observé à la chaîne de Markov, sont couramment appelées les lois d'attache aux données. Ces distributions représentent la variabilité des observations, c'est-à-dire la répartition naturelle des observations de chaque classe, ainsi que les diverses dégradations dues à l'acquisition et à la transmission, que l'on qualifie de bruit.

Finalement, en utilisant les Eq. (3.14) et (3.15) on peut calculer la probabilité conjointe de (X, Y) comme :

$$P(X = x, Y = y) = \pi_{x_1} f_{x_1}(y_1) \prod_{n=2}^N t_{x_{n-1}, x_n} f_{x_n}(y_n) \quad (3.16)$$

La distribution $P(X = x, Y = y)$ est généralement inconnue et doit être estimée avant d'envisager une technique de classification bayésienne. Par conséquent, il est nécessaire d'estimer les deux ensembles de paramètres suivants, à partir des seules données observées :

- L'ensemble Γ constitué des paramètres définissant la loi de $\mathbf{X}$, c'est-à-dire la loi initiale π et la matrice de transition $\mathbf{T}$.
- L'ensemble Δ caractérisant la répartition des observations, c'est-à-dire les $\mathbf{K}$ densités de probabilités $f_k(\cdot)$ associées à chaque état de Ω . Dans le cas gaussien, Δ est uniquement constitué des moyennes et des variances associées à chaque classe.

5.2 Estimation des paramètres

Dans un contexte de classification non supervisée, les paramètres du modèle sont inconnus et doivent être estimés à partir des seules données observées. Des méthodes itératives bien connues telles que l'algorithme EM « Estimation-Maximization » ou SEM « Stochastique EM » peuvent être utilisées. Le principe de l'algorithme EM consiste à maximiser la vraisemblance des données bruitées y par rapport à $\Theta = (\Gamma, \Delta)$. A partir d'une valeur initiale, l'algorithme s'approche itérativement d'un maximum local de vraisemblance. Ceci constitue d'ailleurs son point faible puisqu'il le rend fortement dépendant de l'initialisation. L'algorithme SEM, qui est une variante de l'algorithme EM, est basé sur une procédure itérative qui définit une suite d'estimations au moyen de tirages stochastiques selon la loi *a posteriori* de la chaîne. De par sa nature stochastique, SEM est moins sensible à l'initialisation que EM, et peut atteindre le maximum global de vraisemblance.

Dans notre travail, nous avons considéré une troisième méthode appelée ICE « Iterative Conditional Estimation », dont l'intérêt en segmentation a été illustré dans un certain nombre de situations, en particulier dans divers modèles de Markov cachés (champs, chaînes ou arbres) [3, 6, 7, 8, 13, 16, 20, 23]. En traitement d'images et dans le cas de chaînes gaussiennes, ces algorithmes d'estimation montrent peu de différence lorsque l'initialisation des paramètres recherchés est de bonne qualité [13].

5.2.1 Algorithme d'estimation : ICE

L'algorithme ICE ne se réfère pas à la vraisemblance, mais est basé sur l'espérance conditionnelle des estimateurs à partir des données complètes $(\mathbf{x}, \mathbf{y})$. Il s'agit d'une méthode itérative qui produit une séquence d'estimations $\theta^{[l]}$ des paramètres $\theta \in \Theta$ selon le principe suivant :

1. Initialisation :

L'estimation de $\theta^{[0]}$ est effectuée à partir des données complètes $(y, \tilde{x}^{[0]})$, $\tilde{x}^{[0]}$ étant obtenu à l'aide d'un algorithme de classification initiale (*k-moyennes par exemple*).

2. Induction :

$\theta^{[q+1]} = E[\hat{\theta}(X, Y) | Y = y, \Theta^{[q]}]$, où $\hat{\theta}(X, Y)$ est un estimateur de θ à partir des données complètes.

3. Arrêt :

En pratique, on stoppe l'algorithme à l'itération Q si $\theta^{[Q-1]} \approx \theta^{[Q]}$, ou quand $\|\theta^{[Q-1]} - \theta^{[Q]}\|$ atteint une valeur seuil donnée.

Cet algorithme conduit à deux situations distinctes que nous détaillons par la suite.

- Pour les paramètres de la chaîne de Markov (Γ), l'espérance conditionnelle des estimateurs $\Gamma^{[q-1]}$ est calculable analytiquement (cf. section 5.2.3).
- Par contre, $\delta^{[q+1]}$ ne peut pas être calculée directement pour les paramètres des lois d'attache aux données (Δ). Dans ce cas, l'estimation est effectuée en calculant la moyenne empirique de plusieurs estimations selon $\delta^{[q+1]} = \frac{1}{L} \sum_{l=1}^L \hat{\delta}(\tilde{\mathbf{x}}^{[q,l]}, \mathbf{y})$, avec $\tilde{\mathbf{x}}^{[q,l]}$ une réalisation *a posteriori* de $\mathbf{X}$ conditionnellement à $\mathbf{Y}$, simulée en utilisant $\Theta^{[q]}$ (cf. section 5.2.4).

L'intérêt de la modélisation par les chaînes de Markov cachées repose essentiellement sur l'utilisation des probabilités "progressives" (**forward**) et "rétrogrades" (**backward**).

5.2.2 Procédure Forward-Backward

Le principe de base à l'origine de la procédure de Baum est de décomposer la réalisation bruitée y en deux sous-séquences :

- une sous-séquence de l'instant initial 1 à l'instant n , où l'on attribue à X_n la classe i

$$\alpha_n(i) = P(Y_{1 \rightarrow n} = y_{1 \rightarrow n}, X_n = i) \quad (3.17)$$

- une sous-séquence de l'instant $n+1$ à l'instant final N , sachant que l'on attribue à X_n la classe i .

$$\beta_n(i) = P(Y_{n+1 \rightarrow N} = y_{n+1 \rightarrow N} \mid X_n = i) \quad (3.18)$$

La mise en œuvre de ces probabilités se heurte à des difficultés d'ordre numérique : en effet, les quantités calculées peuvent devenir extrêmement petites, et atteindre les limites de précision des ordinateurs. Afin de s'affranchir de ce problème, il a été proposé d'utiliser les probabilités normalisées (ou *a posteriori*).

– Les probabilités “forward” normalisées sont définies par :

$$\alpha_n^*(i) = P(X_n = i \mid Y_{1 \rightarrow n} = y_{1 \rightarrow n}) \quad (3.19)$$

– et les probabilités “backward” normalisées sont définies par :

$$\beta_n^*(i) = \frac{P(Y_{n+1 \rightarrow N} = y_{n+1 \rightarrow N} \mid X_n = i)}{P(Y_{n+1 \rightarrow N} = y_{n+1 \rightarrow N} \mid Y_{1 \rightarrow n} = y_{1 \rightarrow n})} \quad (3.20)$$

Dans la suite, nous utiliserons le terme de probabilités forward et backward pour désigner les probabilités définies par les Eq. (3.19) et (3.20), et nous omettrons l'astérisque. L'intérêt de l'utilisation de ces probabilités réside dans le fait qu'elles se calculent de façon récursive à l'aide des équations suivantes :

➤ **Initialisation** : $n = 1$:

$$\alpha_1(i) = \frac{\pi_i f_i(y_1)}{\sum_{j \in \Omega} \pi_j f_j(y_1)}, \quad \forall i \in \Omega \quad (3.21)$$

➤ **Induction** : $n > 1$:

$$\alpha_{n+1}(i) = \frac{f_i(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j) t_{j,i}}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j) t_{j,k}}, \quad \forall i \in \Omega \quad (3.22)$$

➤ **Initialisation** : $n = N$:

$$\beta_N(i) = 1, \quad \forall i \in \Omega \quad (3.23)$$

➤ **Induction** : $n < N$:

$$\beta_n(i) = \frac{\sum_{j \in \Omega} t_{j,i} f_i(y_{n+1}) \beta_{n+1}(j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j) t_{j,k}}, \quad \forall i \in \Omega \quad (3.24)$$

A partir de ces définitions, nous pouvons exprimer la probabilité conjointe *a posteriori*

$P(X_n = i, X_{n+1} = j \mid Y = y)$, notée $\Psi_n(i, j)$, par :

$$\Psi_n(i, j) = \frac{\alpha_n(i) t_{j,i} f_i(y_{n+1}) \beta_{n+1}(j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{j \in \Omega} \alpha_n(l) t_{j,k}} \quad (3.25)$$

Et la probabilité marginale *a posteriori* $P(X_n = i \mid Y = y)$, notée $\eta_n(i)$, selon :

$$\eta_n(i) = \sum_{j \in \Omega} \Psi_n(i, j) = \alpha_n(i) \beta_n(i) \quad (3.26)$$

5.2.3 Ré-estimation des paramètres de la chaîne de Markov

Les paramètres de Γ peuvent être ré-estimés de façon déterministe en utilisant les probabilités *a posteriori* définies ci-dessus.

➤ Pour les probabilités initiales, le calcul de $E[\hat{\pi}_i \mid Y, \Theta^{[q]}]$ donne :

$$\pi_i^{[q+1]} = \frac{1}{N} \sum_{n=1}^N \eta_n^{[q]}(i), \quad \forall i \in \Omega \quad (3.27)$$

➤ L'estimation des probabilités de transition $E[\hat{t}_{i,j}|Y, \Theta^{[q]}]$ peut être effectuée selon :

$$t_{i,j}^{[q+1]} = \frac{\sum_{n=1}^{N-1} \Psi_n^{[q]}(i, j)}{\sum_{n=1}^{N-1} \eta_n^{[q]}(i)} \quad \forall i, j \in \Omega \quad (3.28)$$

5.2.4 Ré-estimation des paramètres des lois d'attache aux données :

Pour les paramètres de Δ , $\delta^{[q+1]}$ n'est pas calculable directement. Cependant, ces paramètres peuvent être ré-estimés en calculant la moyenne empirique de plusieurs estimations

selon $\delta^{[q+1]} = \frac{1}{L} \sum_{l=1}^L \hat{\delta}(\tilde{x}^{[q,l]}, y)$, avec $\tilde{x}^{[q,l]}$ une réalisation *à posteriori* de $\mathbf{X}$

conditionnellement à Y , simulée en utilisant $\Theta^{[q]}$. En effet, il a été montré que la loi de $\mathbf{X}$ conditionnellement à Y (noté $\tilde{\mathbf{X}}$) est une chaîne de Markov non stationnaire, dont les paramètres $\tilde{\pi}$ et $\tilde{T}$ peuvent être calculés à partir des probabilités Forward-Backward selon :

$$\tilde{\pi}_i = \frac{1}{N} \sum_{n=1}^N \alpha_n(i) \beta_n(i), \quad \forall i \in \Omega \quad (3.29)$$

$$\tilde{t}_{i,j}(n) = \frac{t_{i,j} f_i(y_n) \beta_n(j)}{\sum_{k \in \Omega} t_{i,k} f_k(y_n) \beta_n(k)}, \quad \forall i, j \in \Omega \quad (3.30)$$

Dans le contexte du traitement d'images, l'expérience a montré qu'une seule réalisation est généralement suffisante, nous utilisons donc dans la suite une unique réalisation ($L=1$) pour estimer $\delta^{[q+1]}$.

Notons $\tilde{x}$ une réalisation de la chaîne *a posteriori* selon $\tilde{\pi}$ et $\tilde{T}$. En considérant que la répartition des observations de chacune des classes peut être modélisée par des distributions gaussiennes, les paramètres associés à $f_i(y_n)$ sont les moments d'ordre 1 (μ_1) et 2 (μ_2). Connaissant y et $\tilde{x}$, les moyennes et les variances de chaque classe sont estimées à l'aide des estimateurs empiriques classiques.

$$\mu_1(i) = \frac{1}{N_i} \sum_{n=1}^N (y_n \times 1_{[\tilde{x}_n=i]}), \quad \forall i \in \Omega \quad (3.31)$$

$$\mu_2(i) = \frac{1}{N_i} \sum_{n=1}^N ([y_n - \mu_1(i)] \times 1_{[\tilde{x}_n=i]})^2, \quad \forall i \in \Omega \quad (3.32)$$

Avec $N_i = \sum_{n=1}^N 1_{[\tilde{x}_n=i]}$, et la fonction $1_{[\tilde{x}_n=i]} = \begin{cases} 1 & \text{si } \tilde{x}_n = i \\ 0 & \text{sinon} \end{cases}$

L'ensemble des étapes nécessaires à l'estimation des paramètres par ICE est synthétisé dans le schéma de la figure (3.6).

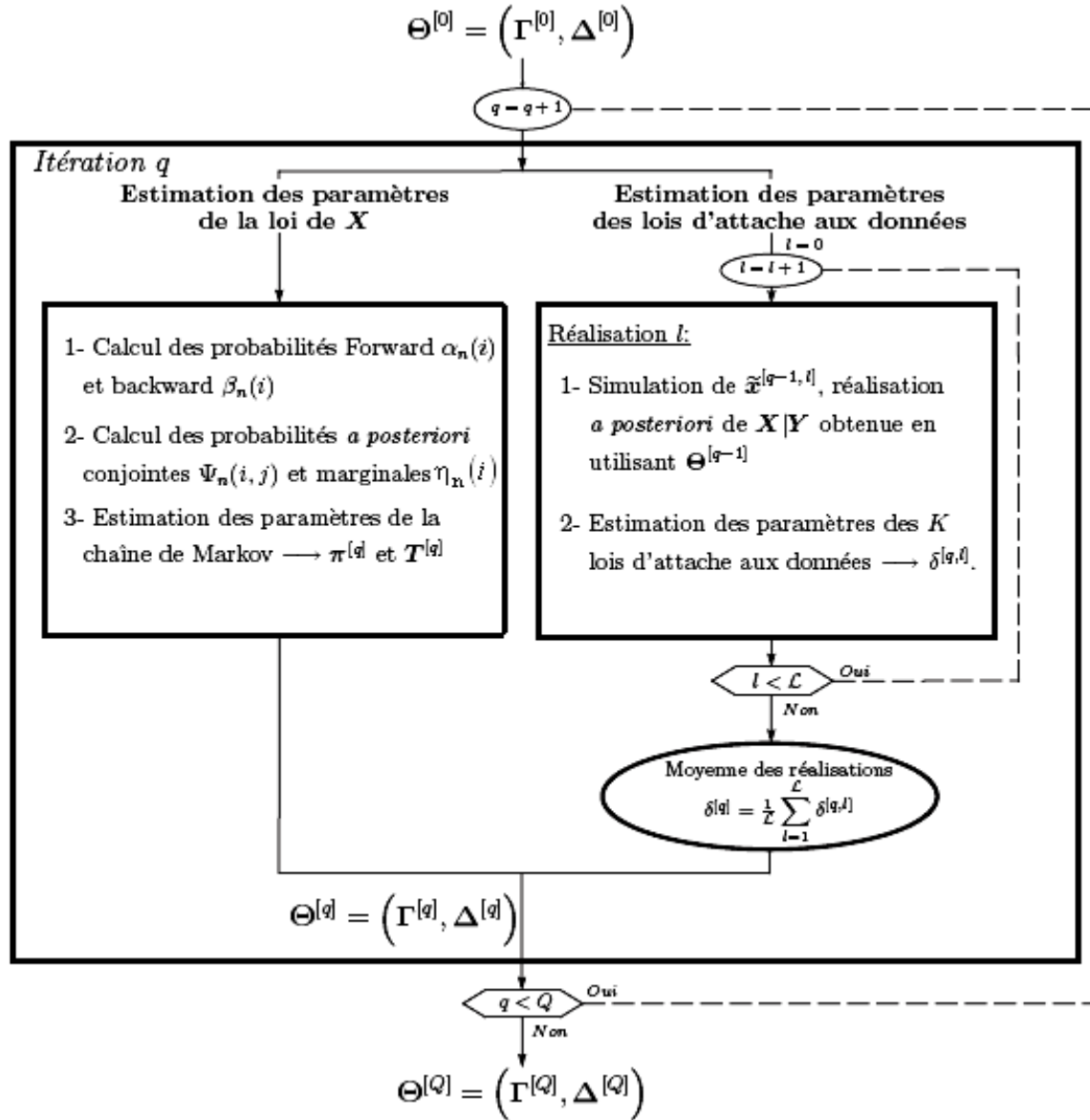


Figure 3.6 : Organigramme de l'algorithme d'estimation de paramètres par la méthode ICE dans le contexte des chaînes de Markov cachées.

5.2.5 Classification bayésienne

Dans le contexte des CMC, le calcul des solutions bayésiennes est explicite, contrairement au modèle des HMRF (Hidden Markov Random Field) qui fait appel à des méthodes d'approximation, coûteuses en temps de calcul, de type Monte-Carlo (**MPM**), ou recuit simulé (**MAP**).

- **Solution du MAP (algorithme de Viterbi) :** Pour calculer la solution du **MAP** dans le cas des CMC, l'algorithme de Viterbi constitue une solution efficace, dont la mise en oeuvre permet d'estimer l'ensemble des états $(x_n)_{1 \leq n \leq N}$ qui maximise la probabilité de **X**

a posteriori. Cette méthode de classification est couramment employée dans les domaines de la reconnaissance de l'écriture manuscrite ou le traitement de la parole.

Cependant, l'estimateur du **MAP** présente l'inconvénient de pénaliser de la même façon une erreur sur un unique pixel et des erreurs sur plusieurs pixels, ce qui n'est pas satisfaisant dans le contexte de la segmentation des images. De plus, il a été montré dans divers travaux [13] que l'estimateur du **MPM** semble être un critère mieux adapté à la problématique de la segmentation d'images, chose que nous avons confirmé après la programmation.

- **Solution du MPM** : Le principe du MPM consiste à estimer en chaque instant n , l'état i qui maximise la probabilité $P(\mathbf{X}_n = i / \mathbf{Y} = \mathbf{y})$. Il en découle donc :

$$\hat{x}_{n \text{ MPM}}(\mathbf{y}) = \arg \max_{i \in \Omega} (\eta_n(i))$$

Dans ce contexte, le calcul explicite de la solution bayésienne du **MPM** permet une mise oeuvre rapide et fiable de la classification (contrairement au modèle des HMRF qui fait appel à des méthodes d'approximation de type (**Monte-Carlo**)). Dans notre travail, nous avons donc principalement utilisé l'estimateur du MPM.

6. Résultats et évaluation

6.1. Introduction

Après avoir présenté la méthode de segmentation par Chaîne de Markov Cachée, nous allons à travers ce paragraphe, présenter les résultats obtenus ainsi que leurs évaluations.

Il existe différentes manières pour évaluer les résultats d'une segmentation d'images. Dans le cas de la segmentation d'images médicales, l'évaluateur final est le clinicien car c'est à lui que revient la dernière décision. On considère une autre méthode d'évaluation des résultats dite quantitative qui possède aussi son importance, elle traduit numériquement la qualité de la segmentation et offre la possibilité de comparer différentes méthodes en spécifiant le domaine où elles donnent les meilleurs résultats.

Dans notre travail nous appliquons la segmentation sur deux types d'images, les images fantômes proposées sur le site de BrainWeb du MNI (Montreal Neurological Institute), dans ce type on peut agir sur les paramètres d'épaisseur de coupe, de bruit et d'hétérogénéité dans les

différentes pondérations et les images réelles acquises dans le service d'IRM au Centre Hospitalo-universitaire de Bab El Oued.

Pour évaluer notre méthode, nous utilisons une évaluation quantitative, puis nous la comparons avec d'autres méthodes de segmentation. Pour ce faire, nous considérons deux méthodes, La première est l'approche coopérative (FCM - croissance de régions basée sur les systèmes multi agents) réalisée par R. Haroun dans [11], la deuxième est la méthode basée sur les ondelettes réalisée par M. Brahim dans [12].

L'implémentation a été faite sur une plate forme de Windows XP en utilisant le langage Builder C++.

6.2 Images utilisées

6.2.1 Images simulées : les fantômes du MNI

Ces images sont générées à l'aide du simulateur proposé sur le site BrainWeb du MNI. Il s'agit d'images de synthèse réalistes dont il est possible de contrôler les paramètres d'épaisseur de coupe, de bruit et d'hétérogénéité. Leur construction est décrite dans [23, 24, 25 et 26]. MNI offre surtout une image de référence segmentée en MB, MG, LCR, et autres tissus.

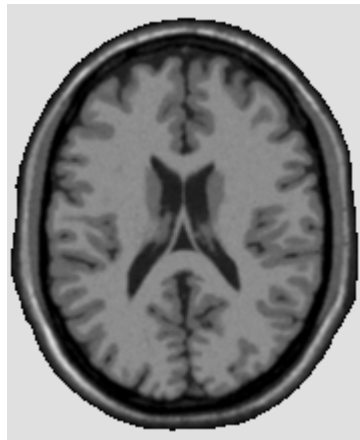


Figure 3.7 : Image de référence du BrainWeb

6.2.2 Images réelles

Les images réelles utilisées ont été obtenues au service d'IRM du Centre Hôpitalo Universitaire de Bab El Oued « CHU » à Alger, Le scanner IRM utilise une induction de 1.5 (tesla).

6.3 Critère d'évaluation quantitative

Afin de quantifier les résultats obtenus, plusieurs critères d'évaluation existent dont nous pouvons citer le critère de Jaccard, celui-ci étant utilisé dans beaucoup de travaux relatifs à la

segmentation d'images de RMN cérébrales [27, 28 et 29]. Ce critère utilise trois facteurs : le nombre de Vrais Positifs (VP), c'est-à-dire de pixels détectés à raison, le nombre de Faux Positifs (FP), c'est-à-dire de pixels détectés à tort et le nombre de Faux Négatifs (FN), c'est-à-dire de pixels non détectés à tort, ces facteurs sont calculés pour chaque classe de tissus (voir Fig. 4.2). A partir de ces valeurs, la similarité entre l'image de référence et l'image segmentée est ensuite quantifiée pour chaque classe de tissu. Le critère de Jaccard est défini de la façon suivante :

$$J = \frac{VP}{VP + FN + FP}$$

L'opérateur « + » désigne l'union des ensembles. Une valeur de « 0 » du critère de Jaccard indique une segmentation complètement fausse, tandis qu'une valeur de « 1 » indique une segmentation parfaite.

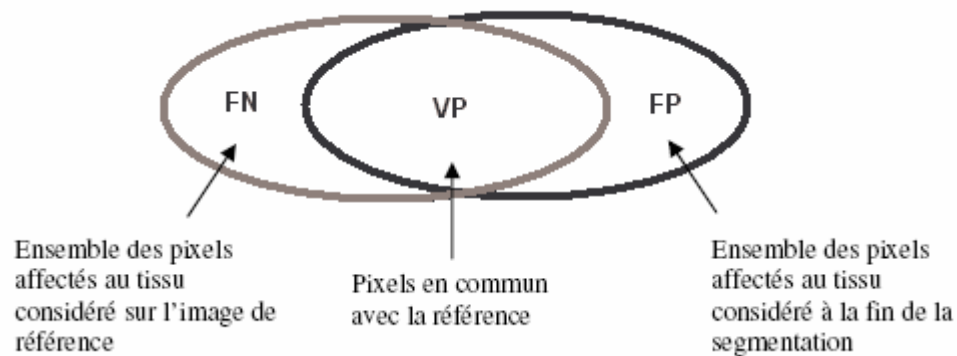


Figure 3.8 : Définition des valeurs de Vrais Positifs (VP), Faux Positifs (FP) et Faux Négatifs (FN) par comparaison de l'image de référence et de l'image segmentée.

Le tableau 3.1 [11] porte les valeurs du critère de Jaccard pour des images simulées de différents niveaux de bruit et de non-uniformité, mais avec une connaissance au départ sur la distribution des pixels pour chaque tissu. Cela peut nous dévier du cadre de la segmentation non supervisée des images IRM.

Non-Uniformité %	Bruit %	MB	MG	LCR
0	0	0.88	0.85	0.60
	20	0.88	0.85	0.60
	40	0.88	0.75	0.59
3	0	0.78	0.78	0.58
	20	0.78	0.77	0.57
	40	0.77	0.77	0.57
5	0	0.78	0.74	0.58
	20	0.77	0.60	0.56
	40	0.77	0.65	0.56
7	0	0.79	0.65	0.57
	20	0.77	0.71	0.56
	40	0.74	0.69	0.56
9	0	0.77	0.71	0.54
	20	0.77	0.71	0.54
	40	0.74	0.69	0.54

Tableau 3.1 : Valeurs du critère de Jaccard pour des images simulées de différents niveaux de bruit et de non-uniformité.

D'après le tableau 3.1, on remarque que le critère de Jaccard pour le LCR reste le plus faible, ceci est dû à la qualité de l'édition de l'encéphale, celle-ci se faisant par une suite de dilatations et d'érosions, généralement le nombre de traitements est sélectionné par l'utilisateur.

La non-uniformité et le bruit font diminuer le critère de Jaccard, les erreurs sont principalement positionnées dans les frontières à cause du volume partiel.

Cependant, nous avons employé une autre évaluation quantitative, celle-ci étant utilisée dans beaucoup de travaux relatifs à la segmentation d'images de RMN cérébrales, s'écrivant comme suit:

$$\tau = \frac{\tilde{X}}{X} \times 100 \% \quad (3.33)$$

Où $\tilde{X}$ est le nombre de pixels détecté dans la segmentation, et X est le nombre total de chaque tissu séparé.

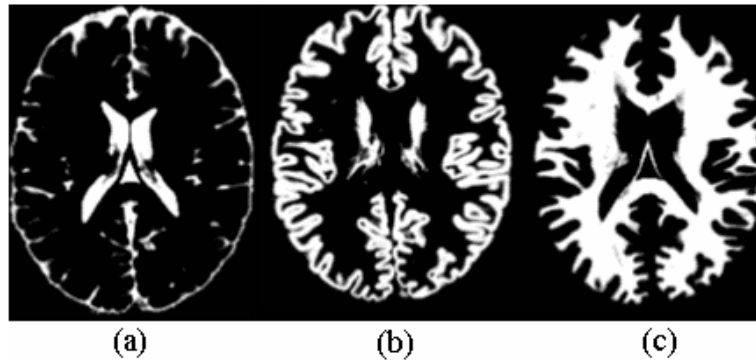


Figure 3.9 : les trois tissus séparés proposés dans le site BrainWeb, (a) LCR, (b) MG, (c) MB

6.4 Segmentation des images

L'objectif de cette section est d'illustrer le modèle classique des CMC sur différentes images IRM. Certains résultats nous serviront ultérieurement comme point de comparaison pour les extensions du modèle des CMC que nous présenterons par la suite.

Il est intéressant de rappeler que dans le contexte non supervisé, les seules données que nous devons introduire sont :

- le nombre de classes : K
- et la forme des lois d'attache aux données (la loi gaussienne dans notre cas).

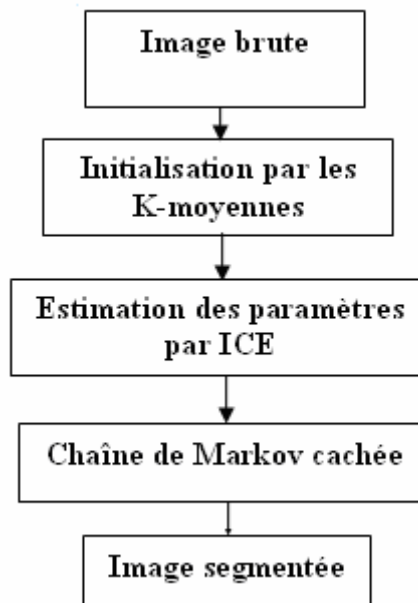


Figure 3.10 : Etapes de la segmentation par Chaîne de Markov cachée.

6.4.1 Segmentation des images de synthèse

Dans cette section nous avons appliqué la méthode de chaîne de Markov sur les images fantômes de différentes coupes, bruit et de non-uniformité pour permettre de visualiser le comportement de cette méthode de segmentation.

Dans le contexte de la segmentation par chaîne de Markov cachée, l'étape de pré segmentation (édition de l'encéphale) sera faite préalablement car elle influe sur le LCR et précisément sur l'évaluation quantitative. Le résultat de la segmentation par CMC donnant alors 4 tissus dont celui représentant l'encéphale qui sera supprimé avant la comparaison des résultats avec les deux autres méthodes (ondelettes et croissance de région) qui elles, utilisent des images éditées.

Pour l'étape de filtrage, nous utilisons un filtre médian de fenêtre 3×3 après la segmentation dans le but d'obtenir un lissage des frontières ; par expérience notre méthode basée sur la chaîne de Markov cachée n'a pas besoin de filtrage, ceci étant du à sa robustesse en présence de bruit.

Nous utilisons parfois indifféremment les termes segmentation en régions ou classification ; cette particularité est liée à la caractéristique propre de l'image IRM, car dans la segmentation en régions on extrait des tissus homogènes par exemple le LCR, MB, MG et dans le cas de la classification (segmentation en classes) on extrait des régions liées par exemple à un niveau de gris ou à une texture.

La figure 3.11 représente une image IRM coronale avec 0% de bruit et 0% de non-uniformité segmentée en trois tissus, le LCR en (noir), la matière grise MG en (orange) et la matière blanche MB en (gris).

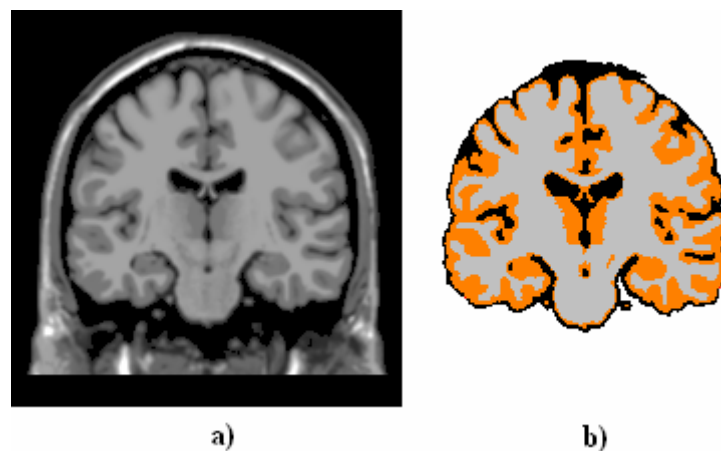


Figure 3.11 : a) image coronale 0% de bruit et 0% de non-uniformité, b) image segmentée en 3 tissus.

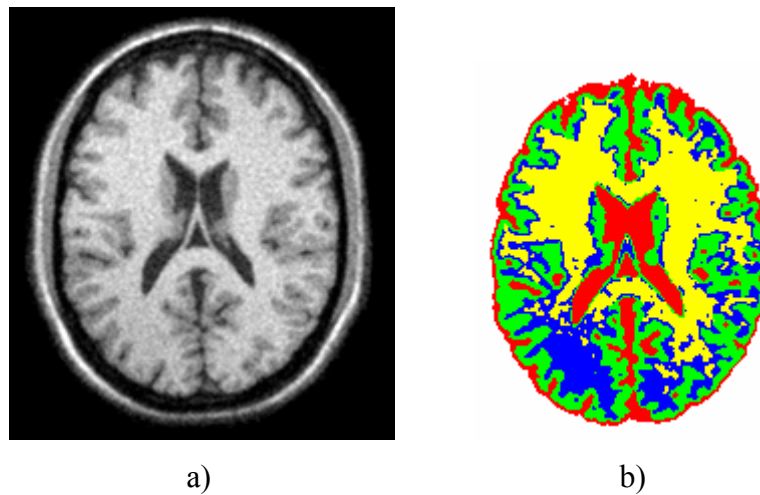


Figure 3.12 : a) image transversale 40% de bruit et 9% de non-uniformité, b) image segmentée en 3 tissus.

Les figures 3.11 et 3.12 représentant deux images IRM avec des pourcentages de bruit et de non-uniformité différents sont segmentées en classes ou en tissus selon la valeur de K donnée à l'initialisation car un tissu peut être représenté par une ou plusieurs classes. Nous avons utilisé la fusion des classes pour extraire un tissu homogène.

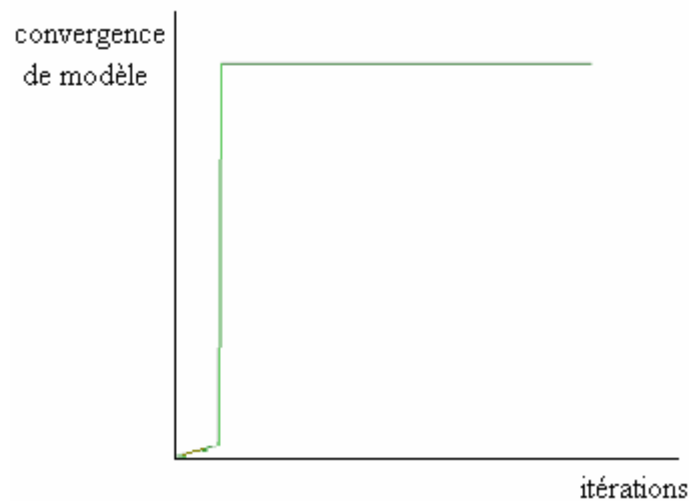


Figure 3.13 : convergence du modèle CMC.

Nous avons appliqué la méthode CMC pour segmenter les images IRM. Les valeurs de l'évaluation quantitative pour différents niveaux de bruit et de non-uniformité sont données par le tableau 3.2. Dans lequel ces valeurs sont relevées par la moyenne de trois fois la segmentation pour chaque image.

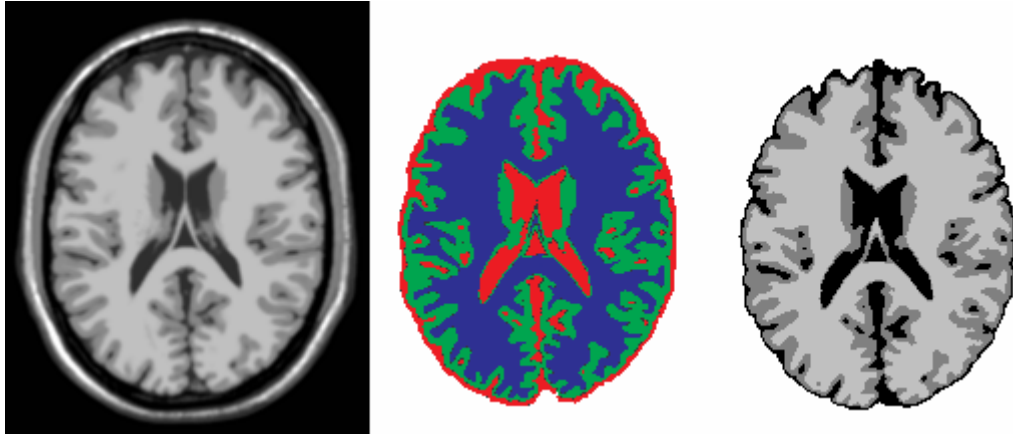
Non-uniformité %	Bruit %	MB %	MG %	LCR %
0	0	82,36	93,86	76,52
3	0	88,55	72,31	72,14
	20	87,74	80,80	76,04
	40	85,05	93,47	69,30
5	0	88,91	77,51	73,44
	20	74,72	94,65	71,89
	40	81,11	93,86	79,27
7	20	84,01	87,64	72,05
	40	85,23	89,45	74,01
9	0	74,27	95,05	75,50
	40	87,95	83,04	70,32

Tableau 3.2 : Valeurs d'évaluation quantitative par rapport à différents paramètres de bruit et de non-uniformité.

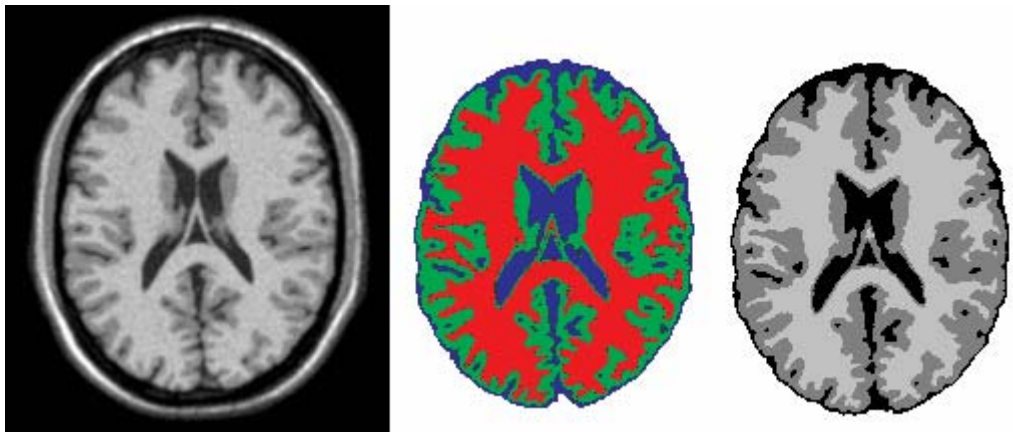
Comme le montre le tableau 3.2, notre méthode de segmentation donne un bon taux de segmentation des trois tissus (par rapport aux résultats obtenus avec l'image référence : 0% de bruit et 0% de non uniformité) malgré la présence de bruit et l'hétérogénéité présents dans l'image.

6.4.1.1 Comparaison avec la méthode coopérative (FCM-croissance de régions basée sur les systèmes multi agents) [11]:

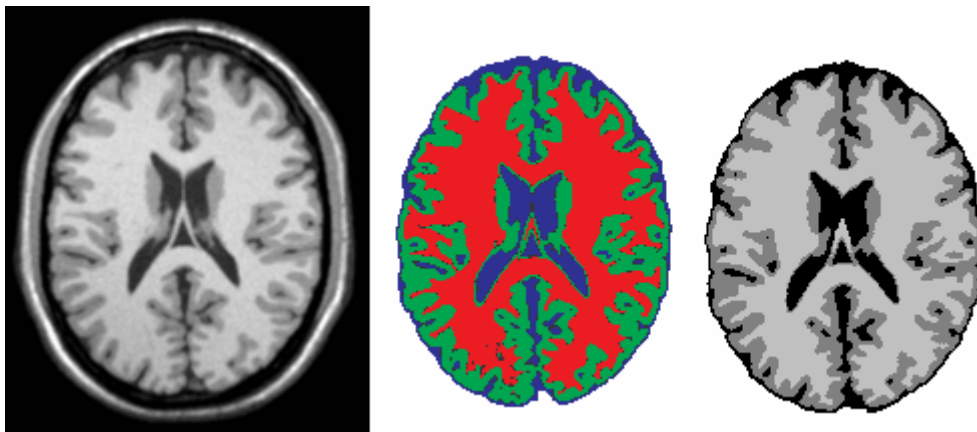
La première colonne à gauche, de la figure 3. 14, représente l'image originale, la deuxième représente l'image segmentée par la méthode de croissance de régions [11] et la troisième celle segmentée par la méthode de chaîne de Markov cachée avec utilisation de la fusion des classes que nous avons implémenté.



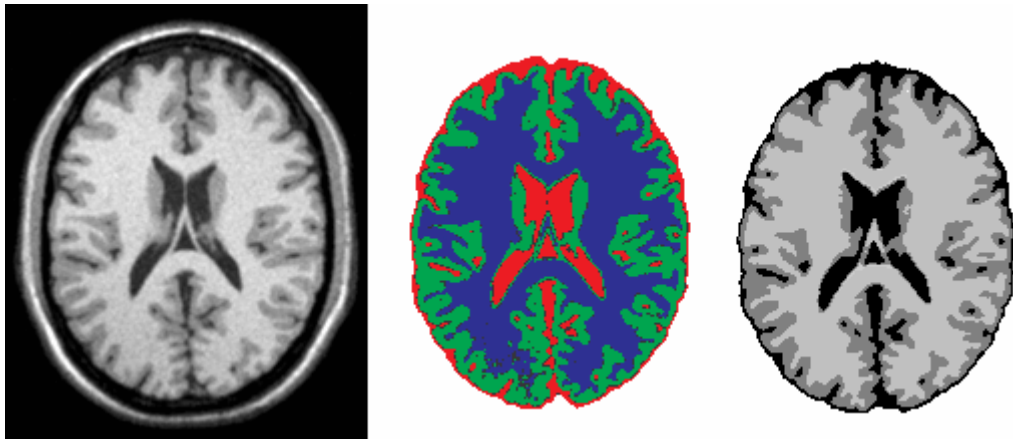
(a) 0% de bruit, 0% de non-uniformité



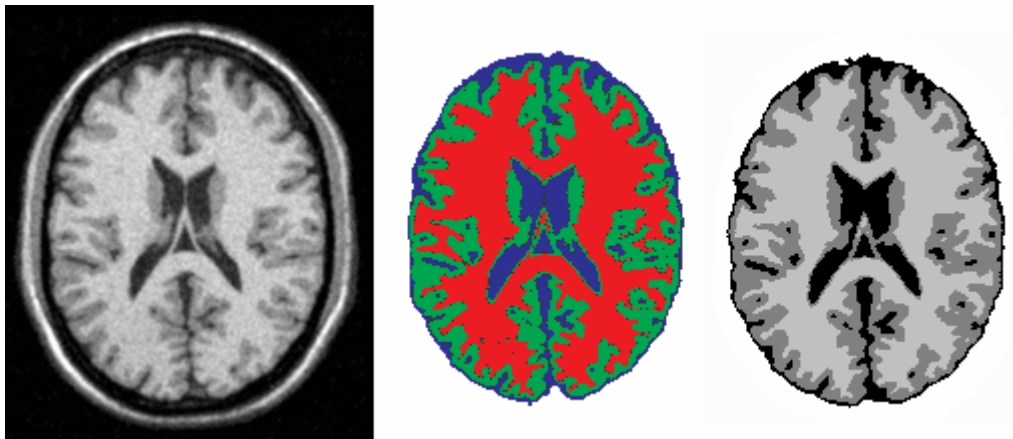
(b) 0% de bruit, 5% non-uniformité



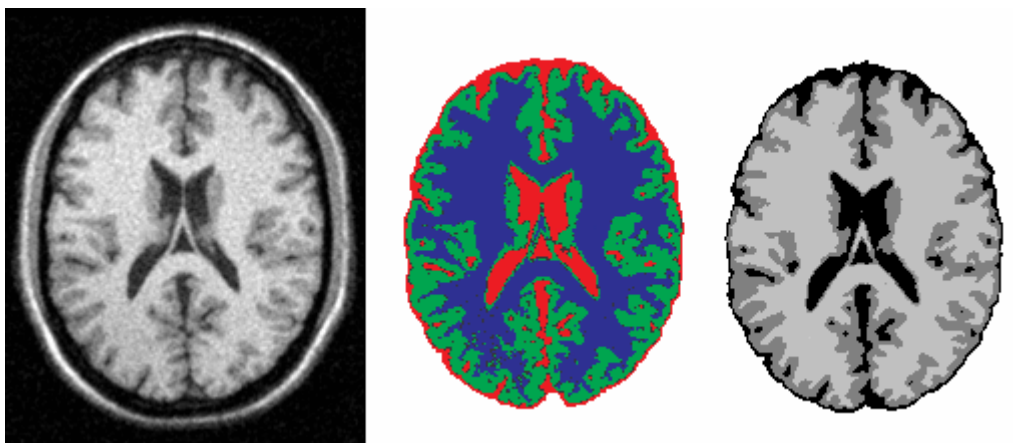
(c) 40% de bruit, 3% non-uniformité



(d) 40% de bruit, 5% non-uniformité



(e) 20% de bruit, 7% de non-uniformité

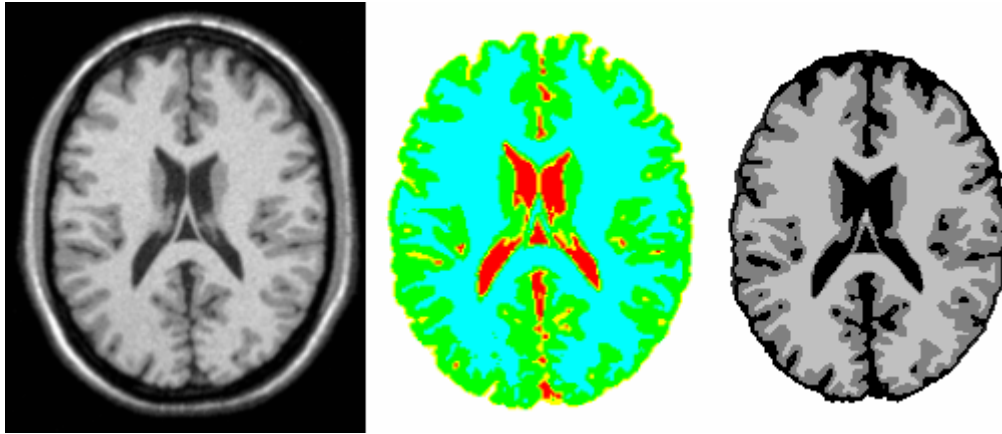


(f) 40% de bruit, 9% de non-uniformité

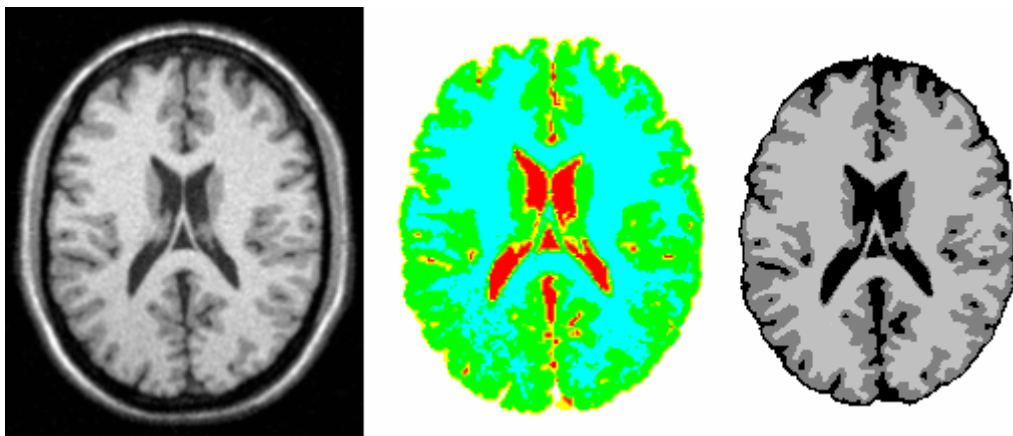
Figure 3.14 : résultats de CMC sur des images fantômes et comparaison avec les résultats de la segmentation par croissance de régions.

6.4.1.2 Comparaison avec la méthode des ondelettes [12]

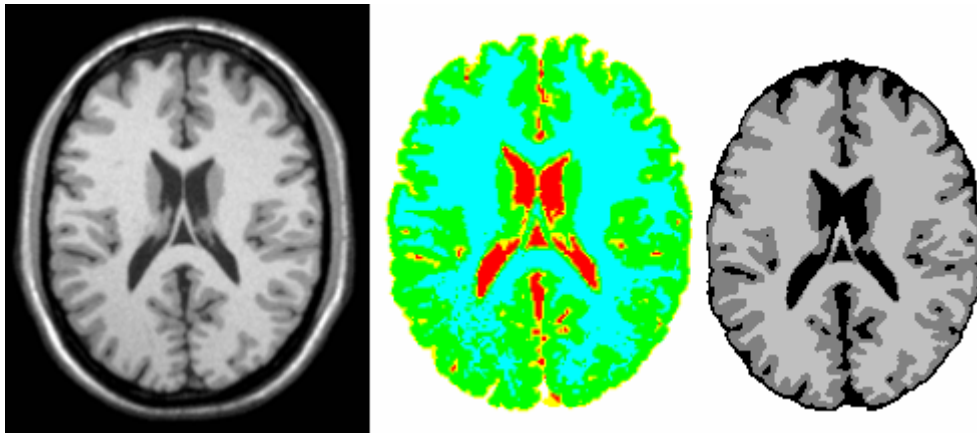
La première colonne à gauche, de la figure 3.15, représente l'image originale, la deuxième représente l'image segmentée par la méthode des ondelettes [12] et la troisième représente l'image segmentée par la méthode de chaîne de Markov cachée avec utilisation de la fusion des classes que nous avons implémentée.



(a) 20% de bruit, 5% non-uniformité



(b) 40% de bruit, 7% de non-uniformité



(c) 40% de bruit, 3% de non-uniformité

Figure 3.15 : résultats de CMC sur des images fantômes et comparaison avec les résultats de la segmentation par ondelettes.

Nous pouvons conclure que la mise en œuvre de la méthode de chaîne de Markov cachée dans la segmentation des images IRM, montre bien sa robustesse au bruit et à la non-uniformité par rapport aux autres méthodes (ondelettes et croissance de région) ; néanmoins, nous avons remarqué que dans certaines images par exemple (image 0% de bruit et 5% de non-uniformité) les tissus sont étendus, ceci est dû à l'artéfact de l'effet de volume partiel.

6.4.2 Segmentation des images réelles

Les images réelles originales sont moyennement bruitées, elles nous ont été fournies en format DICOM, nous les avons ensuite transformées en format BMP lisible par notre logiciel, nous présentons dans ce qui suit, des images originales. L'image de la figure 3.15 b) est celle d'une segmentation en $K = 6$ classes à 100 itérations d'une coupe sagittale de la moëlle cervicale (la moëlle cervicale en couleur vert).

L'utilisation d'un K variable permet de visualiser lisiblement les différents tissus.

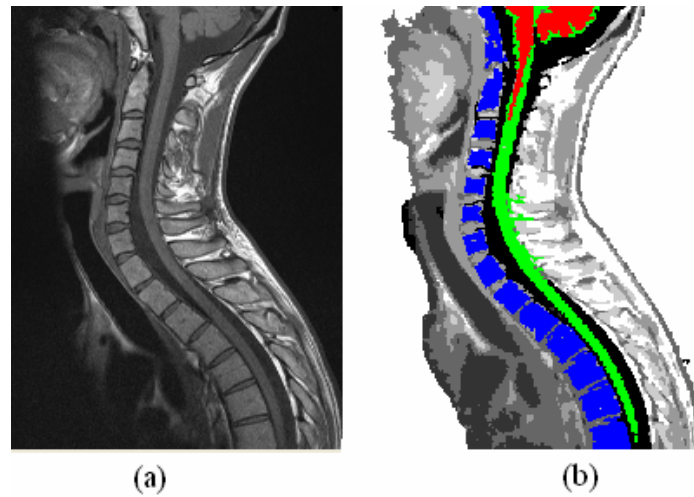


Figure 3.16 : (a) image réelle sagittale, (b) image segmentée en 6 classes par CMC

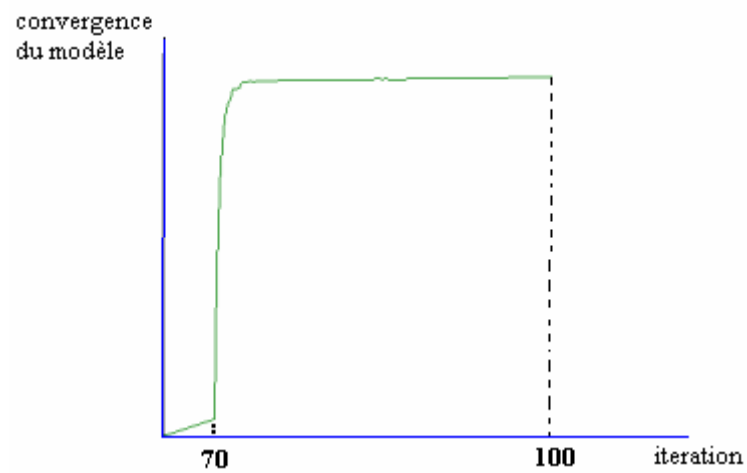


Figure 3.17 : convergence du modèle CMC

La figure 3.16 montre que le modèle CMC converge rapidement (à partir de 70 itérations)

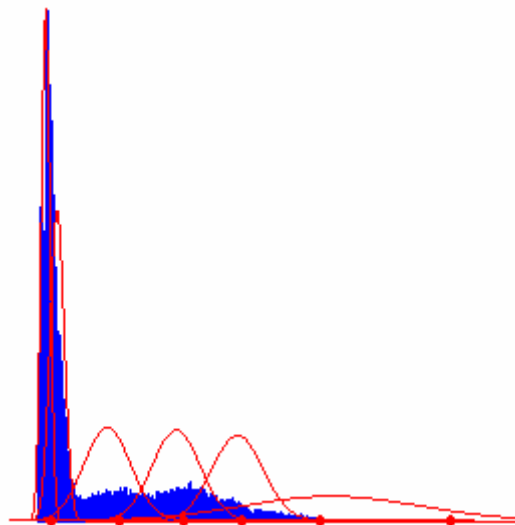


Figure 3.18 : Loi d'attache aux données

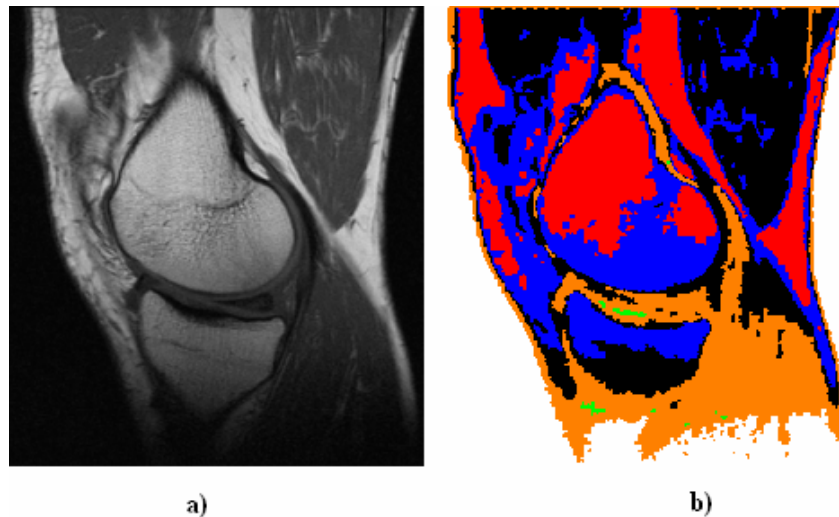


Figure 3.19 : (a) image réelle sagittale du genou, (b) image segmentée en 5 classes par CMC

7. Conclusion

Dans ce chapitre, nous avons présenté le principe d'une segmentation statistique non supervisée d'images basée sur le modèle des chaînes de Markov cachées. Nous avons introduit les outils et les méthodes nécessaires à la mise en oeuvre de cette modélisation dans le contexte de la méthode d'estimation de paramètres ICE.

Les résultats obtenus par cette méthode ont été comparées aux deux méthodes précédemment élaborées par notre équipe (méthode de croissance de régions et celle des ondelettes). Nous avons constaté que la méthode CMC est plus robuste au bruit.

Par extension du modèle classique des CMC détaillé dans ce chapitre, nous allons présenter le modèle des CMC floues, qui constitue une partie de notre travail. Les principes et méthodes présentés dans ce chapitre demeurent valables dans ce nouveau contexte.

Chapitre 4

Chaîne de Markov Cachée floue CMC-f

1 Introduction

Comme nous l'avons illustré dans la section précédente, la restauration bayésienne dans le contexte des CMC est une technique de classification très robuste. Cependant, le modèle des CMC peut s'avérer peu satisfaisant dans certaines situations. Le flou est un effet souvent confondu avec le bruit (section 4.5 Chapitre 1), ce qui déstabilise l'estimation des paramètres du modèle. Ceci conduit généralement à des résultats de segmentation qui occultent la richesse de l'information présente dans l'image bruitée. L'objectif de ce chapitre est de présenter un nouveau modèle de chaîne de Markov cachée flou qui permet de caractériser conjointement le bruit et le flou associés aux observations.

Afin de prendre en compte la nature floue des observations dans les modèles markoviens, il a été proposé dans la littérature d'utiliser, non pas une mesure d'incertitude sur les observations (caractéristique de l'approche probabiliste des modèles markoviens classiques), mais une mesure d'imprécision (caractéristique de l'approche floue). Ces modèles markoviens s'appuient donc sur la théorie des ensembles flous en remplacement de la modélisation probabiliste. Ce modèle, combinant les deux approches, est différent des modèles classiques et flous évoqués précédemment. Ces modèles proposent de tenir compte conjointement des classes dures, et des classes floues. Dans ce paragraphe, nous présentons donc le modèle qui adapte ce point de vue au contexte des CMC floue.

2. Modélisation d'une chaîne de Markov cachée floue [13]

Pour bien montrer la notion de flou dans le nouveau contexte de CMC-f, nous considérerons le cas de la segmentation d'une image en deux classes dures, dénotées 0 et 1, et une classe floue. Chaque pixel est donc susceptible d'appartenir à l'une des deux classes dures, ou à la classe floue. Selon la théorie des sous-ensembles flous, initialement introduite par Zadeh (sec 3.2.2), nous avons alors : $\mathbf{x}_n = (\varepsilon_n^0, \varepsilon_n^1) \in [0, 1]^2$ où, ε_n^0 représente le degré d'appartenance du pixel x_n à la classe dure 0, et ε_n^1 représente le degré d'appartenance du pixel $\mathbf{x}_n$ à la classe dure 1.

Par comparaison avec le cas classique, X_n ne prend donc plus ses valeurs dans l'ensemble discret $\Omega = \{0, 1\}$, mais dans l'intervalle continu $\Omega_f = [0, 1]$. La nouvelle représentation de $\mathbf{X}_n$ est la suivante :

$$X_n = \varepsilon_n \quad \text{avec} \quad \begin{cases} \varepsilon_n = 0 & \text{Si le pixel appartient à la classe dure 0} \\ \varepsilon_n \in]0, 1[& \text{Si le pixel est flou,} \\ \varepsilon_n = 1 & \text{Si le pixel appartient à la classe dure 1} \end{cases}$$

Ainsi la chaîne $\mathbf{X}$ est un processus à valeur dans l'ensemble $\Omega_f^N = [0, 1]$. La nouvelle approche requiert donc la définition des probabilités de la chaîne sur Ω_f . La loi de chaque X_n est donnée par une densité qui inclue à la fois des composantes discrètes (classes « dures »), et une composante continue (classe « floue »). En notant δ_0, δ_1 les mesures de Dirac en 0 et 1 et ν la mesure de Lebesgue sur $]0, 1[$ et en prenant $\nu = \delta_0 + \nu + \delta_1$ comme mesure sur Ω_f , la distribution de X_n peut alors être définie par une densité h sur $[0, 1]$ selon la mesure ν . La loi de X_n s'écrit alors :

$$P(X_n = \varepsilon_n) = \begin{cases} h(\varepsilon_n = 0) = \pi_0, \\ h(\varepsilon_n \in [a, b]) = \int_a^b h(\varepsilon) d\nu(\varepsilon), \text{ avec } 0 < a \leq b < 1 \\ h(\varepsilon_n = 1) = \pi_1. \end{cases}$$

h vérifie donc :

$$h(0) + h(1) + \int_0^1 h(\varepsilon) d\nu(\varepsilon) = 1 \quad (4.1)$$

La loi initiale de la chaîne, qui correspond à la distribution de $\mathbf{X}_1$, est donc maintenant définie par la fonction h , et non pas par un vecteur comme dans le cas classique. De façon comparable, les probabilités de transition sont aussi définies par une fonction, et non pas par une matrice.

Remarque 1 :

L'établissement du modèle des CMC floues dans le cas général à K classes dures implique la définition d'une mesure ν sur $[0, 1]^k$, ce qui ne pose pas de difficulté majeure pour le modèle, mais devient rapidement insupportable d'un point de vue numérique. Dans ce paragraphe, nous limitons la présentation des CMC floues au cas de $K = 2$ classes dures.

De façon similaire aux CMC classiques, il convient d'estimer $P(X, Y)$ pour appliquer une décision bayésienne. Nous considérons toujours les deux hypothèses H_1 et H_2 évoquées dans la section (5.1.5), et l'expression de la probabilité conjointe de (X, Y) présentée par l'Eq (3.25)

demeure inchangée. Pour réaliser l'estimation des paramètres du nouveau modèle, nous utilisons un nouveau l'algorithme ICE. Dans ce nouveau contexte, nous devons donc estimer les paramètres de la chaîne de Markov floue (Γ), et l'ensemble Δ composé des paramètres des lois d'attache aux données des deux classes dures et de la classe floue. La phase de l'estimation des paramètres de la chaîne de Markov floue devient comme définie dans le paragraphe suivant.

3. Estimation des paramètres de la chaîne de Markov floue

3.1 Les probabilités Forward-Backward

L'adaptation des probabilités Forward-Backward de **Baum Welech** au nouveau contexte flou est comme suit :

$$\alpha_1(\xi) = \frac{\pi_\xi f_\xi(y_{n+1})}{\int_0^1 \pi_\xi f_\xi(y_{n+1}) dv(\zeta)} \quad (4.2)$$

$$\alpha_{n+1}(\xi) = \frac{\int_0^1 \alpha_n(\zeta) t_{\zeta, \xi} f_\xi(y_{n+1}) dv(\zeta)}{\int_0^1 f_\Gamma(y_{n+1}) \int_0^1 \alpha_n(\zeta) t_{\zeta, \Gamma} f_\xi(y_{n+1}) dv(\zeta) dv(\Gamma)} \quad \forall \xi \in \Omega_f \quad (4.3)$$

$$\beta_N(\xi) = 1 \quad (4.4)$$

$$\beta_n(\xi) = \frac{\int_0^1 \alpha_n(\zeta) t_{\zeta, \xi} f_\xi(y_{n+1}) dv(\zeta)}{\int_0^1 f_\Gamma(y_{n+1}) \int_0^1 \alpha_n(\zeta) t_{\zeta, \Gamma} f_\xi(y_{n+1}) dv(\zeta) dv(\Gamma)} \quad \forall \xi \in \Omega_f \quad (4.5)$$

Remarque 2 :

La seule différence entre ces probabilités et celles définies dans le cas des CMC classiques est le passage d'une somme discrète sur $\Omega = \{0, 1\}$, à une intégrale continue sur $\Omega_f = [0, 1]$.

De façon similaire au cas classique, les probabilités *a posteriori* conjointes $\Psi_n(\xi, \zeta)$ et marginales $\eta_n(\xi)$ s'écrivent :

$$\Psi_n(\xi, \zeta) = \frac{\alpha_n(\xi) t_{\zeta, \xi} f_{\xi}(y_{n+1}) \beta_{n+1}(\zeta)}{\int_0^1 f_{\Gamma}(y_{n+1}) \int_0^1 \alpha_n(\zeta) t_{\zeta, \Gamma} dv(\zeta) dv(\Gamma)} \quad \forall \xi, \zeta \in \Omega_f \quad (4.6)$$

$$\eta_n(\xi) = \int_0^1 \Psi_n(\xi, \zeta) dv(\zeta) = \alpha_n(\xi) \beta_n(\xi), \quad \forall \xi \in \Omega_f \quad (4.7)$$

3.2 Ré-estimation des paramètres de la chaîne de Markov floue

La ré-estimation des paramètres de la loi *a priori* de X s'effectue avec les mêmes équations que celles utilisées dans le cas classique ; la seule différence provenant de l'expression des probabilités $\alpha_n(\xi)$, $\beta_n(\xi)$, $\psi_n(\xi, \zeta)$, $\eta_n(\xi)$ définies ci-dessus.

➤ Pour les probabilités initiales, l'estimation est effectuée selon :

$$\pi_{\xi}^{[q+1]} = \frac{1}{N} \sum_{n=1}^N \eta_n^{[q]}(\xi), \quad \forall \xi \in \Omega_f \quad (4.8)$$

➤ L'Estimation des probabilités de transition peut être effectuée selon :

$$t_{\xi, \zeta}^{[q+1]} = \frac{\sum_{n=1}^{N-1} \Psi_n^{[q]}(\xi, \zeta)}{\sum_{n=1}^{N-1} \eta_n^{[q]}(\xi)} \quad \forall \xi, \zeta \in \Omega_f \quad (4.9)$$

3.3 Ré-estimation des paramètres des lois d'attache aux données

L'ensemble des paramètres constituant Δ doit aussi être ré-estimé dans ce nouveau contexte. Nous devons alors distinguer le cas des deux classes dures de celui de la classe floue. Pour simplifier les explications, sans pour autant restreindre la généralité des propos, nous supposons que la répartition des observations de chaque classe peut être modélisée par une loi gaussienne, notée $\Delta(\mu_1, \mu_2)$.

3.3.1 Cas des classes dures

Pour les paramètres des classes dures, l'estimation s'effectue de façon similaire aux CMC classiques, à l'aide d'une réalisation *a posteriori* de $\mathbf{X}$ conditionnellement à $\mathbf{Y}$. Les paramètres de la chaîne de Markov *a posteriori* sont calculés à partir des nouvelles probabilités Forward-Backward (Eq. (4.2) à (4.5)) selon :

$$\tilde{\pi}_{\xi} = \frac{1}{N} \sum_{n=1}^N \alpha_n(\xi) \beta_n(\xi), \quad \forall \xi \in \Omega_f \quad (4.10)$$

$$\tilde{t}_{\xi, \zeta}(n) = \frac{t_{\xi, \xi} f_{\zeta}(y_n) \beta_{n+1}(\zeta)}{\int_0^1 t_{\xi, \Gamma} f_{\Gamma}(y_n) \beta_n(\zeta) dv(\zeta)}, \quad \forall \xi, \zeta \in \Omega_f \quad (4.11)$$

Comme dans le cas classique, les moyennes et variances des deux lois d'attache aux données dures sont alors estimées à l'aide des estimateurs empiriques classiques. La répartition des observations dures est alors caractérisée par les distributions gaussiennes $\Delta(\mu_1^{(0)}, \mu_2^{(0)})$ pour la classe dure 0 et $\Delta(\mu_1^{(1)}, \mu_2^{(1)})$ pour la classe dure 1.

3.3.2 Cas de la classe floue

Afin de définir la loi d'attache aux données de la classe floue, nous considérons deux variables aléatoires indépendantes $X^{(0)}$ et $X^{(1)}$ correspondant respectivement aux classes dures 0 et 1, caractérisées par les distributions. En considérant qu'un pixel flou contient une « proportion » ε de classe dure 1, (et par conséquent une proportion $1 - \varepsilon$ de classe dure 0), nous pouvons alors exprimer la variable aléatoire correspondant à la classe floue selon :

$$X^{(\varepsilon)} = (1 - \varepsilon) X^{(0)} + \varepsilon X^{(1)} \quad (4.12)$$

Le paramètre de distribution qui caractérise la variable aléatoire $X^{(\varepsilon)}$ est $\Delta(\mu_1^{(\varepsilon)}, \mu_2^{(\varepsilon)})$ avec :

- Pour la moyenne :

$$\mu_1^{(\varepsilon)} = (1 - \varepsilon) \mu_1^{(0)} + \varepsilon \mu_1^{(1)} \quad (4.13)$$

- Pour la variance :

$$\mu_2^{(\varepsilon)} = (1 - \varepsilon)^2 \mu_2^{(0)} + \varepsilon^2 \mu_2^{(1)} \quad (4.14)$$

3.4 Classification bayésienne

Comme nous l'avons vu dans le cas classique, nous utilisons la solution du **MPM**, qui s'exprime simplement en fonction des nouvelles probabilités *a posteriori* définies par :

$$\hat{x}_{n\text{MPM}}(y) = \arg \max (\alpha_n(\xi) \beta_n(\xi)) \quad \xi \in [0, 1] \quad (4.14)$$

Cette procédure se déroule en deux temps :

1. Application de la règle bayésienne classique sur les trois classes $\{0, 1,]0, 1[\}$.
2. Si l'élément choisi est dans $\{0, 1\}$, on arrête. Sinon, on choisit dans $]0, 1[$ la valeur qui maximise à nouveau la règle de classification bayésienne.

4. Approximation numérique du modèle

Les intégrales mises en oeuvre dans les nouvelles probabilités de **Baum Welech** ne peuvent pas être calculées de façon analytique, et il convient de proposer une approximation numérique qui permet leur implémentation informatique. Pour cela, on divise l'intervalle $[0, 1]$

en F sous-intervalles flous $[a_0 = 0, a_1 = 1/F], [a_1 = 1/F, a_2 = 2/F], \dots [a_{F-1} = (F-1)/F, a_F = 1]$. À chaque sous-intervalle flou, nous associons la valeur médiane correspondante. À titre d'exemple:

- $F = 2$ implique $[a_0 = 0, a_1 = 0.5], [a_1 = 0.5, a_2 = 1]$ et par conséquent $\varepsilon \in \{0, 0.25, 0.75, 1\}$.
- $F = 3$ implique $[a_0 = 0, a_1 = 0.33], [a_1 = 0.33, a_2 = 0.66], [a_2 = 0.66, a_3 = 1]$ et par conséquent $\varepsilon \in \{0, 0.165, 0.5, 0.825, 1\}$.

Nous remarquons que plus F est grand, plus l'estimation des paramètres est précise. D'un point de vue numérique et algorithmique, les traitements associés à une chaîne de Markov cachée floue à deux classes dures et F sous-intervalles flous sont très proches de ceux mis en oeuvre pour une chaîne de Markov cachée classique à $F + 2$ états. Néanmoins, il est important de noter que les paramètres des lois d'attache aux données floues sont contraints par ceux des classes dures et du choix de l'expression de la variable aléatoire $X^{(\varepsilon)}$.

5. Segmentation par chaînes de Markov floues

D'après les résultats de chaîne de Markov CMC sur les images fantômes, nous avons remarqué que les classes (tissus), sont bien segmentées dans les images ayant un bruit élevé de 20-40 %, ce qui montre bien la robustesse des CMC par rapport à la présence de bruit, par contre dans les images non bruitées, les classes sont étendues grâce à l'effet de floue (figure 4.1).

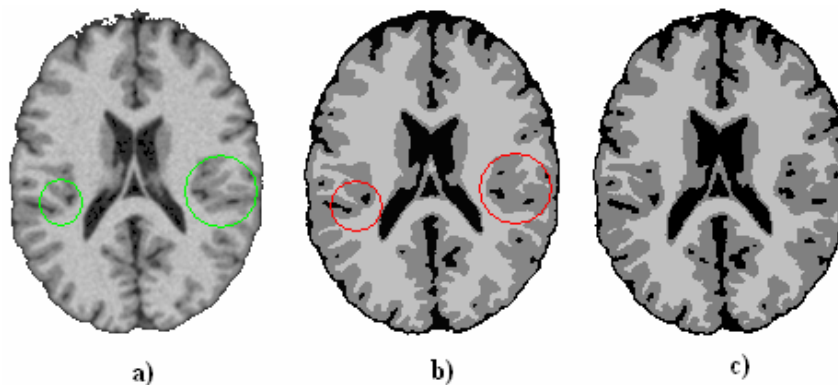


Figure 4.1 : a) image fantôme 0% de bruit et 5% de non-uniformité, b) image segmentée en 3 tissus, c) image 0% de bruit et de 3% de non-uniformité segmentée en 3 tissus.

Dans ce contexte, et afin d'améliorer les capacités de la chaîne de Markov, nous avons utilisé une autre variante de CMC qui traite les deux artéfacts en même temps (bruit et effet de flou) dite chaîne de Markov floue CMC-f.

Pour l'implémentation de la CMC-f nous nous sommes limités sur deux classes dures à la fois (MB-MG, MG-LCR) et pour les sous-classes floues nous nous sommes limités à 2 pour $\Omega_1=[MB, MG]$ et 1 pour $\Omega_2 = [MG, LCR]$, cette configuration permet de localiser le traitement dans un tissu touché par un cancer, et par expérience le tissu le plus touché par les cancers est la matière blanche. Ceci permet d'atteindre l'objectif de la segmentation.

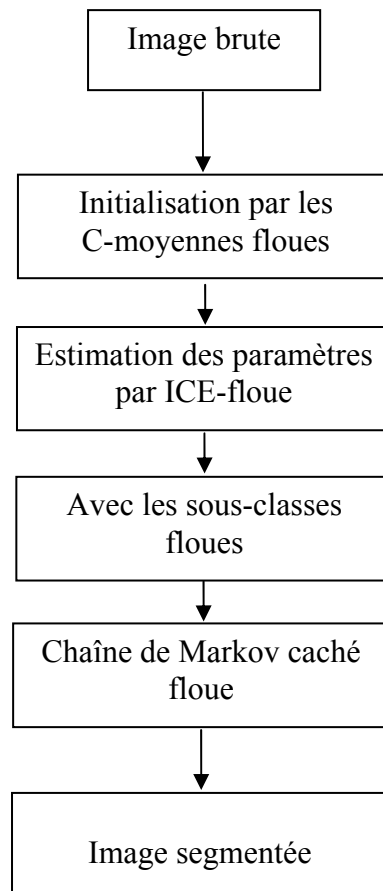


Figure 4.2 : Etapes de la segmentation par chaîne de Markov floue

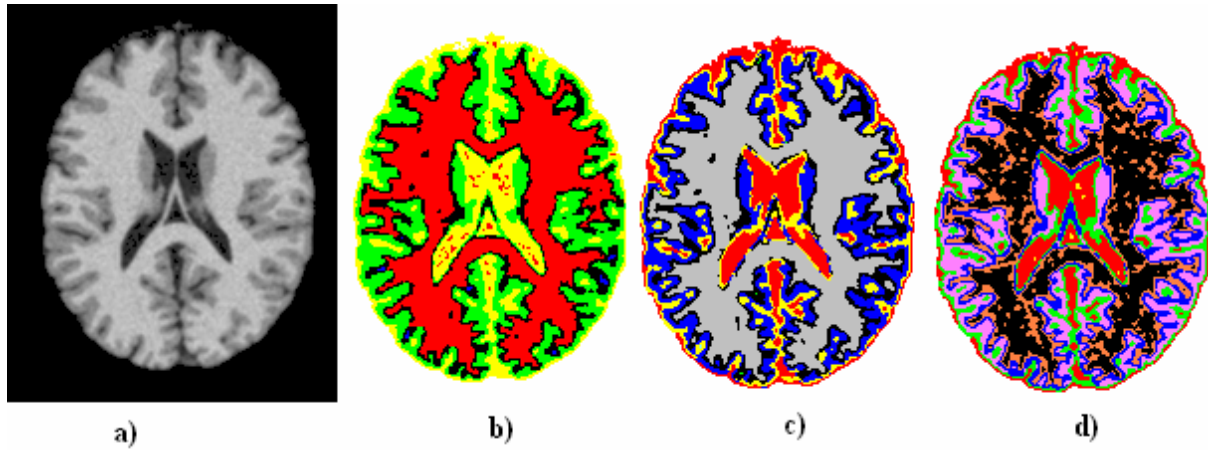


Figure 4.3 : a) image brute 0% de bruit, 5% de non uniformité éditée, b) image segmentée CMC-f où $f(\Omega_1)=1$, c) CMC-f où $f(\Omega_1)=f(\Omega_2)=1$, d) CMC-f où $f(\Omega_1)=2$, $f(\Omega_2)=1$.

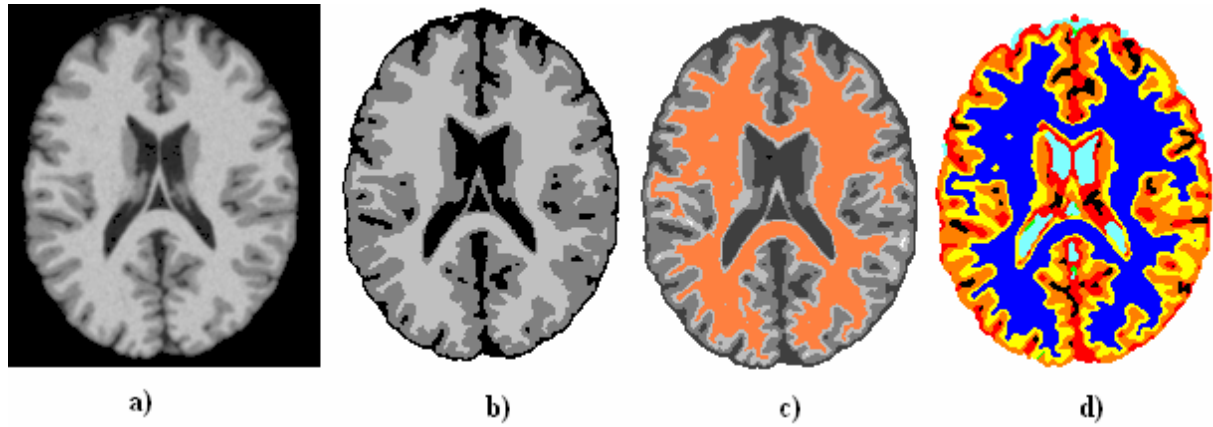


Figure 4.4 : a) image brute 0% de bruit, 3% de non-uniformité, b) CMC classique, c) CMC-f où $f(\Omega_1)=1$, d) CMC-f où $f(\Omega_1)=2$ et $f(\Omega_2)=1$.

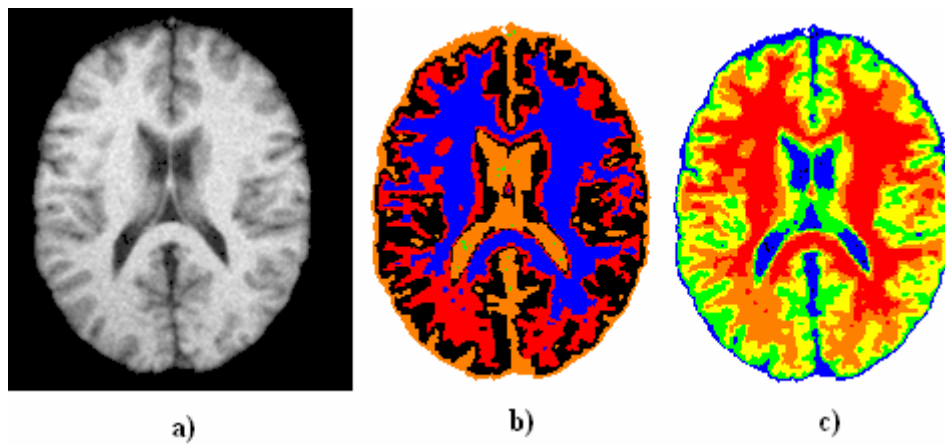


Figure 4.5 : a) image brute 40% de bruit et 9% de non-uniformité coupe 3mm, b) CMC-f où $f(\Omega_1)=1$, c) CMC-f où $f(\Omega_1)=f(\Omega_2)=1$

Les images segmentées par l'approche de chaîne de Markov floue CMC-f visualisées dans les figures 4.3, 4.4 et 4.5 montrent bien l'intérêt d'utiliser des sous classes floues pour bien séparer les tissus et pour améliorer la segmentation (figure 4.5).

En ce qui concerne la qualité de la segmentation par utilisation de chaîne de Markov floue, nous remarquons dans les images segmentées que les tissus (MB, MG) ne sont pas encore bien segmentés. Pour résoudre ce problème nous utilisons la diffusion des classes pour extraire les tissus propres (MB, MG, LCR). Le résultat de cette opération est montré dans les deux figures (4.6 et 4.7).

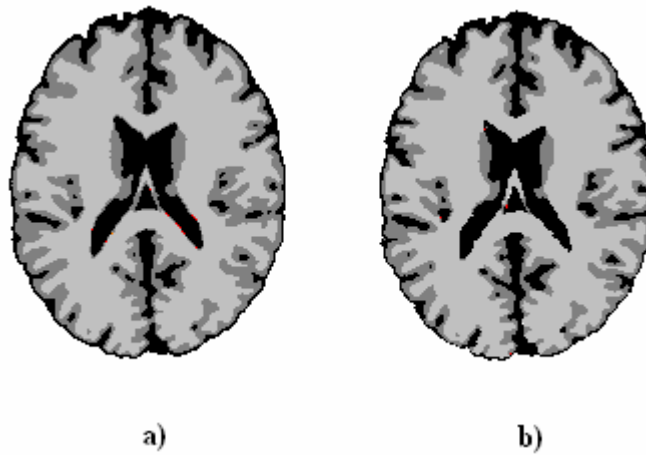


Figure 4.6 : a) image segmentée par CMC-f où $f(\Omega_1) = 1$, b) CMC-f où $f(\Omega_1) = f(\Omega_2) = 1$.

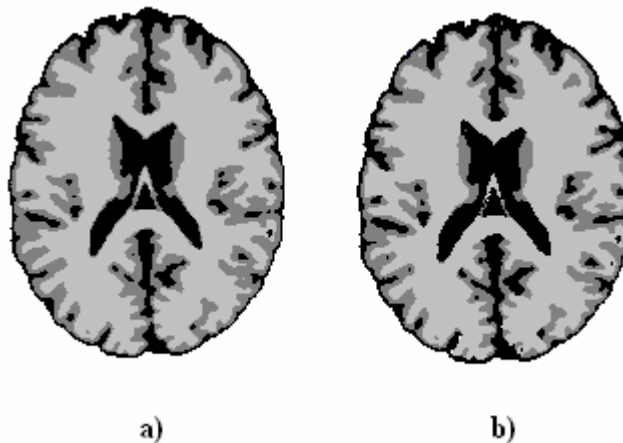


Figure 4.7 : a) CMC-f où $f(\Omega_1) = 1$, b) CMC-f où $f(\Omega_1) = 2$ et $f(\Omega_2) = 1$.

6. Conclusion

Dans ce chapitre, nous avons présenté la méthode de segmentation floue basée sur le modèle des chaînes de Markov cachées, et son application en segmentation d'images non supervisée. L'intérêt essentiel de ce modèle repose sur la prise en compte conjointe des classes dures et de la classe floue dans la modélisation markovienne. Afin de traiter le problème de la classification bayésienne dans ce nouveau contexte, nous avons proposé et mis en oeuvre une extension floue de l'algorithme d'estimation des paramètres ICE.

À partir des expériences, nous avons pu mettre en évidence que dans un contexte bruité, le modèle des CMC floues s'avère intéressant en comparaison avec les modèles classiques, comme l'algorithme des *C-moyennes* floues. La prise en compte simultanée des caractéristiques de bruit et de flou dans le modèle permet d'obtenir des résultats de segmentation pertinents. De façon plus générale, le modèle des CMC floues semble être utile dans les situations où l'objectif est de détecter et caractériser des dégradés bruités, situés de part et d'autre de zones plus homogènes. La plupart des algorithmes de segmentation sont capables de segmenter une image possédant des frontières clairement définies, mais peu d'algorithmes sont capables d'identifier une zone de « transition » entre deux classes.

Chapitre 5

Présentation de logiciel CMC_IRM _Segmentation

1. Introduction

Le but principal de ce travail, est la conception d'un logiciel effectuant la segmentation d'images IRM par Chaîne de Markov cachée. Pour atteindre ce but, nous avons été amené à utiliser le langage de programmation Builder C++ 5.

Dans ce chapitre, nous allons tout d'abord présenter brièvement les différents logiciels utilisés, en expliquant leur choix et les avantages qu'ils présentent. Puis nous passerons à la présentation du logiciel que nous avons développé. Par la suite, nous allons exposer les différentes palettes qui constituent le logiciel et expliquer le rôle de chacune d'entre elles.

2. Logiciel utilisé Builder C++ :

Tout d'abord C++ est un outil RAD, c'est à dire tourner vers le développement rapide d'applications (Rapid Application Development) sous Windows. En un mot, C++ Builder permet de réaliser de façon très simple l'interface des applications et de relier aisément le code utilisateur aux événements Windows, quelle que soit leur origine (souris, clavier, événement système, etc.) Pour ce faire, C++ Builder repose sur un ensemble très complet de *composants visuels* prêts à l'emploi. La quasi-totalité des contrôles de Windows (boutons, boîtes de saisies, listes déroulantes, menus et autres barres d'outils) y sont représentés, regroupés par famille. Leurs caractéristiques sont éditables directement dans une fenêtre spéciale intitulée *éditeur d'objets*. L'autre volet de cette même fenêtre permet d'associer du code au contrôle sélectionné. Il est possible d'ajouter à l'environnement de base des composants fournis par des sociétés tierces et même d'en créer soit même. Un outil RAD c'est également un ensemble de *squelettes* de projets qui permettent de créer plus facilement une application SDI ou MDI , une DLL , des objets OLE , etc. A chacun de ces squelettes est habituellement associé un *expert* qui par une série de boîtes de dialogues permet de fixer une partie des options essentielles à la réalisation du projet associé.

Ainsi, c'est pour la simplicité de son utilisation et pour la diversité des objets (prés à l'emploi) qu'il offre, que nous nous sommes intéressés au langage Builder C++.

3. Logiciels développés :

On donne ci-dessous l'interface et le résultat de l'exécution de notre programme.

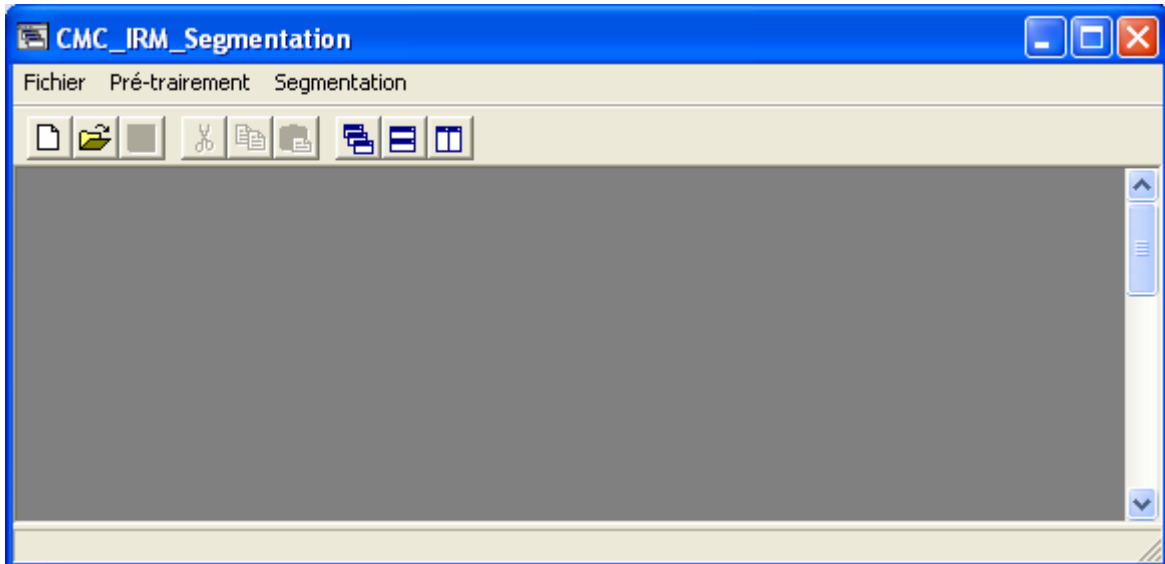


Figure 5.1 : Fenêtre principale

3.1 Les menus

3.1.1 Menu Fichier

Pour le chargement des images, on clique sur l'article Ouvrir



Figure 5.2 : Menu Fichier

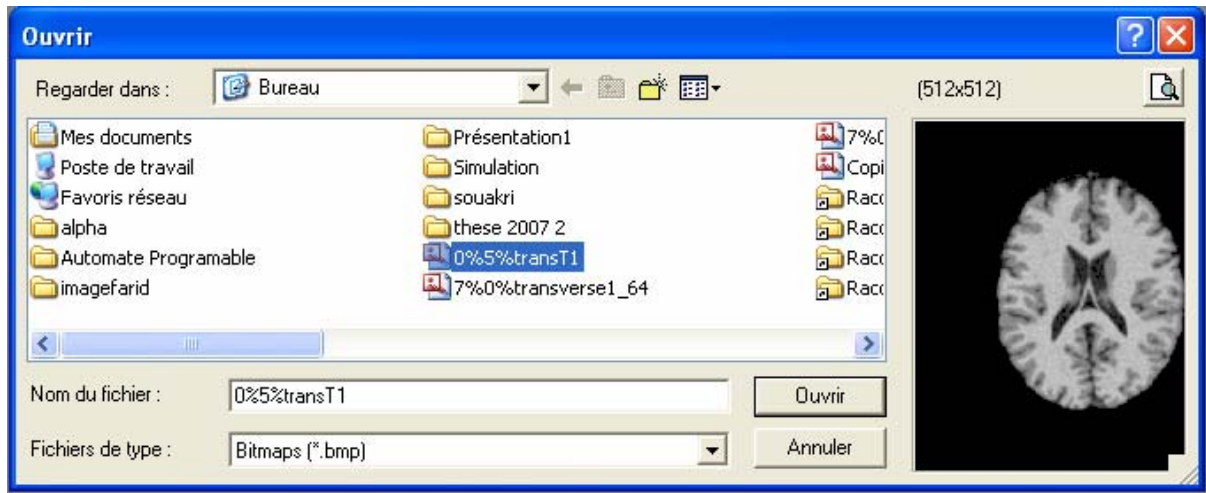
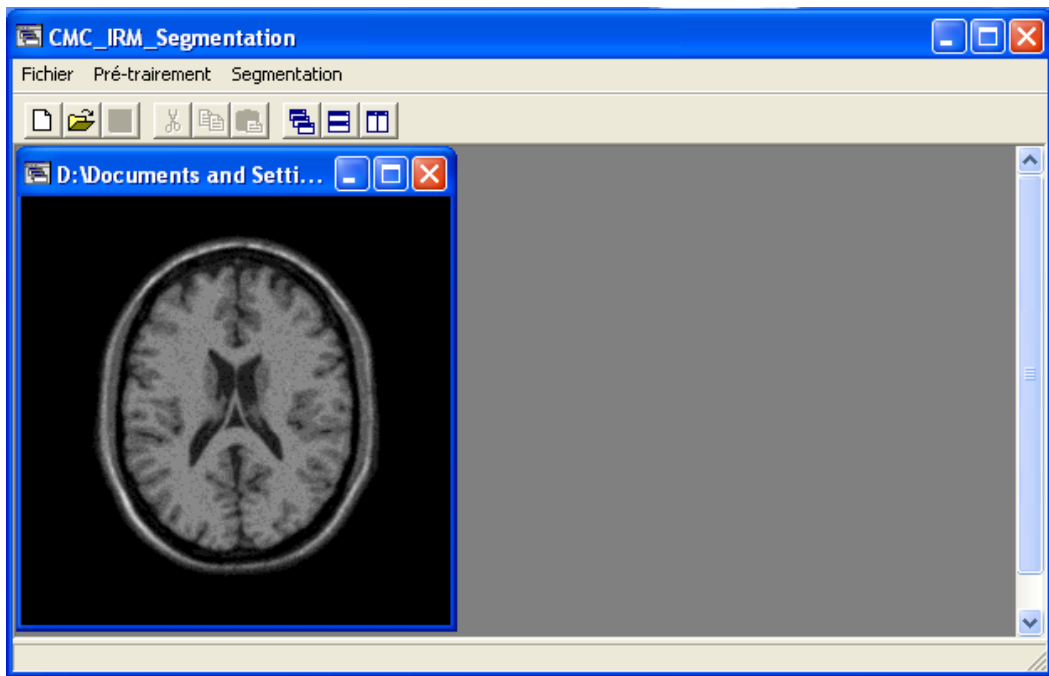


Figure 5.3 : Boîte de sélection de l'image à traiter



L'article Quitter permet de fermer les fenêtres filles et pour mettre fin à l'application on exécute l'article quitter.

3.1.2 Menu Pré-traitement

L'activation de cette commande permettra de réaliser les opérations :

- ✚ La normalisation des images (rendre carré $2^n \times 2^n$).
- ✚ Le filtrage.
- ✚ Parcours de Hilbert.
- ✚ L'histogramme de l'image.

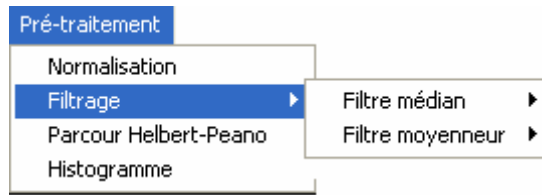


Figure 5.4 : Menu Pré-traitement

3.1.3 Menu Segmentation

L'activation de cette commande permettra de réaliser les opérations :

- ✚ Appliqué la méthode des K-moyennes pour l'initialisation de la chaîne de Markov.
- ✚ Appliqué la méthode des C-moyennes floue pour l'initialisation de la chaîne de Markov floue.
- ✚ La segmentation des images.

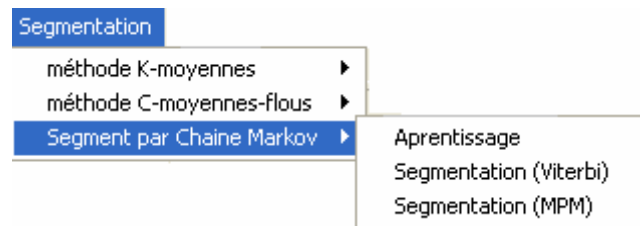


Figure 5.5 : Menu Segmentation

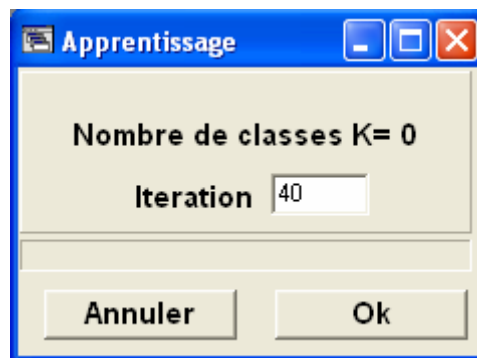


Figure 5.6 : Fenêtre d'apprentissage

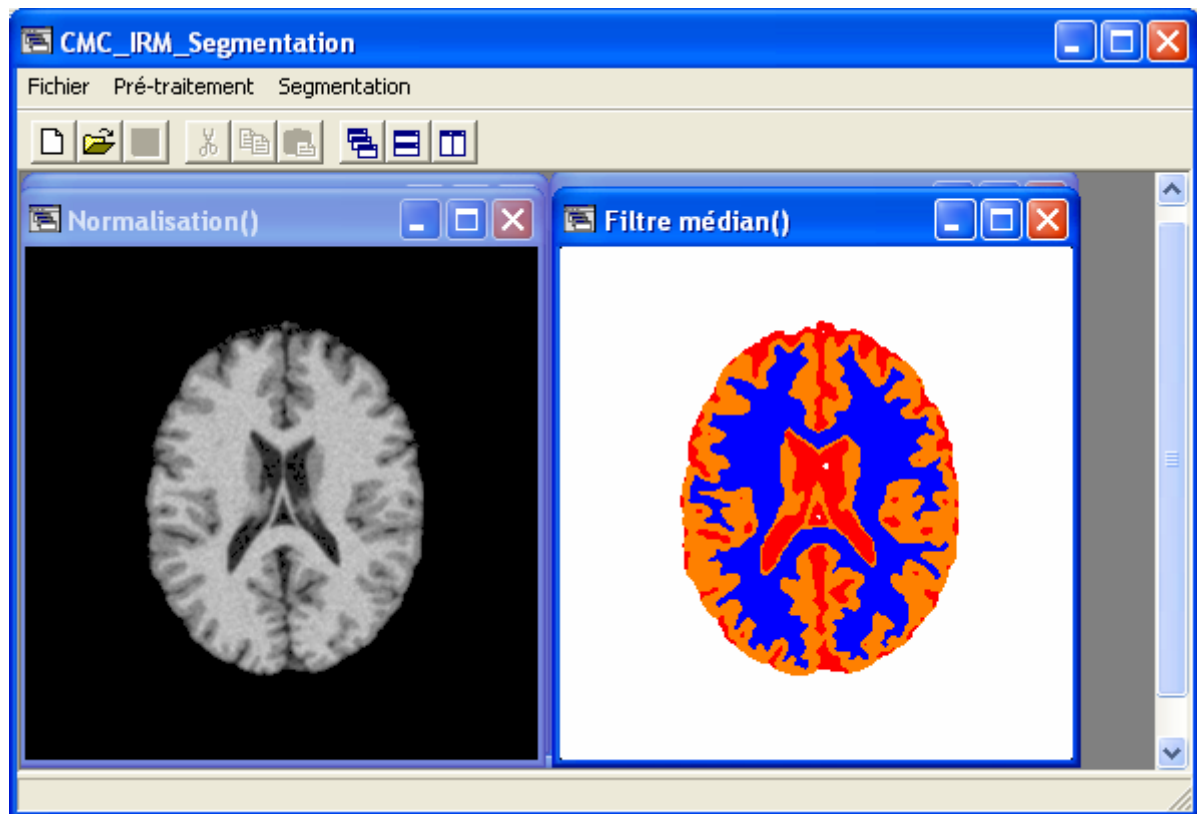


Figure 5.7 : Segmentation d'image 0% de bruit et 5% de non-uniformité en 4 classes et filtré par un filtre médian

Conclusion générale

L'objectif de notre travail concerne la segmentation des images médicales par résonance magnétique nucléaire, en vue de segmenter le cerveau en matière grise, matière blanche et liquide céphalo-rachidien.

Nous avons tout d'abord effectué une étude sur le principe physique de la formation des images IRM. La recherche bibliographique menée concernant les méthodes de segmentation d'images nous a informé sur la diversité des approches qui existent dont nous avons fait l'étude dans le chapitre 2. Nous avons ensuite fait une étude concernant l'approche markovienne « Chaîne de Markov Cachée ».

L'approche que nous avons mise en œuvre appartient à la famille des « approches statistiques globales ». Comme nous l'avons vu, tout traitement basé sur les chaînes de Markov débute par une transformation de l'image bruitée 2D en un vecteur 1D par le parcours de Hilbert, pour cela tous les traitements qui suivent seront effectués sur ce vecteur ; une fois la segmentation de la chaîne achevée, l'image thématique est finalement reconstruite à l'aide du parcours de Hilbert inverse. Dans le contexte de la segmentation non supervisée, les paramètres du modèle sont inconnus et doivent être estimés ; une méthode d'estimation ICE « Itérative Conditionnelle Estimation » a été utilisée grâce à son indépendance à l'initialisation (la méthode des K-moyennes ou C-moyennes floue) par rapport aux autres méthodes d'estimation telles que l'EM « Expectation Maximisation ». A la fin nous avons utilisé une classification bayésienne par le MPM.

En dernier lieu, nous avons effectué une évaluation des résultats obtenus. Nous avons présenté les résultats de la segmentation d'images fantômes et réelles. Une évaluation quantitative a été possible grâce à des images simulées ; nous avons pu tester l'approche pour des images de différents degrés de bruit et d'inhomogénéité. Nous avons également vu l'influence de la variation du nombre de classes **K** sur la qualité de la segmentation. L'approche que nous proposons donne des résultats meilleurs pour la segmentation des deux matières grise et blanche, le LCR néanmoins présente quelques imperfections en ce qui concerne les ventricules et sur la périphérie entre la matière blanche et les tissus environnant l'encéphale. La segmentation du LCR de la périphérie dépend étroitement du type de pondération utilisée. La mise en œuvre de l'approche markovienne illustre bien sa robustesse en présence de bruit et d'effet de volume partiel par un taux de reconnaissance important, de plus de 90 %.

En perspective à notre travail, l'approche markovienne peut être utilisée dans le cadre de la segmentation des images IRM 3D, grâce à la transformation de parcours de Hilbert 3D qui donne une issue pour le traitement d'images 3D avec un traitement parallèle.

Références bibliographiques

- [1] M. Cyril JAGGI, Segmentation par méthodes markovienne de l'encéphale humain en imagerie par résonance magnétique : Théorie, mise en œuvre et évaluation, Thèse Doctorat de l'université de CAEN, décembre 1998.
- [2] Nathalie Richard, Une approche située, coopérative et décentralisée pour l'inter-prétation d'images cérébrales par RMN, Thèse Doctorat de l' Institut Nationale Polytechnique de Grenoble de France, septembre 2004.
- [3] Stéphane Derrode, Wojciech Pieczynski, Signal and Image Segmentation Using Pairwise Markov Chains, IEEE transactions on signal processing, vol. 52, pp. 2477–2489, September 2004.
- [4] S. Derrode, G. Mercier, J. Le Caillec and R. Garello, Estimation of sea-ice SAR clutter statistics from Pearson's system of distributions, in IEEE Int, Conf. Geosci, Remote Sensing, Sydney, Australia, July 2001.
- [5] Thierry Géraud, Segmentation des structures internes du cerveau en imagerie par résonance magnétique, Thèse doctorat de l'Ecole nationale supérieure des télécommunications, juin 1998.
- [6] Nathalie Giordana, Wojciech Pieczynski, Estimation of Generalized Multisensor Hidden Markov Chains and Unsupervised Image Segmentation, IEEE transactions on pattern analysis and machine intelligence, vol. 19, No. 5, MAY 1997.
- [7] Btissam Benmiloud, Wojciech Pieczynski, Estimation des paramètres dans les chaînes de Markov cachées et segmentation d'images, Traitement du signal, Vol.12, N°5, pp 433-454, 1995.
- [8] Nicolas Brunel, Sur quelques extensions des chaînes de Markov cachées et couples. Applications à la segmentation non-supervisée de signaux radar, Thèse de Doctorat de l'université de Paris 6, décembre 2005.
- [9] Jeremy Michael Nett, The study of MS using MRI, Image processing and visualization. Thesis of the University of Louisville, Décembre 2001.
- [10] Germond Laurence, Trois principes de coopération pour la segmentation en imagerie de résonance magnétique cérébrale, Thèse de Doctorat de l'université Joseph Fourier, octobre 1999.

- [11] Haroun Radia, Segmentation des tissus cérébraux sur des images par résonance magnétique, Mémoire de Magister d'Université des Sciences et de la Technologie Houari Boumediene, Mai 2005.
- [12] Brahim Mohamed, Utilisation des ondelets dans la segmentation d'images médicales, Mémoire de Magister de l'Ecole Nationale Polytechnique, Janvier 2006.
- [13] Cyril Carincotte, Segmentation Markovienne floue d'images : Application en détection de changements entre images radar, Thèse de Doctorat de l'université Paul Cezanne, novembre 2005.
- [14] Benoît Scherrer, Segmentation Markovienne Distribuée et Coopérative des Tissus et Structures présents dans des IRM Cérébrales, Mémoire de Master Institut National Polytechnique de Grenoble, Juin 2005.
- [15] Mme Weibei Dou, Segmentation d'images multi spectrales basée sur la fusion d'informations : application aux images IRM, Thèse de Doctorat de l'université de CAEN, Septembre 2006.
- [16] Hafed Abdelhalim, Dahmoune Mohamed, Reconnaissance des visages à base des HMM Structurés, Mémoire d'Ingénieur d'Etat de Institut National de formation en Informatique, Alger Juin 2006.
- [17] Jonathan Lawder, The Application of Space-filling Curves to the Storage and Retrieval of Multi-dimensional Data, Thèse Doctorat de l'université de London, Décembre 1999.
- [18] Valery A. Petrushin, Hidden Markov Models: Fundamentals and Applications, Center for Strategic Technology Research publication, 2000.
- [19] Jalal Fadili, Bruno Moretti, Su Ruan, Daniel Bloyet, Méthode coopérative et quantification de la segmentation de l'encéphale en Imagerie par Résonance Magnétique, Laboratoire de Traitement du Signal de l'ISMRA-ENSI de Caen France – Volume 15 - n°5, décembre 1998.
- [20] W. Pieczynski. Pairwise Markov chains, IEEE Trans. Pattern Anal, Machine Intell., vol. 25, no. 5, pp. 634–639, 2003.
- [21] W. Pieczynski. Chaînes de Markov triplet, Comptes Rendus de l'Académie des Sciences - Mathématique, vol. 335, no. 3, pp. 275–278, 2002.
- [22] Jesús Angulo Lpóez, Morphologie Mathématique et indexation d'images couleurs : Application à la microscopie en biomédecine, Thèse de Doctorat de l'École des Mines de Paris, Décembre 2003.

- [23] R.K.-S. Kwan, A.C.Evans, G.B.Pike, An Extensible MRI Simulator for Post- Processing Evaluation, Visualization in Biomedical Computing (VBC'96), Lecture Notes in Computer Science, 1131, Springer-Verlag, pp. 135-140, 1996.
- [24] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, A.C. Evans, BrainWeb: Online Interface to a 3D MRI Simulated Brain Database, Neuro Image, vol.5, no.4, part 2/4, S425, 1997-- Proceedings of 3-rd International Conference on Functional Mapping of the Human Brain, Copenhagen, 1997.
- [25] D.L.Collins, A.P.Zijdenbos, V.Kollokian, J.G.Sled, N.J.Kabani, C.J.Holmes, A.C, Evans, Design and Construction of a Realistic Digital Brain Phantom, IEEE Transactions on Medical Imaging, 17(3), pp. 463-468, 1998.
- [26] R.K.-S. Kwan, A.C. Evans, G.B. Pike, MRI simulation-based evaluation of image processing and classification methods, IEEE Transactions on Medical Imaging, 18(11), pp. 1085-97, 1999.
- [27] A. Noe, J. C. Gee, Partial volume segmentation of cerebral MRI scans with mixture model clustering, In: B. Likar (ed.), Proceedings de Sixth Computer Vision Winter Workshop CVWW'01, Bled, Slovenija, pp.71-81, 2001.
- [28] J. Warning, Délinéation des aires visuelles rétinotopiques chez l'homme par IRM fonctionnelle, Thèse de doctorat de l'Université Joseph-Fourier Grenoble, 2002.
- [29] N. Richard, Une approche située, coopérative et décentralisée pour l'interprétation d'images cérébrales par RMN, Thèse de doctorat, institut Nationale Polytechnique de Grenoble, 2004.
- [30] W. Pieczynski, Modèles de Markov en traitement d'images, Traitement du Signal, vol. 20, no. 3, pp. 255-278, 2003.
- [31] O.Cappé, *Ten years of HMMs*, March 2001, URL <http://www.tsi.enst.fr/~cappe/docs/hmmbib.html>.
- [32] A. Cohen, Hidden Markov models in biomedical signal processing, in IEEE Int, Conf. Eng. Med. Biol, Society, vol. 3, (pp. 1145-1150), 29 Oct.-1 Nov, 1998.
- [33] K. Aas, L. Eikvil and R. Huseby, Applications of hidden Markov chains in image analysis, Pattern Recognition, vol. 32, no. 4, pp. 703-713, 1999.
- [34] S. Faisan, L. Thoraval, J.-P, Armspach and F. Heitz, Hidden semi-Markov event sequence models: application to brain functional MRI sequence analysis, in IEEE Int. Conf, Image Processing, Rochester, New York, USA, Sept 2002.

- [35] R. Fjørtoft, Y. Delignon, W. Pieczynski, M. Sigelle and F. Tupin, Unsupervised segmentation of radar images using hidden Markov chains and hidden Markov random fields, IEEE Trans. Geosci. Remote Sensing, 2003.
- [36] J. Warning, Délinéation des aires visuelles rétinotopiques chez l'homme par IRM fonctionnelle, Thèse de doctorat de l'Université Joseph-Fourier Grenoble, 2002.
- [37] N. Richard, Une approche située, coopérative et décentralisée pour l'interprétation d'images cérébrales par RMN, Thèse de doctorat, institut Nationale Polytechnique de Grenoble, 2004.
- [38] A. Noe, J. C. Gee, Partial volume segmentation of cerebral MRI scans with mixture model clustering, In: B. Likar (ed.), Proceedings de Sixth Computer Vision Winter Workshop CVWW'01, Bled, Slovenija, pp.71-81, 2001.