



HAL
open science

Longitudinal segmentation of MS lesions in multimodal brain MRI

Laure S. Aït-Ali, Sylvain Prima, Gilles Edan, Christian Barillot

► **To cite this version:**

Laure S. Aït-Ali, Sylvain Prima, Gilles Edan, Christian Barillot. Longitudinal segmentation of MS lesions in multimodal brain MRI. RFIA, University of Tours, Jan 2006, Tours, France, France. inria-00001160

HAL Id: inria-00001160

<https://inria.hal.science/inria-00001160>

Submitted on 22 Mar 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Segmentation longitudinale des lésions de SEP en IRM cérébrale multimodale

Longitudinal segmentation of MS lesions in multimodal brain MRI

L.S. Aït-Ali¹

S. Prima¹

G. Edan²

C. Barillot¹

¹ IRISA Rennes (INRIA, CNRS)

² CHRU Ponchaillou

Campus universitaire de Beaulieu

35042 Rennes Cedex

laitali@irisa.fr

Résumé

La sclérose en plaques (SEP) est une maladie inflammatoire auto-immune du système nerveux central. Des marqueurs cliniques sont utilisés aujourd'hui pour le diagnostic et pour l'évaluation thérapeutique. Cependant ces marqueurs sont subjectifs et sujets à une grande variabilité intra- et inter-observateur. L'imagerie par résonance magnétique (IRM) devient donc un examen complémentaire indispensable pour une meilleure compréhension de la maladie. Dans le but d'automatiser un processus long et fastidieux pour le clinicien, nous proposons de segmenter de façon automatique les lésions de SEP au cours du temps dans des séquences longitudinales d'IRM multimodales. Nous utilisons un algorithme robuste permettant la segmentation spatio-temporelle de ces anomalies et proposons un schéma original de rejet de données atypiques. Nous validons quantitativement nos résultats en utilisant le simulateur BrainWeb. Des résultats très prometteurs sont obtenus dans le cas de données cliniques.

Mots Clefs

Algorithme EM, segmentation robuste, séquence longitudinale, multimodalité, IRM, sclérose en plaques.

Abstract

Multiple sclerosis (MS) is a prototype inflammatory autoimmune disorder of the central nervous system. Today, clinical markers are used for the diagnosis as well as for the drugs assessment. However, these markers are subjective and show poor inter- and intra-rater reliability. Magnetic resonance imaging (MRI) becomes a mandatory surrogate exam for a better understanding of the disease. To make the lesions detection less fastidious and error prone, we propose to automatically segment MS lesions overtime in multidimodal magnetic resonance longitudinal sequences. We use a robust algorithm that allows the spatio-temporal segmentation of the abnormalities and

propose an original rejection scheme for outliers. We validate our method using the BrainWeb simulator. Very promising preliminary results on clinical data are shown.

Keywords

EM algorithm, robust segmentation, longitudinal sequence, multi-modality, MRI, multiple sclerosis.

1 Introduction

La sclérose en plaques (SEP), maladie évolutive de l'adulte jeune, est caractérisée par l'inflammation de la myéline et l'apparition de lésions sclérotiques dans la matière blanche du système nerveux central. Les causes de la maladie sont liées à des facteurs environnementaux et à des facteurs génétiques pour le moment mal identifiés. Les symptômes sont nombreux et variés : fatigue anormale, troubles de la vue, de l'équilibre, problèmes de coordination, tremblements... Parmi les manifestations de la maladie, on note une inflammation aiguë des axones, une destruction de la myéline, parfois un rétablissement de cette myéline, des glioses post-inflammatoires et une dégénérescence neuronale. Le déroulement séquentiel de ces processus est à la base de l'évolution clinique caractérisée par des épisodes avec rétablissement, des épisodes laissant des déficits définitifs, et des épisodes de progression secondaire. Les traitements ont pour but de réduire la fréquence et de limiter les effets des rechutes, des symptômes persistants et de prévenir la progression de la maladie. Malgré cela, la SEP reste une maladie complexe pour laquelle les solutions semblent encore lointaines. Durant ces dernières années, l'imagerie par résonance magnétique (IRM) a eu un impact majeur sur le diagnostic et la connaissance de la maladie. Elle permet notamment de suivre l'évolution de certaines structures ou lésions au cours du temps sans avoir recours à des explorations invasives. Par ailleurs, les éléments recueillis pour l'examen clinique sont à la fois très nombreux et très riches en informations, surtout depuis le développe-

ment des études multi-centriques. La quantité de données disponibles pour chaque patient devient alors de plus en plus grande, nécessitant soit plus de temps d'analyse, soit une sélection très restrictive des données à utiliser. L'informatique, qui permet d'automatiser en partie la tâche de traitement des informations, prend alors toute sa place dans l'univers médical moderne.

2 État de l'art

En IRM, le contraste est dominé par trois principales propriétés physiques : la densité de proton, les temps de relaxation longitudinale (T1) et transversale (T2). Le contraste peut être à dominante T1 en accentuant les différences dans les temps de relaxation T1 (pondération en T1). Le contraste peut être aussi à dominante T2 [1]. Les images appelées à *densité de protons* (DP) sont obtenues en minimisant les effets des contrastes T1 et T2. Ces différentes pondérations seront appelées modalités.

Plusieurs approches ont été développées pour automatiser le processus de segmentation des structures cérébrales en IRM. Une de ces approches consiste à modéliser les distributions d'intensité comme un mélange fini de gaussiennes. Chaque composante de ce mélange représente une structure cérébrale, typiquement, la matière blanche (MB), la matière grise (MG), et le liquide céphalo-rachidien (LCR). Un tel modèle de mélange de gaussiennes (MMG) permet de modéliser les intensités de l'image par un nombre réduit de paramètres. Ces paramètres peuvent être estimés dans un cadre statistique bien défini basé sur l'estimateur du maximum de vraisemblance (MV) en utilisant des méthodes d'optimisation telles que l'algorithme *Expectation - Maximization* (EM) [2]. Une telle approche est malheureusement inadaptée au cas pathologique caractérisé par des intensités anormales en IRM. Certains auteurs ont proposé des modifications à cette approche classique de façon à résoudre ce problème d'intensités atypiques comme celles causées par les lésions de SEP. Schroeter *et al.* [3] ont ajouté au MMG une composante uniforme pour modéliser les intensités des lésions. Van Leemput *et al.* [4] ont introduit une pondération reflétant le degré de "typicalité" de chacun des voxels¹. Ces derniers sont considérés comme des données atypiques si leur distance de Mahalanobis à chaque composante de mélange est supérieure à un seuil donné. Dugas-Phocion *et al.* [5] ont utilisé directement la distance de Mahalanobis au sein des itérations de l'EM pour le rendre plus robuste aux données atypiques. Ces méthodes sont développées pour utiliser plusieurs modalités d'images IRM mais pour un temps donné. Le suivi longitudinal des lésions n'est donc pas traité. D'un autre côté, d'autres méthodes proposent d'étudier l'évolution de la SEP au cours du temps en utilisant des tests statistiques ou des champs de déformations [6, 7, 8, 9]. En revanche, ces méthodes ne permettent pas la segmentation précise des lésions pour un temps donné. Shahar *et al.* [10] ont proposé une combinaison de ces deux approches pour la

segmentation spatio-temporelle des lésions de SEP. Ils effectuent conjointement la segmentation des lésions pour un temps d'acquisition donné et le suivi des lésions au cours du temps à partir d'images 2D pondérées en T2 en introduisant en plus des intensités, le paramètre temps ainsi que les coordonnées spatiales dans le modèle à estimer.

Nous proposons un algorithme original, robuste et multidimensionnel (plusieurs modalités et plusieurs temps d'acquisition) traitant des données $3D + t$. La première contribution est la modélisation originale du problème multidimensionnel et l'estimation robuste de ses paramètres via un schéma d'optimisation itératif. Nous obtenons alors une bonne segmentation des tissus sains. La seconde étape de notre travail est l'extraction des lésions de SEP comme étant des données atypiques du modèle précédemment établi. Une segmentation des tissus sains et des lésions de SEP est obtenue. Enfin, la dernière contribution consiste à améliorer la segmentation en utilisant l'information *a priori* provenant de la connaissance clinique sur l'apparence des lésions selon les différentes modalités d'IRM.

Le pré-traitement des données est expliqué dans la section 3. La méthode proposée est décrite dans la section 4, la validation et les résultats sont présentés dans la section 5. Nous discutons notre méthode et tirons des conclusions dans la section 6.

3 Pré-traitements

3.1 Extraction des structures cérébrales

Les IRM cérébrales contiennent l'image de la tête entière, or nous avons besoin de nous focaliser sur l'étude des tissus cérébraux seulement. Nous séparons donc ces derniers de la peau, de la graisse et de l'os du crâne. La méthode de contours actifs implicites (*level sets*) de Baillard *et al.* [11] est utilisée avec une initialisation automatique pour segmenter le cerveau.

3.2 Correction du biais

En IRM, les inhomogénéités dans le champ de radiofréquence appliqué lors de l'acquisition des images induisent des variations d'intensité au sein d'un même tissu. Ce *biais* d'intensité peut donc causer une mauvaise classification pour des techniques de segmentation fondées uniquement sur les intensités de l'image. L'algorithme de Prima *et al.* [12] est utilisé pour corriger ce biais.

3.3 Recalage géométrique

Comme nous l'avons expliqué précédemment, nous travaillons avec des modalités et des temps différents, nous devons donc aligner géométriquement les images traitées. Nous utilisons l'algorithme de recalage rigide de Maes *et al.* [13] basé sur la maximisation de l'information mutuelle. Toutes les images (différentes modalités, différents temps), corrigées en biais, sont recalées de façon rigide vers une image de référence de telle sorte que chaque voxel représente le même point anatomique pour chaque image.

1. pixels 3D

3.4 Normalisation d'intensité

Nous appliquons une méthode de classification basée intensité, or nous travaillons sur plusieurs temps et, il est possible que pour une même modalité, les intensités moyennes des tissus varient au cours du temps. En effet, le calibrage de la machine peut changer un petit peu entre deux temps d'acquisition. Nous modélisons ces changements par une fonction polynomiale de bas degré (inférieur ou égale à trois) dont les coefficients sont calculés de façon robuste via un M-estimateur.

4 Méthode

Nous rappelons dans la section Sec. 4.1 le principe de l'algorithme EM pour des données multivariées. Dans la section Sec. 4.2, nous proposons une modélisation originale du problème multidimensionnel puis nous présentons notre algorithme robuste estimant les paramètres de ce modèle. Enfin dans la section Sec. 4.3, nous expliquons comment nous segmentons les lésions de SEP.

4.1 Une méthode classique de segmentation multimodale paramétrique : l'EM

Modélisation. Le bruit d'acquisition d'une IRM suit une distribution de Rice, qui peut être assez bien approchée par une distribution gaussienne [14]. Supposons que la distribution des intensités de chaque structure cérébrale est aussi bien approchée par une loi gaussienne, il est alors usuel de considérer que les intensités de l'image sont générées par un mélange de gaussiennes. Si on considère m différentes modalités simultanément, chaque voxel i , $i \in \{1, \dots, n\}$ est décrit par un vecteur d'intensité m -dimensionnel $\mathbf{y}_i = (y_i^1, \dots, y_i^m)$. Ces vecteurs sont vus comme les réalisations d'une variable aléatoire suivant un modèle de mélange de gaussiennes m -dimensionnel défini comme suit :

$$p(\mathbf{y}_i; \theta) = \sum_{j=1}^k \alpha_j (2\pi)^{-m/2} |\Sigma_j|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{y}_i - \mu_j)^T \Sigma_j^{-1} (\mathbf{y}_i - \mu_j)\right), \quad (1)$$

où k est le nombre de composantes du modèle. $\theta = (\alpha, \mu, \Sigma)$ est l'hyper-paramètre du modèle :

- $(\alpha_j)_{j \in \{1, \dots, k\}}$ sont les proportions de mélange,
- $(\mu_j)_{j \in \{1, \dots, k\}}$ sont les vecteurs de moyennes,
- $(\Sigma_j)_{j \in \{1, \dots, k\}}$ sont les matrices de covariance.

Pour les IRM cérébrales, trois structures ont des intensités relativement homogènes et constituent les composantes du modèle gaussien : la matière blanche (MB), la matière grise (MG) et le liquide céphalo-rachidien (LCR). Le but est de donner une étiquette x_i à chaque voxel i , avec $x_i \in \{c_1, \dots, c_k\}$, c_j étant une des trois composantes. L'image résultante $X = (x_1, \dots, x_n)$ est l'image de segmentation des données $Y = (\mathbf{y}_1, \dots, \mathbf{y}_n)$.

Pour des raisons d'efficacité algorithmique, nous traitons l'histogramme joint des données multimodales construit à partir de ces vecteurs \mathbf{y}_i .

Estimation. Une fois le problème modélisé, il faut en calculer les paramètres si on veut pouvoir attribuer une étiquette à chaque voxel. Ce calcul est basé sur l'estimation au sens du maximum de vraisemblance de l'hyper-paramètre θ du modèle :

$$\hat{\theta} = \arg \max_{\theta} \prod_{i=1}^n p(\mathbf{y}_i; \theta). \quad (2)$$

Les vecteurs \mathbf{y}_i sont supposés indépendants.

Nous sommes face à un problème dit de *données incomplètes* puisque ni les étiquettes x_i des voxels ni les paramètres des mélanges de gaussiennes ne sont connus. Seuls les vecteurs d'intensité \mathbf{y}_i sont connus.

Optimisation. L'algorithme *Expectation - Maximization* (EM) [2] est une méthode efficace pour résoudre ce type de problème. Cet algorithme est basé sur un schéma itératif en deux étapes.

- Étape (E) : Calcul de l'espérance conditionnelle de la vraisemblance des données complètes étant données les valeurs observées Y et une valeur *a priori* de l'hyper-paramètre $\theta^{(q-1)}$.
- Étape (M) : Calcul de l'hyper-paramètre $\theta^{(q)}$ maximisant l'espérance obtenue dans l'étape (E).

L'étape (E) revient à calculer la probabilité pour un voxel i d'appartenir à c_j sachant les données Y et l'hyper-paramètre $\theta^{(q-1)}$:

$$P(x_i = c_j | Y; \theta^{(q-1)}) = \frac{\alpha_j^{(q-1)} p(\mathbf{y}_i | x_i = c_j; \theta^{(q-1)})}{\sum_{k=1}^k \alpha_k^{(q-1)} p(\mathbf{y}_i | x_i = c_k; \theta^{(q-1)})} \quad (3)$$

Puis, à l'étape M, les paramètres des classes et les proportions de mélange sont estimés en maximisant la vraisemblance des données :

$$\alpha_j^q = \frac{1}{\text{card}(Y)} \sum_i P(c_i = j | Y, \theta^{(q-1)}) \quad (4)$$

$$\mu_j^q = \frac{\sum_i P(c_i = j | Y, \theta^{(q-1)}) \mathbf{y}_i}{\sum_i P(c_i = j | Y, \theta^{(q-1)})} \quad (5)$$

$$\Sigma_j^q = \frac{\sum_i P(c_i = j | Y, \theta^{(q-1)}) (\mathbf{y}_i - \mu_j^q)(\mathbf{y}_i - \mu_j^q)^T}{\sum_i P(c_i = j | Y, \theta^{(q-1)})} \quad (6)$$

Ces deux étapes sont itérées jusqu'à convergence vers un maximum, au moins local, de la vraisemblance des données.

Classification. Après convergence, nous obtenons une estimation $\hat{\theta}$ de l'hyper-paramètre. Les probabilités finales $P(x_i = c_j | Y; \hat{\theta})$ peuvent donc être calculées. On attribue ensuite à chaque vecteur \mathbf{y}_i l'étiquette de la classe c_j de plus grande probabilité.

4.2 L'EM robuste spatio-temporel

Le modèle spatio-temporel. Pour le suivi du patient et l'étude de la maladie, l'analyse longitudinale est essentielle. Dans ce contexte, nous proposons de développer une méthode de segmentation multimodale pour plusieurs temps d'acquisition $(1, \dots, t)$ avec une approche robuste. L'ensemble des données peut être noté $Y = \{Y^{(1)}, \dots, Y^{(t)}\}$, où $Y^{(l)} = (\mathbf{y}_1^{(l)}, \dots, \mathbf{y}_n^{(l)})$ est l'ensemble des données multimodales au temps d'acquisition l . Une fois les images pré-traitées comme indiqué Section 3, on construit pour chaque temps un histogramme joint m -dimensionnel. Ensuite, chacun de ces t histogrammes joints sont combinés en un seul histogramme joint global contenant les informations de toutes les modalités et tous les temps d'acquisition.

L'estimateur de vraisemblance tamisée. Pour que la classification soit moins sensible au bruit, aux erreurs de recalage, aux effets de volume partiel et surtout aux anomalies pathologiques (*e.g.* les lésions de SEP) qui ne sont pas bien gérés par le modèle de mélange de gaussiennes, nous proposons de remplacer l'estimateur du maximum de vraisemblance (MV) par l'estimateur de vraisemblance tamisée² (TLE) proposé par Neykov *et al.* [15]. L'hyper-paramètre optimal est celui qui maximise le TLE des données :

$$\tilde{\theta} = \arg \max_{\theta} \prod_{i=1}^h p(\mathbf{y}_{\nu(i)}; \theta) \quad (7)$$

où $\nu(i)$ sont les permutations de $i = \{1, \dots, n\}$ telles que $r_{\nu(i)} \leq r_{\nu(i+1)}$, les $r_i = -\log p(\mathbf{y}_i; \theta)$ étant les résidus. Le TLE est à l'estimateur MV ce que le LTS³ est au critère des moindres carrés (LS⁴). Ainsi le TLE permet d'éliminer les données atypiques du calcul de l'hyper-paramètre caractérisant les tissus sains.

Optimisation. Nous proposons un schéma itératif pour estimer $\tilde{\theta}$. La preuve de la convergence de ce schéma, au moins vers un minimum local du critère, a été démontrée par Neykov *et al.* [15].

On peut résumer le schéma comme suit :

1. initialisation aléatoire de l'hyper-paramètre $\tilde{\theta}^0$,
2. maximisation de la vraisemblance de toutes les données en utilisant un EM, donnant une première estimation de l'hyper-paramètre $\tilde{\theta}^{(q-1)}$,

$$\tilde{\theta}^{(q-1)} = \arg \max_{\theta} \prod_{i=1}^n p(\mathbf{y}_i; \theta),$$

3. tri des résidus $r_i = -\log p(\mathbf{y}_i; \tilde{\theta}^{(q-1)})$,
4. maximisation de la vraisemblance pour les vecteurs ayant les h plus petit résidus ($h > 50\%$ des données, typiquement nous avons choisi $h > 90\%$ données)

donnant une nouvelle estimation $\tilde{\theta}^{(q)}$ (toujours en utilisant un EM),

5. retour à l'étape 3 jusqu'à convergence. La convergence est considérée comme atteinte lorsque les proportions de mélange n'évoluent plus à 10^{-9} près.

Classification. Finalement, l'ensemble des images des segmentations $X = (X^1, \dots, X^t)$, avec $X^l = (x_i^l, x_i^l \in \{c_1, \dots, c_k\})$ la segmentation pour l'acquisition l , est construit selon les vecteurs originaux \mathbf{y}_i et leur étiquette x_i . L'étiquette x_i attribuée au voxel i est celle donnant la plus grande probabilité *a posteriori* Eq.(3) une fois l'hyper-paramètre θ estimé. Les images X^1, \dots, X^t sont les segmentations des tissus sains pour les différents temps d'acquisition. Les lésions ne sont pas extraites par cet algorithme puisqu'il n'y a aucune preuve que leur intensité suit une distribution gaussienne. Nous appelons cet EM robuste spatio-temporel STREM⁵.

4.3 Segmentation des lésions

Une première approche intuitive. Nous pouvons tout simplement segmenter les lésions comme étant les voxels ayant les $(n - h)$ résidus les plus grands après convergence de l'algorithme. C'est à dire les $(n - h)$ données correspondant le moins au modèle de tissus sains estimé. On attribue à ces données l'étiquette c_{k+1} . Les autres voxels sont classés comme dans STREM. Nous notons cette méthode de segmentation des tissus sains et des lésions STREM1. Mais la valeur de h dépend alors de la proportion de lésions pour un patient donné et n'est pas généralisable. Pour remédier à ce problème de seuil *ad hoc*, nous proposons deux raffinements.

Extraction des lésions. Une fois le modèle spatio-temporel établi de façon robuste par STREM, nous cherchons à en extraire les données atypiques. Pour chaque voxel $i \in \{1, \dots, n\}$, la distance de Mahalanobis entre le vecteur \mathbf{y}_i et chaque classe c_j est :

$$d_{i,j} = (\mathbf{y}_i - \mu_j)^T \Sigma_j^{-1} (\mathbf{y}_i - \mu_j).$$

Cette distance mesure le degré d'appartenance de i au modèle. Pour des données multivariées m -dimensionnelles suivant une loi gaussienne, la valeur de la distance de Mahalanobis $d_{i,j}$ suit une loi du *chi* 2 à m degrés de liberté (χ_m^2). Donc si cette distance, pour un vecteur \mathbf{y}_i donné est supérieure à la valeur critique de la distribution du χ^2 pour toutes les classes, alors le vecteur n'est pas représentatif du modèle et doit être considéré comme une donnée atypique. Les données atypiques ainsi extraites sont étiquetées c_{k+1} , les autres sont classées comme dans STREM. Cette méthode de segmentation des lésions et des tissus sains est appelée STREM2.

Introduction de connaissances *a priori*. Nous avons choisi un seuil de Mahalanobis très sensible de façon à ce qu'aucune lésion ne soit oubliée. Cette forte sensibilité

2. En anglais : *Trimmed Likelihood Estimator* (TLE)

3. Least Trimmed Squares

4. Least Squares

5. Spatio-Temporal Robust Expectation-Maximization

m	$s_{p=0.5}$	$s_{p=0.1}$	$s_{p=0.05}$	$s_{p=0.01}$
1	0.4549	2.7055	3.8414	6.6349
2	1.3862	4.6051	5.9914	9.2103
3	2.3659	6.2513	7.8147	11.3449
4	3.3567	7.7794	9.4877	13.2767

TAB. 1 – Valeurs critiques s de la distribution du χ^2 selon le nombre de degré de liberté m et la valeur de la p_value . Si on souhaite avoir une probabilité d'appartenance au modèle de 95%, ce qui correspond à une $p_value = 0.05$, alors le seuil s sur la distance de Mahalanobis est 3.8414 si on utilise une modalité d'IRM, 5.9914 pour deux modalités différentes d'IRM, etc.

peut entraîner le rejet de données atypiques comprenant les effets de volume partiel, le bruit et les vaisseaux sanguins. De façon à discriminer les lésions réelles des faux positifs, nous apportons des connaissances *a priori* sur la SEP et son incidence sur les intensités de l'IRM. Il s'avère en effet que les lésions de SEP apparaissent en hypo- ou hypersignal par rapport à la MB selon les modalités d'IRM. La combinaison de ces différentes informations nous permet de définir des "vecteurs-types" propres aux différents types de lésions. Les images pondérées en T2 et DP mettent en évidence la composante myélinique dans les lésions caractérisées par des œdèmes qui apparaissent comme hyper-intenses par rapport à la matière blanche. La double validation grâce à la combinaison de ces deux modalités permet aux cliniciens de supprimer de nombreux artefacts. Par ailleurs, la modalité T1 met en évidence la destruction irréversible des tissus avec l'apparition dans la matière blanche de "trous noirs" (hypo-signal) persistants. Le gadolinium (Gd) est un produit de contraste paramagnétique qui réduit fortement le temps de relaxation T1 (apparition d'un hypersignal). Cette substance caractérise la destruction de la barrière hémato-encéphalique. Cette destruction est caractéristique de la phase inflammatoire accompagnant l'apparition de nouvelles lésions. Cette connaissance *a priori* est synthétisée dans la Tab.1, où -1, +1 et 0 dénotent respectivement un hypo-signal, un hyper-signal, et une apparence normale par rapport à la MB.

	T1	T1-Gd	T2	DP
lésions jeunes	0	+1	+1	+1
trous noirs	-1	0	+1	+1
œdèmes autres lésions	0	0	+1	+1

TAB. 2 – Connaissances *a priori* sur la SEP en IRM. Les lésions dites "jeunes" (4 à 6 semaines) apparaissent comme un hyper-signal sur les images pondérées en T1 avec injection de produit de contraste (T1-Gd). Les trous noirs apparaissent plus ou moins sombres sur les images pondérées en T1. Les autres lésions (ni jeunes, ni caractéristiques de destructions cellulaires) sont hyper-intenses par rapport à la MB sur les images pondérées en T2 et DP.

Les voxels considérés comme données atypiques par STREM2 et ayant été classifiés comme lésions en utilisant les connaissances *a priori* prennent alors la valeur de l'étiquette c_{k+1} , les autres voxels prennent l'étiquette maximisant la probabilité *a posteriori* Eq.(3). Cette méthode est notée STREM3.

L'avantage d'une telle analyse multi-dimensionnelle est double. D'un point de vue technique, l'utilisation d'un maximum de données permet d'augmenter le rapport signal sur bruit et donc d'obtenir une segmentation plus précise. D'un point de vue clinique, nous utilisons l'information des différentes modalités ce qui nous permet de mettre en évidence les différents types de lésions. Nous utilisons l'information des différentes acquisitions temporelles, ce qui nous permet de pouvoir étudier l'évolution des lésions au cours du temps et donc aider à l'étude de la maladie et au suivi des patients.

La méthode complète pour la segmentation des lésions de SEP est résumée dans le schéma Fig.1.

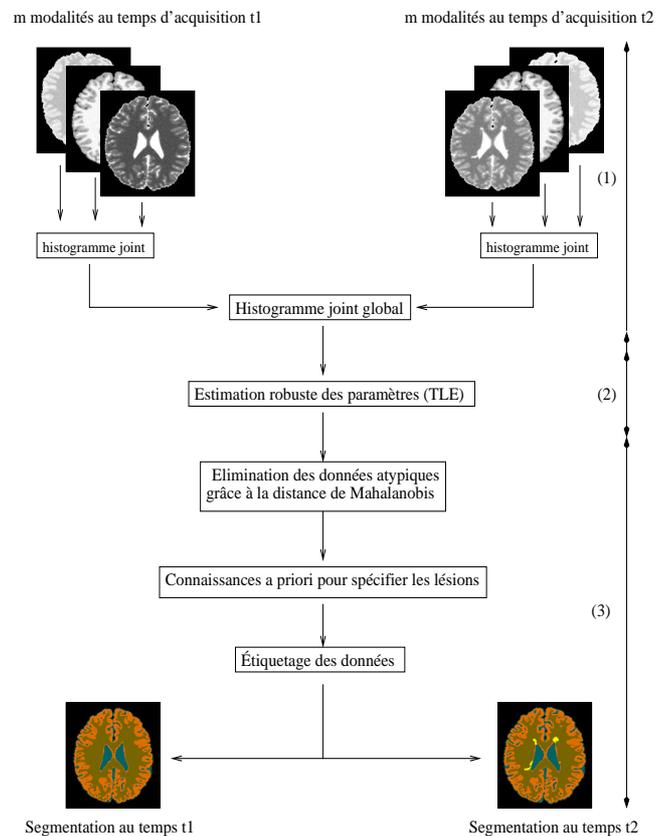


FIG. 1 – Méthode proposée pour la segmentation spatio-temporelle robuste d'IRM 3D, exemple pour deux instants t_1 et t_2 . Étape 1 : modélisation spatio-temporelle des données et construction d'un histogramme joint global utilisant tous les temps d'acquisition. Étape 2 : estimation robuste des paramètres du modèle via un schéma d'optimisation itératif. Étape 3 : classification des tissus sains et des lésions.

5 Validation et résultats

5.1 Validation quantitative sur les données simulées par BrainWeb

Les données. BrainWeb est une base d'images synthétiques réalistes 3D générées par un simulateur d'IRM utilisant les premiers principes des équations de Bloch [16]. Pour chaque type de tissu (MB, MG, LCR, graisse, ...), le modèle anatomique utilisé est constitué d'un ensemble de volumes contenant les degrés d'appartenance au tissu. Les intensités des voxels dans ces volumes représentent la proportion de tissu présent dans chaque voxel dans un intervalle $[0, 1]$. Nous choisissons de simuler deux temps, la base A et la base B , contenant chacune trois modalités d'images : images pondérées en T1, en T2 et DP. Les volumes générés ont une épaisseur de coupe de 1 mm et un niveau de bruit de 3%. La base A , première acquisition considérée, ne contient pas de lésions de SEP, la base B , seconde acquisition, contient des lésions modérées, péri-ventriculaires notamment.

La mesure utilisée. Le coefficient de similarité utilisé pour évaluer la méthode est le DSC⁶. Le DSC mesure le recouvrement spatial entre deux segmentations R_1 et R_2 . Il est défini de la façon suivante :

$$DSC(R_1, R_2) = 2 \cdot \text{card}(R_1 \cap R_2) / \text{card}(R_1 + R_2).$$

Il est démontré que cette mesure est un cas particulier de la statistique *kappa* communément utilisée en analyse de fiabilité [17]. Ce coefficient de similarité est calculé entre les régions segmentées par notre algorithme et les régions définies par la vérité terrain donnée par BrainWeb.

Dans un premier temps, nous démontrons l'amélioration apportée par l'utilisation de plusieurs modalités. Ensuite, nous mettons en évidence l'utilité de l'étude longitudinale. Pour conclure, nous montrons l'impact de nos raffinements.

STREM vs. EM : base A. La segmentation jointe des images pondérées en T2 et DP montre la pertinence de la méthode robuste. En effet, l'image de DP est très peu contrastée et l'EM classique n'arrive pas à différencier la MB de la MG, même avec les informations apportées par l'image pondérée en T2. Alors que la méthode robuste que nous utilisons permet de pallier ce problème (Fig. 2).

Impact de l'analyse multimodale : base A. Nous avons vu que l'analyse multimodale est très intéressante d'un point de vue clinique pour étudier les différents types de lésions de SEP. Nous nous intéressons maintenant à l'impact de l'analyse multimodale sur l'estimation du modèle des tissus sains. Nous étudions donc des IRM normales, sans lésions de SEP (base A). Nous montrons empiriquement que l'utilisation de plusieurs modalités améliore la segmentation grâce à trois expériences :

- (i) STREM monomodale sur une image pondérée en T2 donnant une image de segmentation notée $X_{T_2}^A$,

6. Dice Similarity Coefficient

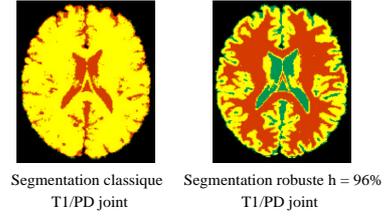


FIG. 2 – Apport de l'algorithme robuste par rapport à l'EM classique. L'EM classique n'est pas en mesure de dissocier la MB et la MG, le contraste étant trop peu prononcé. L'algorithme robuste pallie ce problème.

et STREM sur une image DP donnant une image de segmentation notée X_{DP}^A ;

- (ii) STREM multimodale T2/DP donnant une image de segmentation notée $X_{T_2, DP}^A$;
- (iii) STREM multimodale T1/T2/DP donnant une image de segmentation notée $X_{T_1, T_2, DP}^A$.

Nous avons choisi $h = 96\%$ des données et les images de segmentation obtenues ont été reconstruite en attribuant une des trois étiquettes à chaque voxel.

Nous comparons le DSC de la MB et de la MG (le LCR n'étant pas bien segmenté dans X_{DP}^A) pour les segmentations monomodales $X_{T_2}^A$ et X_{DP}^A avec les DSC de la segmentation multimodale $X_{T_2, DP}^A$. Le fait de traiter conjointement les deux modalités augmente l'indice de similarité. Pour compléter cette étude, l'ajout d'une troisième modalité améliore encore le DSC (Tab.2).

	MB	MG
$X_{T_2}^A$	0.86	0.84
X_{DP}^A	0.88	0.83
$X_{T_2, DP}^A$	0.93	0.88
$X_{T_1, T_2, DP}^A$	0.95	0.89

TAB. 3 – Comparaison des DSC entre l'étude sur une modalité et l'étude conjointe de plusieurs modalités. La segmentation est améliorée quantitativement en utilisant plusieurs modalités pour l'estimation de l'hyper-paramètre du modèle.

Cette amélioration se note sur les images de classification. L'étude de l'image DP seule ne donne pas une image de segmentation correcte, le contraste étant trop peu prononcé sur les images utilisées. En utilisant l'information de la modalité T2 conjointement à la DP, la segmentation est bien meilleure (cf. Fig. 3).

L'analyse multimodale améliore donc l'estimation du modèle pour les tissus sains. Pour l'analyse de la SEP, l'utilisation des différentes modalités est primordiale car elles apportent des informations complémentaires.

Impact de l'analyse longitudinale : base A et B. L'analyse longitudinale permet de détecter à la fois les lésions de SEP et leur évolution. L'étude de cette évolution est

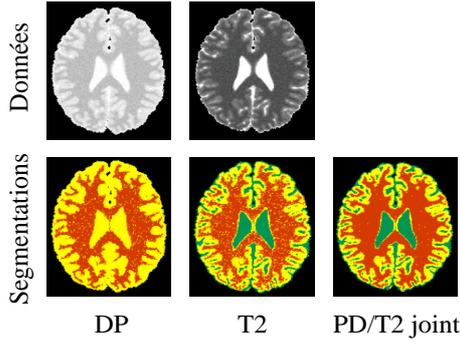


FIG. 3 – Impact de l’analyse multimodale. L’image de DP seule n’est pas bien segmentée. L’image pondérée en T2 seule est mieux segmentée mais il reste quelques artefacts dans la MB (les points jaunes dans le rouge). Mais lorsque l’on utilise ces deux modalités conjointement l’image de segmentation jointe $X_{T_2, DP}^A$ obtenue est bien meilleure que X_{DP}^A et $X_{T_2}^A$.

très importante pour le suivi des patients, pour l’évaluation de thérapie, pour la connaissance de la maladie. La segmentation pour une acquisition donnée est utilisée pour améliorer le suivi des lésions au cours du temps. Inversement, toute l’information de la séquence temporelle est utilisée pour améliorer la segmentation des lésions pour une acquisition donnée. Comme nous l’avons vu dans le paragraphe précédent, l’utilisation de plusieurs modalités améliore la segmentation. Pour cette seconde validation, nous conservons les trois modalités mais en plus nous simulons une séquence longitudinale contenant deux temps d’acquisition. Le premier temps considéré (base A) comporte les trois modalités sans lésions $Y^{(A)} = Y_{T_1, T_2, DP}^A$. Le second temps d’acquisition (base B) est représenté par l’ensemble des images simulées 3D avec lésions $Y^{(B)} = Y_{T_1, T_2, DP}^B$. On simule donc une apparition de lésions au cours du temps. Les expériences suivantes sont effectuées :

- (i) STREM2 multimodale sur les données $Y^{(A)}$ donnant l’image $X_{STREM2}^{(A)}$,
- (ii) STREM2 multimodale sur les données $Y^{(B)}$ donnant l’image $X_{STREM2}^{(B)}$,
- (iii) STREM2 multidimensionnelle sur les données traitées conjointement $Y = (Y^{(A)}, Y^{(B)})$ donnant les images $X_{STREM2} = (X_{STREM2}^{(A/B)}, X_{STREM2}^{(B/A)})$.

$X^{(A/B)}$, réciproquement $X^{(B/A)}$, est le résultat de la segmentation de la base A , réciproquement B , calculée en utilisant conjointement la base A et la base B . On prend $h = 96\%$ des données pour estimer le modèle de tissus sains. Ensuite les données atypiques sont rejetées en utilisant la distance de Mahalanobis pour une valeur critique correspondant à une $p_value = 0.05$. Les voxels ayant été rejetés sont attribués à la classe c_{k+1} des lésions.

On compare les DSC entre la segmentation STREM2 pour un temps d’acquisition donné et la segmentation STREM2

d’une séquence longitudinale. L’analyse multidimensionnelle améliore le résultat de la segmentation (voir Tab.4).

	MB	MG	LCR	lésions
(i) $X_{STREM2}^{(A)}$	0,95	0,89	0,56	-
(iii) $X_{STREM2}^{(A/B)}$	0,95	0,95	0,65	-
(ii) $X_{STREM2}^{(B)}$	0,95	0,91	0,59	0,61
(iii) $X_{STREM2}^{(B/A)}$	0,95	0,91	0,59	0,72

TAB. 4 – Comparaison des coefficients de similarité entre l’analyse spatiale seule et l’analyse spatio-temporelle. L’analyse longitudinale améliore quantitativement la segmentation des lésions ($h = 96\%$ des données, $p_value = 0.05$) et des tissus sains (notamment pour la base A).

Segmentation des lésions : base A et B . Les données utilisées sont les bases A et B conjointement. Nous faisons les trois expériences suivantes :

- (i) STREM1 pour $Y = (Y^{(A)}, Y^{(B)})$ avec $h = 96\%$ donnant $X_{STREM1} = (X_{STREM1}^{(A/B)}, X_{STREM1}^{(B/A)})$,
- (ii) STREM2 pour $Y = (Y^{(A)}, Y^{(B)})$ avec $h = 96\%$, $p_value = 0.05$, on obtient $X_{STREM2} = (X_{STREM2}^{(A/B)}, X_{STREM2}^{(B/A)})$,
- (iii) STREM3 pour $Y = (Y^{(A)}, Y^{(B)})$ avec $h = 96\%$, $p_value = 0.05$, on obtient $X_{STREM3} = (X_{STREM3}^{(A/B)}, X_{STREM3}^{(B/A)})$.

Dans le cas de l’expérience (i), nous avons étiqueté comme lésions les données ayant les $(n - h)$ plus grand résidus. Certaines de ces données atypiques ne sont pas des lésions mais du bruit (cf. Fig. 4, première ligne) notamment au niveau des ventricules (LCR interne, en bleu sur la figure), ce qui a pour effet de diminuer la qualité de la segmentation des lésions. Nous avons trop de faux positifs, le DSC de l’image de segmentation $X_{STREM1}^{(B/A)}$ n’est donc pas très bon (cf. Tab.5). On pourrait alors décider de retirer moins de données atypiques mais alors le modèle serait moins bien estimé.

Dans le cas de l’expérience (ii), une fois le modèle estimé de façon robuste, nous excluons les données pour lesquelles la distance de Mahalanobis à chacune des classes de tissu sain est supérieure au seuil du χ^2 . Dans cet exemple nous prenons une $p_value = 0.05$ de façon à être suffisamment sensible pour détecter toutes les lésions (le taux de vrais positifs *v. p.* est 95%). Nous obtenons alors une classe de lésions plus spécifique (cf. Fig. 4, deuxième ligne). Cependant il reste encore des artefacts dans les ventricules, hors on sait que anatomiquement les lésions ne peuvent être dans le LCR. Nous pourrions alors avoir un seuil plus spécifique pour la distance de Mahalanobis mais nous risquons d’éliminer des vrais positifs alors qu’il est très important pour le clinicien que toutes les lésions ainsi que leur évolution soient détectées.

Nous introduisons donc, pour différencier les lésions des artefacts, des informations *a priori* sur la SEP et son influence sur les intensités d'IRM. Parmi les vecteurs d'intensité \mathbf{y}_i dont la distance de Mahalanobis dépasse la valeur critique du χ^2 pour les trois classes de tissus sains, certains vérifient les règles répertoriées Tab.1 et seront segmentés comme des lésions (étiquette c_{k+1}). Les autres vecteurs d'intensité, \mathbf{y} compris ceux qui avaient été initialement exclus en utilisant la distance de Mahalanobis, sont attribués à la classe de plus grande probabilité.

	DSC MB	DSC MG	DSC LCR	DSC lésions	v. p. lésions
$X_{STREM1}^{(B/A)}$	0.95	0.90	0.58	0.58	0.96
$X_{STREM2}^{(B/A)}$	0.95	0.90	0.58	0.72	0.95
$X_{STREM3}^{(B/A)}$	0.95	0.90	0.58	0.73	0.95

TAB. 5 – DSC pour la base B contenant des lésions. En utilisant seulement le schéma robuste pour extraire les lésions, soit nous basons le choix de h soit nous obtenons une mauvaise segmentation des lésions. En utilisant la distance de Mahalanobis pour extraire les lésions une fois le modèle de tissus sain estimé de façon robuste, la segmentation des lésions est améliorée, notamment au niveau de la spécificité.

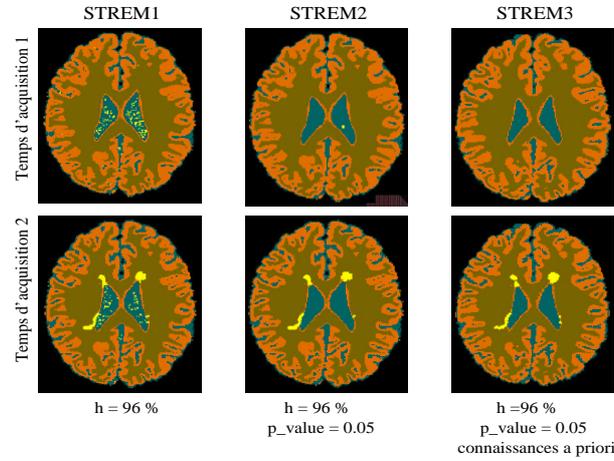


FIG. 4 – Apport des raffinements pour la segmentation robuste multidimensionnelle des lésions de SEP. Beaucoup de faux positifs apparaissent si le schéma robuste seul est utilisé pour extraire les lésions (points jaunes au niveau des ventricules bleus). Si on utilise la distance de Mahalanobis pour extraire les lésions une fois le modèle de tissus sains estimé de façon robuste, on réduit ce nombre de faux positifs. Si en plus, on exige que ces données atypiques pour être des lésions doivent vérifier les règles de la Tab.1, la segmentation est encore améliorée.

Même si l'amélioration due à STREM3 vs. STREM2 n'est pas flagrante sur les données simulées, elle s'avéra utile sur

les données cliniques.

Remarques. Il est important de noter que pour tous les exemples proposés ci-dessus, la même initialisation pour l'hyper-paramètre θ a été utilisée. De plus pour la méthode robuste, le paramètre h utilisé garde la même valeur, soit 96% des données. Lorsque la distance de Mahalanobis a été utilisée pour extraire les données atypiques, la *p_value* est toujours égale à 0.05. Ainsi les résultats peuvent être comparés de façon rigoureuse.

5.2 Résultats qualitatifs sur des données cliniques

Nous présentons un exemple de segmentation robuste sur un ensemble de données cliniques multidimensionnelles. Cette séquence longitudinale est acquise sur un scanner GE 1.5T et correspond au suivi d'un patient tous les trois mois durant un an. Chaque acquisition est composée de trois modalités : DP (TR = 2740 ms, TE = 22.9 ms), pondération en T2 (TR = 2740 ms, TE = 91.7 ms), et pondération en T1 après injection de contraste (TR = 460 ms, TE = 20.0 ms). L'épaisseur de coupe est 3 mm.

L'estimation des différents tissus est calculée pour les trois modalités et les cinq temps conjointement. Les lésions sont segmentées en utilisant STREM2. Seules les segmentations pour les deux premiers temps d'acquisition sont présentées Fig. 5 car pour les instants suivants il n'y a pas d'apparition ou de disparition de lésions, les images de segmentation ne présentent donc que peu d'intérêt.

L'hyper-paramètre θ est estimé de façon robuste avec $h = 95\%$ des données. Ensuite les données atypiques sont extraites en utilisant la valeur critique du χ^2 pour la distance de Mahalanobis, *p_value* = 0.01.

Notre méthode permet de segmenter de très petites lésions SEP même évolutives (cf Fig. 5) sur des images cliniques de très faible résolution (cf Fig. 6). On peut cependant remarquer la présence de nombreux faux positifs. Nous rappelons que pour cette expérience l'information *a priori* sur les lésions SEP n'a pas encore été introduite. D'autre part ces faux positifs sont localisés dans le LCR ou sur les bords du cerveau, ce qui anatomiquement prouve que ce ne sont pas des lésions. Ce type de faux positifs ne perturbera donc pas l'analyse du clinicien mais dans le but de quantifier numériquement les lésions il nous faudra ajouter la connaissance *a priori*.

Remarque. L'image considérée comme référence pour le recalage géométrique est la modalité T1 au premier temps d'acquisition.

L'initialisation du schéma itératif est aléatoire. Plusieurs initialisations sont donc faites (de l'ordre d'une dizaine), la convergence estimée globale est celle de plus grande vraisemblance. On peut préciser qu'en utilisant STREM, il a été constaté que le schéma itératif converge fréquemment vers le même minimum alors qu'avec un EM classique, différents minima sont atteints pour différentes initialisations. D'autre part, il est important de noter que ces images ont été acquises dans le cadre d'un protocole d'étude bien

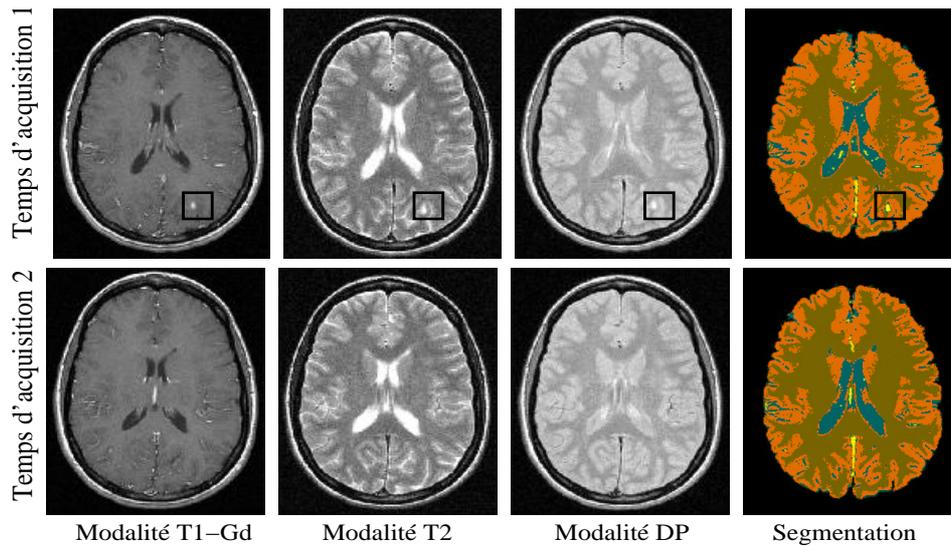


FIG. 5 – *STREM2* pour des données cliniques. Une lésion jeune apparaît au temps d'acquisition 1 et disparaît au temps d'acquisition 2. Les IRM sont stables pour les temps d'acquisition suivants. Les résultats préliminaires sur la segmentation jointe montre la détection de très petites lésions évolutives.

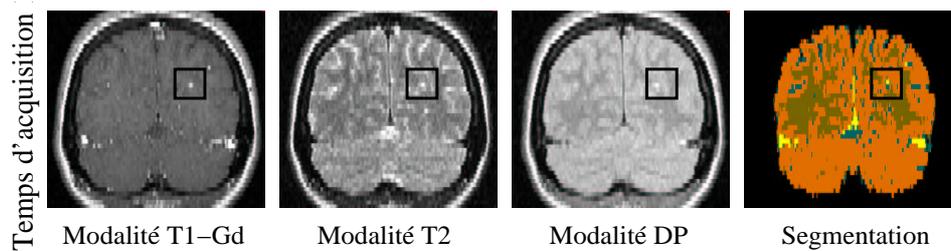


FIG. 6 – *STREM2* pour des données cliniques, coupes coronales pour le temps d'acquisition 1. Les images sont basse résolution et nous obtenons néanmoins une bonne segmentation.

défini. Les coupes de 3mm d'épaisseur sont considérées comme suffisantes, les lésions de taille inférieures à 3mm peuvent ne pas être visibles mais le clinicien estime que l'étude de lésions aussi petites n'est pas nécessaire. Par ailleurs, des coupes 1mm nécessitent un temps d'acquisition plus long qui serait pénible pour le patient. Notre algorithme doit donc pallier à ce problème de faible résolution en étant robuste aux effets de volume partiel (lorsque plusieurs tissus sont mélangés dans un même voxel).

6 Conclusions et discussion

Nous proposons un algorithme robuste multidimensionnel de segmentation sur des IRM multimodales $3D + t$. La convergence d'un tel schéma itératif a été démontré. Nous avons mise en évidence l'utilité d'une telle méthode robuste par rapport à un algorithme EM classique. Nous avons montré l'utilité de l'analyse multimodale puis longitudinale. Enfin, dans le cadre de notre application à la SEP, nous avons apporté des connaissances *a priori* utiles à la segmentation des lésions.

L'initialisation de l'hyper-paramètre θ est faite de façon aléatoire, dans un futur proche nous souhaitons l'initialiser en utilisant un atlas cérébral recalé vers nos données de façon non rigide. Les connaissances *a priori* sur l'apparition des lésions de SEP en IRM vont être introduites dans la segmentation des données cliniques. Notre algorithme va être validé sur une plus grande base de données cliniques en comparant la classification des lésions obtenue à celle faite par un expert.

Nous affinerons encore notre classification selon le type des lésions mais aussi leur évolution [18], de façon à obtenir des marqueurs IRM plus spécifiques et plus corrélés avec les signes cliniques que ceux déjà existant. D'autre part, les manifestations de la SEP en IRM entraînent en plus des lésions sclérotiques une atrophie cérébrale. Les tissus sains étant bien segmentés, il nous sera alors facile de calculer cette atrophie volumique et de suivre son évolution dans le temps.

Les applications de notre méthode ne sont pas restreintes à la SEP. En effet, une telle méthode de segmentation robuste

avec rejet de données atypiques peut s'appliquer dans le cadre d'autres anomalies cérébrales telles les tumeurs, ou les œdèmes présents chez les traumatisés crâniens graves par exemple. De plus, l'analyse longitudinale permet d'extraire des données atypiques dans une séquence temporelle et donc de pouvoir appliquer la méthode à des problèmes de détection d'événements.

Remerciements

Nous tenons à remercier le Dr. Carsin (CHRU Ponchaillou) pour nous avoir fourni les images.

Références

- [1] F. Barkhof and M. van Walderveen. Characterization of tissue damage in multiple sclerosis by nuclear magnetic resonance. *Philosophical Transactions: Biological Sciences*, 354(1390):1675–1686, October 1999.
- [2] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1):1–38, 1977.
- [3] P. Schroeter, J.M. Vesin, T. Langenberger, and R. Meuli. Robust Parameter Estimation of Intensity Distributions for Brain Magnetic Resonance Images. *IEEE Transactions on Medical Imaging*, 17(2):172–186, 1998.
- [4] K. Van Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Suetens. Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE Transactions on Medical Imaging*, 20(8):677–688, 2001.
- [5] G. Dugas-Phocion, M.A. González, C. Lebrun, S. Chanalet, C. Bensa, G. Malandain, and N. Ayache. Hierarchical segmentation of multiple sclerosis lesions in multi-sequence MRI. In *Isbi04*, Arlington, USA, April 2004.
- [6] D. Welti, G. Gerig, E.-W. Radü, L. Kappos, and G. Székely. Spatio-temporal segmentation of active multiple sclerosis lesions in serial MRI data. In M.F. Insana and R.M. Leahy, editors, *17th International Conference on Information Processing in Medical Imaging, IPMI'2001*, volume 2082 of *Lecture Notes in Computer Science*, pages 438–445, Davis, USA, June 2001. Springer.
- [7] M. Bosc, F. Heitz, J.P. Armspach, I. Namer, D. Gounot, and L. Rumbach. Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution. *NeuroImage*, 20(2):643–656, October 2003.
- [8] S. Prima, D.L. Arnold, and D.L. Collins. Multivariate statistics for detection of MS activity in serial multimodal MR images. In R.E. Ellis and T.M. Peters, editors, *6th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI'2003*, volume 2878 of *Lecture Notes in Computer Science*, pages 663–670, Montreal, Canada, November 2003. Springer.
- [9] D. Rey, G. Subsol, H. Delingette, and N. Ayache. Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis. *Medical Image Analysis*, 6(2):163–179, June 2002.
- [10] A. Shahar and H. Greenspan. A probabilistic framework for the detection and tracking in time of multiple sclerosis lesions. In *International Symposium on Biomedical Imaging, ISBI'04*, pages 440–443, Arlington, USA, April 2004.
- [11] C. Baillard, P. Hellier, and C. Barillot. Segmentation of brain 3D MR images using level sets and dense registration. *Medical Image Analysis*, 5(3):185–194, 2001.
- [12] S. Prima, N. Ayache, T. Barrick, and N. Roberts. Maximum likelihood estimation of the bias field in MR brain images: investigating different modelings of the imaging process. In W.J. Niessen and M.A. Viergever, editors, *4th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI'2001*, volume 2208 of *Lecture Notes in Computer Science*, pages 811–819, Utrecht, The Netherlands, October 2001. Springer.
- [13] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, April 1997.
- [14] J. Sijbers, A.J. Den Dekker, P. Scheunders, and D. Van Dyck. Maximum Likelihood estimation of Rician distribution parameters. *IEEE Transactions on Medical Imaging*, 17(3):357–361, 1998.
- [15] N. Neykov, P. Fizmoser, R. Dimova, and P. Neytchev. Mixture of GLMs and the trimmed likelihood methodology. In *Compstat'04*, pages 1585–1592, 2004.
- [16] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, and A.C. Evans. BrainWeb: Online Interface to a 3D MRI Simulated Brain Database. *Proceedings of 3-rd International Conference on Functional Mapping of the Human Brain, Copenhagen, May 1997*, 5(4), 1997.
- [17] A.P. Zijdenbos, B.M. Dawant, R.A. Margolin, and A.C. Palmer. Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE Transactions on Medical Imaging*, 13(4):716–724, December 1994.
- [18] C. Lucchinetti, W. Bruck, J. Parisi, B. Scheithauer, M. Rodriguez, and H. Lassmann. Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. *Annals of Neurology*, 47(6):707–717, June 2000.