



Unifying framework for multimodal brain MRI segmentation based on Hidden Markov Chains

S. Bricq^{a,b,*}, Ch. Collet^a, J.P. Armspach^b

^a Université Strasbourg I, LSIT: UMR CNRS 7005, ENSPS/LSIT, Pole API, Bd S. Brant, BP 10413 F-67412 Illkirch, France

^b Université Strasbourg I, LINC, UMR CNRS 7191, France

ARTICLE INFO

Article history:

Received 22 June 2007

Received in revised form 7 March 2008

Accepted 7 March 2008

Available online 17 March 2008

Keywords:

Multimodal MRI segmentation

Hidden Markov Chain

Partial volume effect

Brain atlas

ABSTRACT

In the frame of 3D medical imaging, accurate segmentation of multimodal brain MR images is of interest for many brain disorders. However, due to several factors such as noise, imaging artifacts, intrinsic tissue variation and partial volume effects, tissue classification remains a challenging task. In this paper, we present a unifying framework for unsupervised segmentation of multimodal brain MR images including partial volume effect, bias field correction, and information given by a probabilistic atlas. Here-proposed method takes into account neighborhood information using a Hidden Markov Chain (HMC) model. Due to the limited resolution of imaging devices, voxels may be composed of a mixture of different tissue types, this partial volume effect is included to achieve an accurate segmentation of brain tissues. Instead of assigning each voxel to a single tissue class (i.e., hard classification), we compute the relative amount of each pure tissue class in each voxel (mixture estimation). Further, a bias field estimation step is added to the proposed algorithm to correct intensity inhomogeneities. Furthermore, atlas priors were incorporated using probabilistic brain atlas containing prior expectations about the spatial localization of different tissue classes. This atlas is considered as a complementary sensor and the proposed method is extended to multimodal brain MRI without any user-tunable parameter (unsupervised algorithm). To validate this new unifying framework, we present experimental results on both synthetic and real brain images, for which the ground truth is available. Comparison with other often used techniques demonstrates the accuracy and the robustness of this new Markovian segmentation scheme.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Segmentation is an important step for quantitative analysis of brain images and for the study of many brain disorders. Indeed structural changes in the brain can be due to some brain disorders. The quantification of these changes, by measuring volumes of structures of interest, can be used to characterize disease severity or evolution. For instance, in morphometry, brain tissue segmentation enables to compare tissue volumes and to follow the evolution of some brain disorders. Before these measurements can be carried out, labelling process must be performed. In this approach, the tissue types of interest are white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF), but other approaches classify voxels according to their anatomical structure. As most studies involve large amounts of data, manual tracing of cerebral structures in MR images by a human expert is obviously a time consuming process. Moreover, these manual segmentations are prone to large intra- and inter-observer variability, which deteriorates the

significance of the resulting segmentation analysis. Thus, due to different artifacts appearing on MR images, segmentation is not obvious and remains a challenging task. The first point deals with spatial regularization required to overcome the disturbance added during the MRI formation. A second artifact that appears in MR images is the corruption by a bias field. These smooth intensity variations are due to equipment imperfections and patient-induced electrodynamic interactions (Sled and Pike, 1998). A third important issue in MR segmentation is the partial volume effect (PVE). Due to the limited resolution of the acquisition system, voxels along the boundaries are composed of two or more tissues. Therefore, it is necessary to consider this partial volume effect, in order to achieve an accurate segmentation of brain tissues. In fact, hard segmentation (WM, GM, CSF) ignores this problem and therefore loses information concerning the tissue structure.

This paper proposes a unifying framework for determining the amount of each substance in each voxel of the brain, including bias field correction, neighborhood information using the Hidden Markov Chain model (HMC) and using a probabilistic atlas as a complementary sensor. To carry out a regularized segmentation, we adopt the Hidden Markov Chain model. A bias field estimation step is added to correct intensity inhomogeneities. To achieve an accurate

* Corresponding author. Address: Université Strasbourg I, LSIT: UMR CNRS 7005, ENSPS/LSIT, Pole API, Bd S. Brant, BP 10413 F-67412 Illkirch, France.

E-mail address: bricq@lsit.u-strasbg.fr (S. Bricq).

segmentation of brain tissues, partial volume effects are included in the model. Instead of assigning each voxel to a single tissue class (i.e., hard classification), the method we proposed estimates the proportion of each tissue type in each voxel without requiring any tunable parameter. In fact, our contribution extends the model presented in (Bricq et al., 2006) by including partial volume effect, bias field correction and prior information given by a probabilistic atlas. Indeed, the atlas represents the *a priori* probability of a voxel being WM, GM, or CSF, after an image has been normalized to the same space. The method we proposed is then extended to multi-modal brain MRI. In this work, we assume that the registration of the images with the atlas has already been performed. Further information on registration methods can be found elsewhere (e.g., in our case, we use the method proposed in Noblet et al. (2005)).

This paper is organized as follows: in the following section, we present the HMC model. Section 3 describes our method that includes partial volume effect, bias field correction and atlas information in the HMC model to segment brain MR images. The method is then validated on synthetic data in Section 4. In Section 5, results obtained on real data are shown and compared with other often-used methods: namely, SPM5¹ (Statistical Parametric Mapping) (Ashburner and Friston, 2005), EMS² (Expectation-Maximization Segmentation) (Van Leemput et al., 2003) and the method developed by Shattuck et al. (2001). Finally, in Section 6, conclusions are drawn and future developments are suggested.

2. Hidden Markov Chain

As the state of instrumentation evolves to produce ever finer resolution in spectral, spatial, and temporal data, ever more sophisticated statistical techniques are required to properly handle these data. In this way, Markovian models provide powerful statistical tools for computer vision applications. Indeed, the major idea consists in taking into account the neighborhood to infer the decision. To phrase it in a different way, the contextual information contained in the labelling of all the other pixels is entirely captured by the neighborhood of the pixel under concern. Then, great care must be taken while choosing the prior model. Different models are available: Markovian Random Field (Graffigne et al., 1995), Markovian Quadtree (Romberg et al., 2001, 2003), Hidden Markov Chain (Pieczyński, 2003), Hybrid Models (Pérez et al., 2000)... thus model selection is a challenging computational problem. The other major difficulty is computational: most of the Markov models are non-causal. As a consequence, inference must be conducted iteratively, which might turn prohibitively expensive when dealing with large data images. For example, Markov Random Fields (MRF) are time consuming, specially in dealing with 3D cubes. Indeed, *a posteriori* realizations of label maps X cannot be created directly, and thus are approximated iteratively using the Metropolis algorithm (Metropolis et al., 1953) or the Gibbs sampler (Geman and Geman, 1984). To overcome this issue, parameter estimation is frequently based on a limited subset of voxels (Ruan et al., 2002, 1999b) on one hand, and to avoid such case, hierarchical or alternative approaches can efficiently provide substantial gains in speed on the other hand. Compared to MRF, the Hidden Markov Chain model requires lower computing costs and does not need any approximation: the *a posteriori* law can be exactly estimated. In our case, we propose to use Hidden Markov Chain model, by using a 3D Hilbert-Peano scan of the data cube (Section 2.1). This approach is based on the assumption that we have introduced *a priori* information, namely that the data are correlated. By this

way, different models can be used: they only model correlation (i.e., homogeneous label map) in different manner. Moreover, because the data are strongly correlated, the 3D scan does not break the neighborhood information, even if two neighbors in the cube can be far away in the chain. In fact, one major advantage to use Markov Chain Model stems from their ability to learn model parameters through re-estimation procedure, and to provide a form of context handling in pattern recognition. Iterative update algorithm re-estimates parameters of a given Hidden Markov model to produce a new model which has a higher probability of generating the given observation sequence. This re-estimation procedure is continued until no more significant improvement in probability can be obtained, i.e., the local minimum is thus found. Our motivations for using such HMC (Hidden Markov Chain) model are to exploit the spatial dependencies between neighboring pixels and to impose a spatial regularity constraint on the classes to provide fast computations and efficient structures to process large data cubes.

HMC has been used to segment 2D images, especially SAR images. Fjortoft et al. (2003) compare the classification obtained on both synthetic and real radar images using HMC and MRF. HMC provides a fast and robust alternative to MRF. The classification obtained with HMC is sometimes less regular than the classification obtained with MRF but fine structures are generally very well preserved.

2.1. Hilbert-Peano scan

Thanks to the fractal Hilbert-Peano scan (Fig. 1), Hidden Markov Chain models can be used to segment 3D data (Bandoh and Kamata, 1999). The interest of Markov chain methods for image segmentation compared to 3D Markov field models is that, being based on 1D modeling, they result in lower computing cost. Nevertheless, contrary to the Markov Random Field, in the Markov chain model the neighboring information is partially translated in the chain (through the 3D-Hilbert-Peano path): two neighbors in the chain are neighbors in the grid, but two neighbors in the cube can be far away in the chain. In fact, the Hilbert-Peano path preserves the neighborhood in the 1D vector as well as possible. However, due to probably strong correlation within the data cube, this scan will weakly influence the segmentation results. The first step of segmentation algorithms based on Markov chains consists in transforming the image into a vector. Once all the processing has been carried out on the vector, we apply the inverse transformation on the segmented chain to obtain the final segmented image. The influence of different Hilbert-Peano scans on the segmentation results is studied in Section 4.2.2.

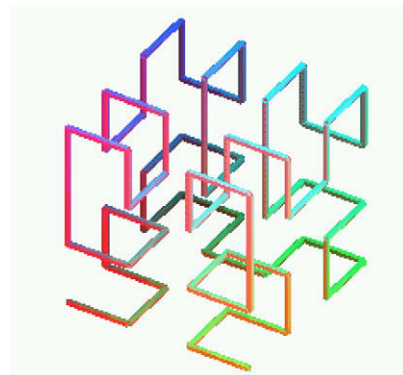


Fig. 1. Hilbert-Peano scan for 3D images.

¹ <http://www.fil.ion.ucl.ac.uk/spm>.

² <http://www.medicalimagecomputing.com/EMS/>.

2.2. HMC theory

Let us now consider two sequences of random variables $X = (X_n)_{n \in S}$ the hidden process, and $Y = (Y_n)_{n \in S}$ the observed one, with S the finite set corresponding to the N voxels of an image. Y_n is a vector containing the observations from different modalities at voxel n : $Y_n = \{Y_n^1, \dots, Y_n^L, \dots, Y_n^L\}$ with L the number of modalities and Y_n^l the observations at modality l at voxel n . For multimodal images, we use multivariate normals with mean μ_k and covariance matrix Σ_k . Each X_n takes its value in a finite set of K classes $\Omega = \{\omega_1, \omega_2, \dots, \omega_K\}$. Each Y_n^l takes its value in the set of real numbers \mathbb{R} .

X is a Markov chain if:

$$P(X_{n+1} = \omega_{i_{n+1}} | X_n = \omega_{i_n}, \dots, X_1 = \omega_{i_1}) = P(X_{n+1} = \omega_{i_{n+1}} | X_n = \omega_{i_n}) \quad (1)$$

Thus, X will be determined by two parameters:

- the initial distribution $\pi_i = P(X_1 = \omega_i)$;
- the transition matrix $t_{ij}^n = P(X_{n+1} = \omega_j | X_n = \omega_i)$.

Furthermore, we assume the homogeneity of the Markov chain, which means that the transition matrix is independent of the location n : $t_{ij}^n = t_{ij}$, for $1 \leq n < N$. Nevertheless, we do not assume the stationarity of the chain, which means that $\pi_i \neq P(X_n = \omega_i) \forall n \neq 1$. Fig. 2 presents the HMC dependency graph.

2.3. HMC algorithm

Let X be an homogeneous Markov chain, then we assume:

- Y_n are conditionally independent given X

$$P(Y|X) = \prod_{n=1}^N P(Y_n = y_n | X) \quad (2)$$

- the conditional probability of Y_n given X is equal to the conditional probability of Y_n given X_n

$$P(Y_n = y_n | X) = p(Y_n = y_n | X_n = \omega_k) \quad (3)$$

Eqs. (2) and (3) lead to

$$P(Y|X) = \prod_{n=1}^N p(Y_n = y_n | X_n = \omega_k) \quad (4)$$

The *a priori* parameters are clustered in $\Phi_x = \{\pi_k, t_{kl}\}$, and the *data-driven* ones are gathered in $\Phi_y = \{\mu_k, \Sigma_k\}$. We will denote $\Phi = \{\Phi_x, \Phi_y\}$ and $f_k(y_n)$ the likelihood $p(Y_n = y_n | X_n = \omega_k)$ of the observations y_n conditionally to $X_n = \omega_k$. The aim consists in finding hidden labels given the observations. X can be recovered from the observed process using different Bayesian classification techniques like Marginal Posterior Mode (MPM) (Gelman et al., 2005) with Baum and Welch forward–backward algorithm (Devijsver, 1985), or Maximum *a posteriori* (MAP) with the Viterbi algorithm (Forney, 1973). Φ_x and Φ_y are unknown parameters. We have based on Baum

and Welch algorithm (Baum, 1972) to estimate these parameters and the Markov chain. We thus need to define forward and backward probabilities (Baum, 1972). These probabilities take an important part in parameter estimation and restoration of Y , they allow to estimate the *a posteriori* probability of each X_n :

- Forward probability

$$\alpha_n(k) = P(X_n = \omega_k, Y_{\leq n} = (y_1, \dots, y_n)) \quad (5)$$

with $Y_{\leq n} = (Y_1, \dots, Y_n)$.

- Backward probability

$$\beta_n(k) = P(Y_{>n} = (y_{n+1}, \dots, y_N) | Y_n = y_n) \quad (6)$$

with $Y_{>n} = (Y_{n+1}, \dots, Y_N)$.

However, with these expressions, we are confronted with underflow problems. To overcome this issue, Devijver (1985) proposes to use conditional probabilities instead of joint probabilities:

- Forward probability

$$\alpha_n(k) = P(X_n = \omega_k | Y_{\leq n} = (y_1, \dots, y_n)) \quad (7)$$

- Backward probability

$$\beta_n(k) = \frac{P(Y_{>n} = (y_{n+1}, \dots, y_N) | X_n = \omega_k)}{P(Y_{>n} = (y_{n+1}, \dots, y_N) | Y_{\leq n} = (y_1, \dots, y_n))} \quad (8)$$

These probabilities can be recursively computed. Forward probabilities are obtained with a recursive computation scanning of the chain from $n = 1$ to $n = N$. Backward probabilities are obtained using an inverse scan of the chain.

Finally, to segment the chain, we use the MPM criterium:

$$\hat{x}_n = \arg \max_{\omega_k} P(X_n = \omega_k | Y = y, \Phi) = \arg \max_{\omega_k} \alpha_n(k) \beta_n(k) \quad (9)$$

2.4. Parameter estimation

Parameter estimation on Markov chains can be stochastic (SEM, Masson and Pieczynski, 1993), deterministic (EM algorithm, Tanner, 1993) or hybrid (ICE, Pieczynski, 1992). In this article, the model parameters have been estimated with the EM algorithm. EM is an iterative optimization method for parameter estimation in the sense of maximum likelihood. Instead of using only observed data Y , observations are augmented with hidden variables X to simplify calculation. Observed data are considered as incomplete. Missing data X are added to these data to lead to complete data (X, Y) . Dempster et al. (1977) propose this two-steps algorithm:

- E-step (Expectation)

$$Q(\Phi, \Phi^{[q]}) = E[\log P(X, Y | \Phi) | Y, \Phi^{[q]}] \quad (10)$$

- M-step (Maximization)

$$\Phi^{[q+1]} = \arg \max_{\Phi} Q(\Phi, \Phi^{[q]}) \quad (11)$$

In the case of Hidden Markov Chains, we have the following relationship:

$$P(X, Y | \Phi) = P(X_1 | \Phi_x) \prod_{n=2}^N P(X_n | X_{n-1}, \Phi_x) \prod_{n=1}^N P(Y_n | X_n, \Phi_y) \quad (12)$$

Thus, we can write

$$\begin{aligned} Q(\Phi, \Phi^{[q]}) &= E[\log P(X, Y | \Phi) | Y, \Phi^{[q]}] \\ &= \sum_{n=1}^N E[\log P(Y_n | X_n, \Phi_y) | Y, \Phi^{[q]}] \\ &\quad + \sum_{n=2}^N E[\log P(X_n | X_{n-1}, \Phi_x) | Y, \Phi^{[q]}] \end{aligned} \quad (13)$$

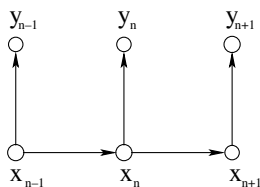


Fig. 2. Dependency graph of Hidden Markov Chain.

$$+ E[\log P(X_1 | \Phi_x) | \mathbf{Y}, \Phi^{[q]}] \quad (14)$$

$$= \sum_i \gamma_i^{[q]}(i) \log \pi_i + \sum_{n \neq 1} \sum_i \sum_j \xi_n^{[q]}(i, j) \log t_{ij} \\ + \sum_n \sum_i \gamma_n^{[q]}(i) \log f_i(y_n) \quad (15)$$

with

- *a posteriori* marginal probability

$$\gamma_n^{[q]}(i) = p(x_n = \omega_i | y, \Phi^{[q]}) \quad (16) \\ = \frac{\alpha_n(i) \beta_n(i)}{\sum_j \alpha_n(j) \beta_n(j)}$$

- *a posteriori* joint probability

$$\xi_n^{[q]}(i, j) = p(x_{n-1} = \omega_j, x_n = \omega_i | y, \Phi^{[q]}) = \alpha_{n-1}(j) t_{ji} f_i(y_n) \beta_n(i) \quad (17)$$

- initial distribution

$$\pi_i = p(x_1 = \omega_i) \quad (18)$$

- transition matrix

$$t_{ij}^n = p(x_{n+1} = \omega_j | x_n = \omega_i) \quad (19)$$

This EM algorithm was developed by Baum (1972). Calculation of forward and backward probabilities can be found in previous section. We maximize Q under the following constraints: $\sum_i \pi_i = 1$ and $\sum_j t_{ij} = 1$. Using Lagrangian multipliers, we obtain

$$\pi_i^{[q+1]} = \gamma_1^{[q]}(i) \quad (20)$$

$$t_{ij}^{[q+1]} = \frac{\sum_{n=2}^N \xi_n^{[q]}(i, j)}{\sum_{n=1}^{N-1} \gamma_n^{[q]}(i)} \quad (21)$$

$$\sum_n \sum_i \gamma_n^{[q]}(i) \log f_i(y_n) = \sum_n \sum_i \gamma_n^{[q]}(i) \left[-\log \sigma_i - \frac{1}{2\sigma_i^2} (y_n - \mu_i)^2 \right] \quad (22)$$

Setting the partial derivative for μ_i and σ_i of $\sum_n \sum_i \gamma_n^{[q]}(i) \log f_i(y_n)$ to zero leads to

$$\mu_i^{[q+1]} = \frac{\sum_n \gamma_n^{[q]}(i) y_n}{\sum_n \gamma_n^{[q]}(i)} \quad (23)$$

$$\sigma_i^{2[q+1]} = \frac{\sum_n \gamma_n^{[q]}(i) [y_n - \mu_i]^2}{\sum_n \gamma_n^{[q]}(i)} \quad (24)$$

Algorithm 1 presents the MPM segmentation based on a Hidden Markov Chain.

Algorithm 1 (MPM segmentation based on a Hidden Markov Chain).

- (1) Image Vectorization using 3D Hilbert-Peano scan

- (2) Initialization

- (a) For *a priori* parameters Φ_x

$$\pi_k = \frac{1}{K} \quad \forall k \quad (25)$$

$$t_{ij} = \begin{cases} \frac{3}{4} & \text{if } i = j \\ \frac{1}{4(K-1)} & \text{else} \end{cases} \quad (26)$$

- (b) Data-driven parameters Φ_y are initialized using K-Means

- (3) For each iteration

- (a) Calculation of $\alpha_n(k)$ and $\beta_n(k)$ using Eqs. (7) and (8).
- (b) Update of parameters Φ using Eqs. (20), (21), (23) and (24).

- (4) MPM segmentation using Eq. (9)

$$\hat{x}_n = \arg \max_{\omega_k} P(X_n = \omega_k | Y = y, \Phi) = \arg \max_{\omega_k} \alpha_n(k) \beta_n(k)$$

- (5) Inverse transformation to obtain the segmented image

3. Unifying framework for multimodal MRI segmentation

In this part, we describe the unifying framework for unsupervised segmentation of multimodal brain MR images including partial volume effect, bias field correction, information given by a probabilistic atlas, and neighborhood information based on the HMC principles which we presented in the previous section.

3.1. Probabilistic atlas

In Bricq et al. (2006), we used the Triplet Markov Chain (TMC) (Lanchantin and Pieczynski, 2004, 2002) to include *a priori* information given by a probabilistic atlas (Fig. 3). Here-proposed model is more general since it includes bias field correction, atlas information, and partial volume modeling (Section 3.3). For a Bayesian approach, the information brought by a probabilistic atlas may help and finally improve the accuracy of segmentation. For this reason, *a priori* information was sometimes introduced in some models (Ashburner and Friston, 2005, 2002). Ashburner and Friston proposed a modified maximum likelihood mixture model algorithm using probabilistic images for the segmentation (Ashburner and Friston, 1997, 2000). The probability images they used are the means of binary images of white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF) obtained from the segmentation of the original images. However, this method does not take into account neighborhood information, and therefore the quality of the segmentation is degraded if noise is present in the images. Other methods (Van Leemput et al., 1999b, 2002) propose to use a digital brain atlas to initialize the iterative process. Each time, the main idea remains the same: the atlas is an interesting tool to drive the segmentation process. In Marroquin et al. (2002), a brain atlas is used to segment the brain from non-brain tissue and to compute prior probabilities for each class at each voxel location. In Al-Zubi et al. (2002), a fuzzy atlas which indicates the probability distribution of each tissue type in the brain is used as *a priori* knowledge to correct misclassified voxels. In this last case, the neighborhood behavior is regularized thanks to a Markov Random Field modeling to gain robustness towards the noise (Fig. 4).

The modeling scheme we propose captures *a priori* information both in the neighborhood and in an atlas and in fact, this atlas is considered as a complementary sensor. In this study, we use an atlas bringing *a priori* information deriving from averaged segmented MR images. In this way, the atlas provides probability maps of the three main tissues in the brain (WM, GM, CSF). $p(B_n | X_n = \omega_k) = b_n(k)$ denotes the tissue probability for class k at voxel n given by the atlas.

The model we propose here is a Hidden Markov Chain with two independent sensors. In this case, we have as information, both the intensity of the observed data and the probability maps. Thus,

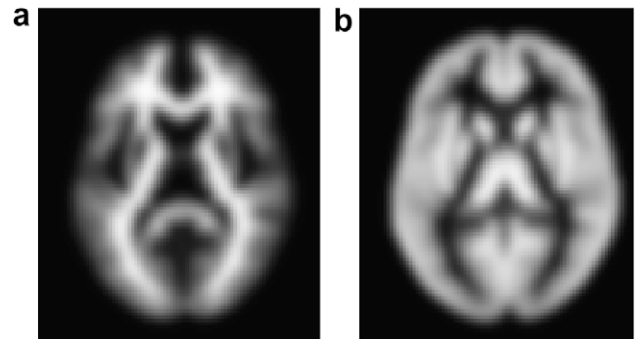


Fig. 3. Example of *a priori* probability images: (a) white matter atlas and (b) gray matter atlas.

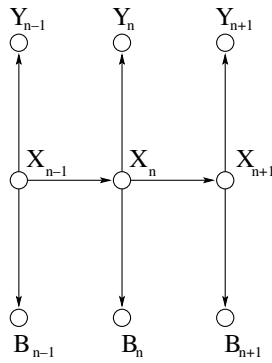


Fig. 4. Hidden Markov Chain X with information given by a probabilistic atlas and the observations Y .

we have to find $\hat{x}_n = \arg \max_{\omega_k} P(X_n = \omega_k | \mathbf{Y}, \mathbf{B})$, where $P(X_n = \omega_k | \mathbf{Y}, \mathbf{B}) = \alpha_n(k) \beta_n(k)$ with $\alpha_n = P(X_n = \omega_k | \mathbf{Y}_{\leq n}, \mathbf{B}_{\leq n})$ forward probability and $\beta_n = \frac{P(\mathbf{Y}_{>n}, \mathbf{B}_{>n} | X_n = \omega_k)}{P(\mathbf{Y}_{>n}, \mathbf{B}_{>n} | \mathbf{Y}_{\leq n}, \mathbf{B}_{\leq n})}$ backward probability. Equations of the forward and backward probabilities are detailed in Appendix A.

With this link to the atlas, *a priori* knowledge from the probabilistic maps of each tissue classes and spatial organization will be both considered for segmentation. The data-driven parameters will depend on the observations and on the probability maps of the different tissue of the brain. The use of an atlas makes the segmentation process more robust by introducing an *a priori* brought by averaged cross-patients features. Ignoring this information could sometimes lead the algorithm to a low accuracy segmentation map.

3.2. Bias field correction

An artifact that appears in MR images is the corruption by a bias field. These smooth intensity variations are due to equipment imperfections and patient-induced electrodynamic interactions (Sled and Pike, 1998). The most commonly used model is a multiplicative bias field with an additive noise:

$$y(x) = \tilde{y}(x)b(x) + n(x) \quad (27)$$

where $y(x)$ is the voxel intensity, $\tilde{y}(x)$ the “true” voxel intensity, $b(x)$ is a multiplicative bias field, and $n(x)$ an additive Gaussian noise. Belaroussi et al. (2006) present an overview of existing methods for intensity non-uniformity correction. This bias field is frequently modeled as a linear combination of smooth basis functions. These functions are either spline functions (Sled and Zijdenbos, 1998) or polynomial basis functions (Van Leemput et al., 1999a, 2000). Sled and Zijdenbos (1998) propose to correct non-uniformity by finding the smooth multiplicative field which maximizes the frequency content of \mathbf{Z} , where \mathbf{Z} denotes the probability function of $\log \tilde{y}$. This method does not need a tissue model and can be applied without data segmentation. Other approaches interleave bias field estimation and tissue classification: bias field correction and segmentation are then included in the same framework by maximizing the same cost function. Most of these approaches are based on the EM algorithm (Van Leemput et al., 1999a, 2002), estimating alternatively classification and bias field. In this way, different modelings for the bias field b are proposed. Wells et al. (1996) model the bias field as a N -dimensional zero mean Gaussian prior probability density function, where Y_n is the observed MRI intensity at voxel n . Van Leemput et al. (1999a) represent the bias field by a linear combination of polynomial basis functions. Ashburner and Friston (2005) parameterize the bias field as an exponential of a linear combination of low frequency basis functions.

To take into consideration bias field inhomogeneities, a bias field estimation step is included to the processing chain we proposed. MRI intensity inhomogeneities are modeled by a bias field

which multiplies the intensity data. Thanks to a logarithmic transformation of the intensities, the artifact can be modeled by an additive bias field (Wells et al., 1996): $\mathbf{Z}_n = \log \mathbf{Y}_n$, where \mathbf{Y}_n is the observed MRI intensity at voxel n .

We model the bias field as a linear combination $\sum_k c_k \psi_k(t)$ of smooth basis functions $\psi_k(t)$, as Van Leemput et al. (1999a), whereas each class is modeled by a normal distribution

$$p(\mathbf{Z}_n = \mathbf{z}_n | X_n = \omega_i) = \frac{1}{\sqrt{\det \Sigma_i} (2\pi)^{L/2}} \exp \left(-\frac{1}{2} \left[\mathbf{z}_n - \boldsymbol{\mu}_i - \sum_k c_k \psi_k(t_n) \right]^T \Sigma_i^{-1} \left[\mathbf{z}_n - \boldsymbol{\mu}_i - \sum_k c_k \psi_k(t_n) \right] \right) \quad (28)$$

Such an approximation was used in Wells et al. (1996), Guillemaud and Brady (1997) and Van Leemput et al. (1999a). The algorithm presented in the previous section is slightly modified to include the bias field correction step. Polynomial basis functions were used as bias field basis functions. Since the likelihood is only increased and not maximized, the algorithm is a generalized EM (GEM). Estimation of the bias field parameters is detailed in Appendix B.

3.3. Hidden Markov Chain with partial volume effect

Instead of assigning each voxel to a single class, some authors estimate the amount of every tissue in each individual voxel. Indeed, due to the limited resolution of acquisition system, voxels along the boundaries between different classes are composed of two or more tissues. The classification of a voxel representing the main tissue type loses information about the tissue content.

In the literature, several papers have incorporated partial volume effect in segmentation: modified EM algorithms have been proposed to include the partial volume effect (Dugas-Phocion et al., 2004). However, their method does not consider pixel neighborhood. Bosc et al. (2003) address issues related to partial volumes by modeling the acquisition process and using a high resolution label image, thereby obtaining high resolution segmentations from standard MR images.

Some approaches model the mixels with a spatial correlation between the proportions of the different tissue types. A Markovian *a priori* is introduced to impose this constraint: Choi et al. (1991) use a Markov Random Field model (MRF) to take into account spatial correlation between neighboring voxels (this method needs multimodal images, which are not always available). Indeed the probability of occurrence of each tissue type for a voxel is influenced by the tissue type of its neighbors. Shattuck et al. (2001) and Tohka et al. (2004) first classify the brain into different PV labels using a simple Potts model. Then they assign to each voxel a tissue fraction for WM, GM, and CSF based on its tissue label and its intensity in the MR image. However, this method requires one parameter to be chosen by the user: the β term of the Potts model that controls the strength of the prior is a user-tunable parameter. Some methods only identify voxels containing PVE or do not try to estimate the amount of each tissue in each voxel: for example, Ruan et al. (2000) propose to first classify voxels in pure tissues and mixture classes using a MRF prior. The segmentation obtained is then used to assign each voxel to one single tissue type using a T1-specific feature, thus losing all notion of partial volume effect. In another paper, Ruan et al. (2002) use a fuzzy Markov Random Field model to take into account partial volume effect. This method is unfortunately time consuming due to the generation of the Markov field which is carried out using a Gibbs sampler. To solve this problem, the estimation algorithm was performed in a region which is about 1/6 of the entire volume, thereby losing information for computational reasons. Van Leemput et al. (2003) use a sub-voxel label image L . The observed image $\tilde{\mathbf{Y}} = \{\tilde{y}_i, i = 1, \dots, I\}$ is considered as the downsampling by a factor M of an ideal image

$Y = \{y_i, i = 1, \dots, J\}$ (with $J = I * M$) without partial volume effect. Given the observed image \tilde{Y} , the aim is to reconstruct the subvoxel label L or rather to estimate the tissue fractions in each voxel. They use a EM algorithm (Dempster et al., 1977) to estimate model parameters and missing data, i.e., unknown data L , Y from the observed one \tilde{Y} . They present three different prior models for the underlying label image L , especially a MRF model. This MRF model is defined on subvoxels, which constrain homogeneous regions of pure tissues bordered by PV voxels. However, this algorithm was implemented only for 2D images and requires about 20 min for one slice of an MR image.

The method we propose in this paper estimates the PV fraction of pure tissue types in each voxel by adapting the HMC model presented in the previous section to include partial volume effects. The proposed method is a post-processing step working on the non-log-transformed data. During this step the atlas is not used and the parameters of the bias field model are kept fixed. Let us denote by \mathbf{A}_n the vector which contains proportions of each tissue class in voxel n

$$\mathbf{A}_n = \begin{bmatrix} a_n^1 \\ \vdots \\ a_n^{K-1} \end{bmatrix} \quad (29)$$

with $\forall n, \forall k, a_n^k$ denoting the proportion of tissue k in voxel n and taking its values within $[0, 1]$. Amounts of tissue K are deduced by: $a_n^K = 1 - \sum_{k=1}^{K-1} a_n^k$. a_n^k can take an infinity of values in $[0, 1]$, but in practice, we discretize it using a step of 0.1 or 0.2, which means that a_n^k takes its values in $\{0, 0.1, 0.2, 0.3, \dots, 0.9, 1\}$ (respectively, $\{0, 0.2, \dots, 0.8, 1\}$).

We represent the intensity y of a voxel by a weighted sum of K pure tissues:

$$y = \sum_{i=1}^K a^i y_i^p \quad (30)$$

with $\sum_{i=1}^K a^i = 1$ and $a^i > 0$. a^i denotes the proportion of the pure class i and y_i^p a random variable representing the pure class i . We assume that each pure class is normally distributed $y_i^p \sim \mathcal{N}(\mu_i, \sigma_i^2)$ in intensity and that Y_i^p are independent. Therefore, for a given \mathbf{A} : $\mathbf{Y}|\mathbf{A} \sim \mathcal{N}(\sum_{i=1}^K a^i \mu_i, \sum_{i=1}^K (a^i)^2 \sigma_i^2)$ (Stark and Woods, 1986).

\mathbf{A} is a Markov chain, thus we can apply the properties of Markov chains presented in Section 2.3. \mathbf{A} is determined by the initial distribution and the transition matrix. Fig. 5 presents the dependency graph of a HMC with partial volume effect. We use the MPM (Gelman et al., 2005) to obtain a labelling of the image.

$$\hat{\mathbf{A}}_n = \arg \max_{\mathbf{A}_n} P(\mathbf{A}_n | \mathbf{Y}) = \alpha_n \beta_n \quad (31)$$

with $\alpha_n = P(\mathbf{A}_n | \mathbf{Y}_{\leq n})$ forward probability and $\beta_n = \frac{P(\mathbf{Y}_{\leq n} | \mathbf{A}_n)}{P(\mathbf{Y}_{> n} | \mathbf{Y}_{\leq n})}$ backward probability. For each voxel, we compute the probabilities

of every feasible label and we pick the one with the highest probability.

Forward and backward probabilities can be computed recursively:

- forward probability

$$\alpha_1 = P(\mathbf{A}_1)P(\mathbf{Y}_1 | \mathbf{A}_1) \quad (32)$$

$$\alpha_n = P(\mathbf{A}_n, \mathbf{Y}_{\leq n}) \quad \forall n > 1 \quad (33)$$

$$= \sum_{\mathbf{A}_{n-1}} \alpha_{n-1} P(\mathbf{A}_n | \mathbf{A}_{n-1}) P(\mathbf{Y}_n | \mathbf{A}_n) \quad (34)$$

Forward probabilities are then normalized.

- backward probability

$$\beta_N = 1 \quad (35)$$

$$\beta_n = \frac{1}{N_{n+1}} \sum_{\mathbf{A}_{n+1}} \beta_{n+1} P(\mathbf{A}_{n+1} | \mathbf{A}_n) P(\mathbf{Y}_{n+1} | \mathbf{A}_{n+1}) \quad \forall n < N \quad (36)$$

Thus, we obtain K segmentation maps containing the amount of each tissue type in each voxel. As a conclusion, we proposed a new HMC model to take into account PVE, which is entirely unsupervised. The results we obtain (cf. following sections) show that contextual information can be correctly captured by the 3D Hilbert-Peano scan. The major advantage to use HMC stands for their ability to learn model parameters and to estimate *a posteriori* probabilities on each site in exact manner with low computing costs.

3.4. Algorithm outline

We will assume that the registration of the images with the atlas (Noblet et al., 2005) has already been performed. Therefore, in the following we will not consider the registration step in computing time. Spatial normalization is performed using a non-rigid transformation described in (Noblet et al., 2005). Then the brain volume is extracted by removing the skull using the Brain Extraction Tool (BET) (Smith, 2002). To initialize the *data-driven* parameters $\Phi_y = \{\mu_k, \sigma_k\}$, we use the K-Means algorithm (Bovik, 2000). *A priori* parameters $\Phi_x = \{\pi_k, t_{kl}\}$ were initialized as described in Algorithm 1. Partial volume effects do not affect bias field correction. Thus, to speed up the algorithm, we first estimate the bias field before adding partial volume effects (see Fig. 6). We apply the HMC algorithm with bias field correction using *a priori* information contained in the probability maps. This leads to the estimation of the bias field and of μ and σ for pure classes. At this step of the algorithm, we could also provide a segmentation into three pure classes.

To reduce the numerical complexity, we only consider two different tissue types in a voxel at the same time. In fact, the probability of mixing more than two tissues is very low. Furthermore, the probability of the mixture of WM and CSF is low (Ruan et al., 2000, 2002, 2001), which is why we enable only the WM/GM and the GM/CSF mixture in the brain. The last step of the algorithm consists in applying HMC segmentation with partial volume effects to obtain three maps containing the amounts of each pure tissue in each voxel (see Fig. 6 the complete processing chain).

Algorithm 2 presents the unifying framework we proposed to segment multimodal brain MRI.

Algorithm 2 (Multimodal brain MRI segmentation based on a Hidden Markov Chain).

- (1) Non-rigid Registration of the images with the atlas (Noblet et al., 2005)
- (2) Brain Extraction using BET (Smith, 2002)
- (3) Image Vectorization using 3D Hilbert-Peano scan (Bandoh and Kamata, 1999)

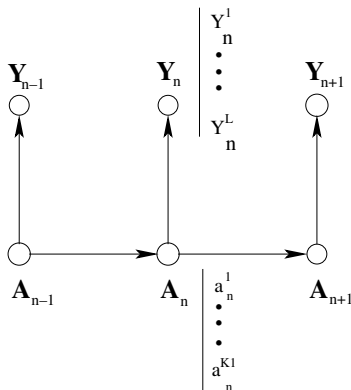


Fig. 5. Hidden Markov Chain with partial volume effect.

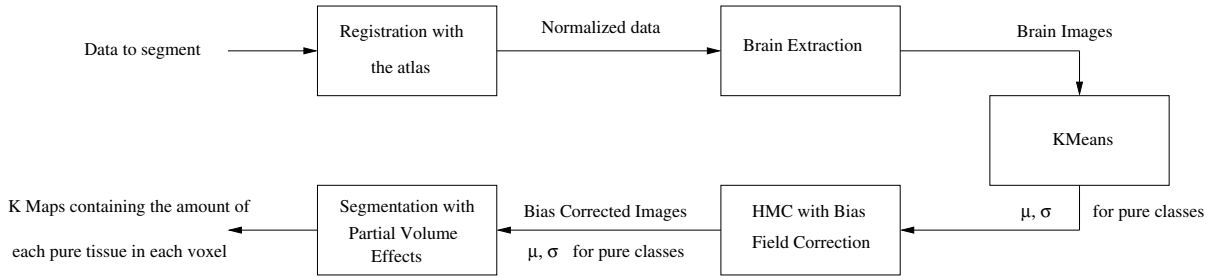


Fig. 6. Algorithm outline.

(4) Initialization

(a) For *a priori* parameters Φ_x

$$\pi_k = \frac{1}{K} \quad \forall k$$

$$t_{ij} = \begin{cases} \frac{3}{4} & \text{if } i = j \\ \frac{1}{4(K-1)} & \text{else} \end{cases}$$

(b) Data-driven parameters Φ_y are initialized using K-Means (Bovik, 2000)

(5) HMC with bias field correction (on log-transformed data), using atlas information. For each iteration

- Calculation of $\alpha_n(k)$ and $\beta_n(k)$ using Eq. (A.6) and (A.8).
- Update of parameters Φ using Eq. (B.7), (B.8), (B.10) and (B.11).

(6) HMC segmentation with partial volume estimation (on non-log-transformed data), without using atlas information and keeping bias field parameters (estimated during the previous step) fixed.

(7) MPM segmentation using Eq. (9)

$$\hat{x}_n = \arg \max_{\omega_k} P(X_n = \omega_k | Y = y, \Phi) = \arg \max_{\omega_k} \alpha_n(k) \beta_n(k)$$

(8) Inverse Hilbert-Peano transformation to obtain the segmented image

4. Validation on synthetic images

4.1. Data

As the ground truth is not known for real data, the developed algorithm has been first validated on simulated images. Knowing the anatomical model, we can have a quantitative assessment of the performance of the different algorithms and then compare them. The Brainweb web-site³ (Kwan et al., 1996, 1997, 1998) offers a large amount of different phantoms of MR brain images with varying noise levels from 0% to 9% and varying levels of non-uniformities from 0% to 40%. From these phantoms the tissue classification in white matter, gray matter, and cerebrospinal fluid is known. We used 18 synthetic volumes of $181 \times 217 \times 181$. The T1-weighted modality are 1 mm \times 1 mm \times 1 mm in spacing. As probability maps, we used the standard atlas provided by SPM to validate our method. Nevertheless, for a specific application, atlas more dedicated to this application can be used to improve segmentation accuracy. All the images of the Brainweb database have been skull stripped using the brain mask obtained from the ground truth.

To evaluate the performance of this algorithm (HMC), we compare the proposed method with three other classic methods: SPM5⁴ (Statistical Parametric Mapping) (Ashburner and Friston,

2005), EMS⁵ (Expectation-Maximization Segmentation) (Van Leemput et al., 1999a) and with the method proposed by Shattuck et al. (2001). The different methods are compared in Table 3. SPM5 and EMS are two free available reference software for brain MRI segmentation. Ashburner and Friston (2005) include tissue classification, bias field correction and image registration with tissue probability maps in the same model. Van Leemput et al. (1999a) propose a framework including bias field correction and tissue classification in a MRF model. Nevertheless, the method presented in Van Leemput et al. (2003) to include partial volume effect is not implemented in EMS. Shattuck et al. (2001) proposed a method including neighborhood information through the use of a Markov prior. Their method compensates for image non-uniformities and combines the estimation of pure tissue fraction with a Gibbs prior modeling of spatial properties of the brain. We first compare the results with PVE presented in Shattuck et al. (2001) with our method. Then we assign each voxel to the most representative class, thereby obtaining a “hard” segmentation and we compare the results with SPM5 and EMS.

4.2. Quantitative results

4.2.1. Evaluation criteria

To evaluate segmentation results with PVE (mixture-based estimation), we compute the RMS error between two fractional volumes (Shattuck et al., 2001, 2004):

$$e_{\text{RMS}}(k) = \sqrt{\frac{1}{\Omega} \sum_{n \in \Omega} |a_n(k) - a_n^f(k)|^2} \quad (37)$$

where Ω is the region of the brain mask, $a_n(k)$ represents the proportion of tissue k at voxel n of the segmented volume, and $a_n^f(k)$ denotes the same proportion in the phantom.

To compare “hard” segmentation results, we use

- True positive fraction (TPF) (or sensitivity):

$$\text{TPF} = \frac{\#(SM \cap GT)}{\#(GT)} = \frac{\#TP}{\#TP + \#FN} \quad (38)$$

- Specificity:

$$\text{SPE} = 1 - \frac{\#(SM - GT)}{\#(GT^c)} = \frac{\#TN}{\#TN + \#FP} \quad (39)$$

This rate is related to the false positive fraction (FPF): $\text{SPE} = 1 - \text{FPF}$

- Kappa index (or Dice similarity coefficient):

$$\text{KI} = 2 \frac{\#(\text{seg} \cap GT)}{\# \text{seg} + \#GT} \quad (40)$$

³ <http://www.bic.mni.mcgill.ca/brainweb/>.

⁴ <http://www.fil.ion.ucl.ac.uk/spm/>.

⁵ <http://www.medicalimagingcomputing.com/EMS/>.

where GT stands for ground truth map, SM for segmented map, TN for true negative map, TP for true positive map, FN for false negative map and FP for false positive map. Furthermore, $\#X$ represents the cardinal of X , $X \cap Y$ is the intersection between the sets X and Y , $X \cup Y$ represents the union of sets X and Y , and $X - Y$ is a set difference: $X - Y = \{u : u \in X \text{ and } u \notin Y\}$. The better are the results, the closer to 100% are the criteria. These three criteria are used when “hard classification” is performed. To evaluate segmentation results using these criteria, we assign each voxel to the most representative tissue class.

4.2.2. Impact of Hilbert-Peano scan on segmentation results

Image vectorization depends on the scan of the 3D image. Hilbert-Peano path is the most suitable scan for image segmentation: indeed this is the scan keeping better neighborhood information.

Table 1

Results obtained on a T1 Brainweb image with 5% noise and 0% inhomogeneity with 3 different Hilbert-Peano scans

| Method | Iter | WM | | GM | | CSF | | τ_{err} |
|-------------|------|--------|----------|-------|--------------|-------|----------|--------------|
| | | μ | σ | μ | σ (%) | μ | σ | |
| HMC, Scan 1 | 16 | 129.87 | 8.54 | 99.25 | 10.56 | 52.16 | 15.85 | 7.75 |
| HMC, Scan 2 | 16 | 129.89 | 8.52 | 99.32 | 10.53 | 52.37 | 16.01 | 7.86 |
| HMC, Scan 3 | 16 | 129.86 | 8.53 | 99.29 | 10.54 | 52.30 | 15.95 | 7.81 |

τ_{err} represents the rate of misclassified voxels.

However, according to the voxel of departure, the vector we obtained will change. Thus, we have tested the impact of this scan on segmentation results. We vectorize the image using the Hilbert-Peano scan starting with three different vertex of the cube (corresponding to the scans 1, 2 and 3). Then HMC segmentation of this vector was carried out (without bias field correction, partial volume effect and atlas information) and the inverse transformation was performed to obtain the segmented image. Results obtained on a T1 Brainweb image with 5% noise and 0% inhomogeneity are shown in Table 1. For each scan, the number of iterations to converge, the rate of misclassified voxels τ_{err} and the estimation of μ and σ for each class are given. We observe that results are similar.

4.2.3. Results on monomodal images

4.2.3.1. Mixture-based segmentation. We test our algorithm on the T1 Brainweb database with 20% non-uniformity and different noise levels. An example of the maps obtained for white matter, gray matter and cerebrospinal fluid is presented in Fig. 7. Each map contains the proportions of each tissue.

Fig. 8 shows the RMS error obtained for both white and gray matter with noise varying from 3% to 9% compared with those presented in Shattuck et al. (2001). Except for white matter with 3% noise, the proposed method yields the best results for both white and gray matter. We notice that the RMS error increases with the noise. In the best case, the HMC method produces RMS error 12.9% better for WM and 17.8% for GM compared to the RMS error given by Shattuck et al. (2001).

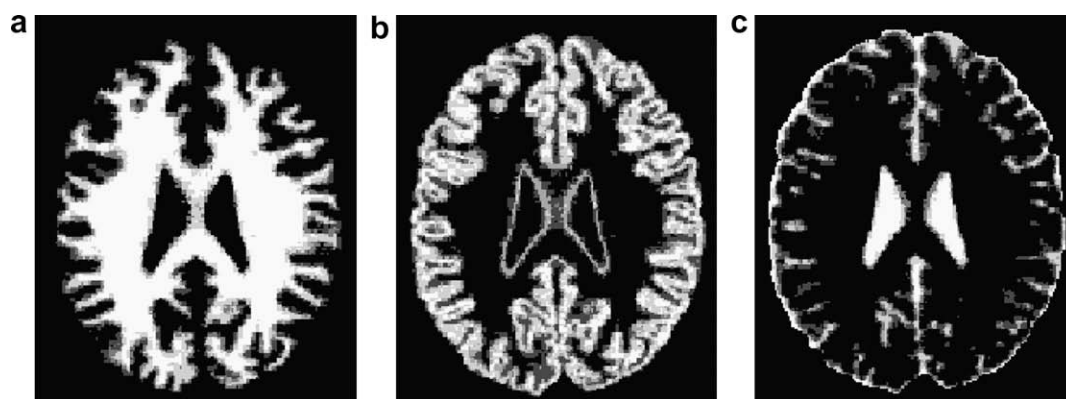


Fig. 7. Maps obtained on Brainweb image with 5% noise level and 20% inhomogeneity. (a), (b) and (c) correspond, respectively, to white matter, gray matter, and cerebrospinal fluid. RMS error obtained for this image is presented in Fig. 8.

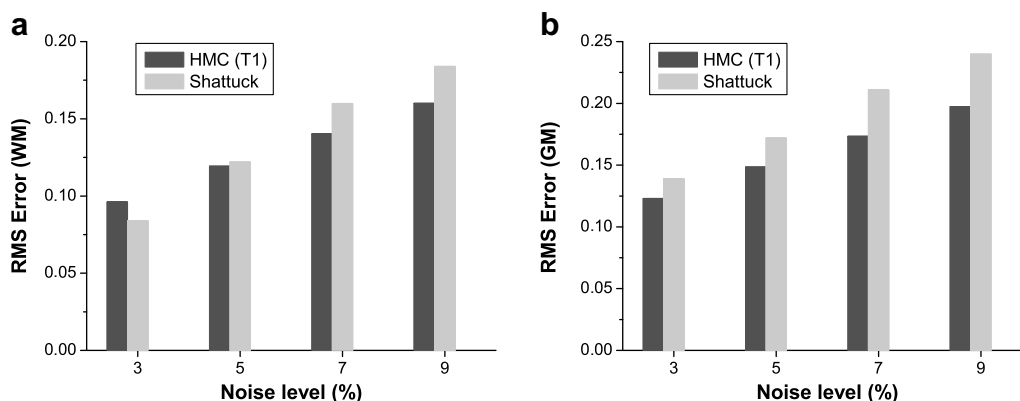


Fig. 8. RMS error obtained on the T1 Brainweb database with 20% inhomogeneity with the proposed method and the method of Shattuck et al. (2001). (a) and (b) correspond, respectively, to the results obtained for white matter (WM) and gray matter (GM).

4.2.4. Hard segmentation

Then we assign each voxel to the most representative tissue class to obtain a hard segmentation in three classes. The hard segmentation obtained with the different methods tested for the Brainweb image with 5% noise and 20% inhomogeneity is presented in Fig. 9. We compare our proposed method (HMC) with EMS and SPM5. Fig. 10 presents the Kappa index obtained with different segmentation methods on the Brainweb database with 20% inhomogeneity and noise level varying from 0% to 9%. For white and gray matter, the proposed method yielded the best results for each noise level. We notice that the index Kappa for the proposed HMC method decreases when the noise level increases. For

EMS and SPM5, the Kappa index is maximum between 3% and 5% noise which corresponds in practice to the noise level in real images. However, even for these noise levels the proposed HMC method included PVE obtains the most accurate segmentation.

4.2.5. Results on multimodal images

4.2.5.1. Mixture-based segmentation. The Brainweb database provides multimodal images with different levels of noise and inhomogeneity. As the approach proposed by Shattuck et al. (2001) has only been applied to single-spectral images, we compare results on T1/T2 multimodal images with those obtained in the monomodal case with our HMC proposed method. Fig. 11 shows

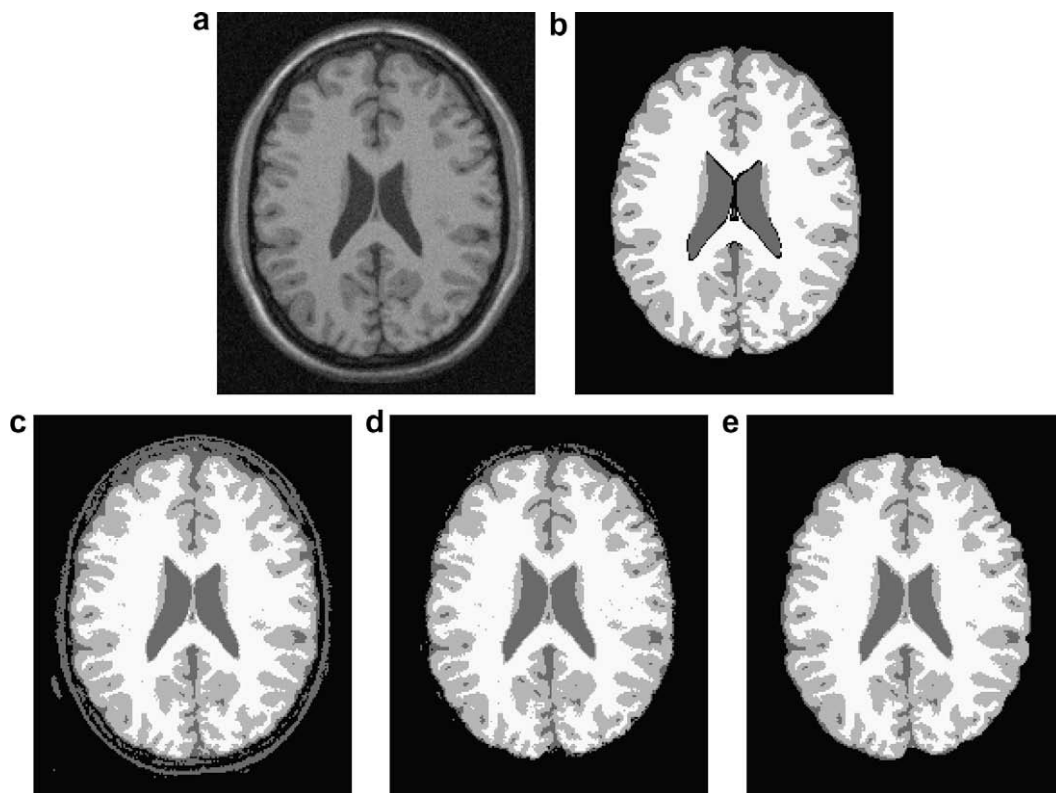


Fig. 9. Results obtained on Brainweb image with 5% noise level and 20% inhomogeneity. (a) corresponds to the data to be segmented and (b) to the ground truth. (c), (d) and (e) correspond, respectively, to the results obtained with SPM5, EMS, and the proposed method. Kappa index obtained for this image with the different methods is presented in Fig. 10.

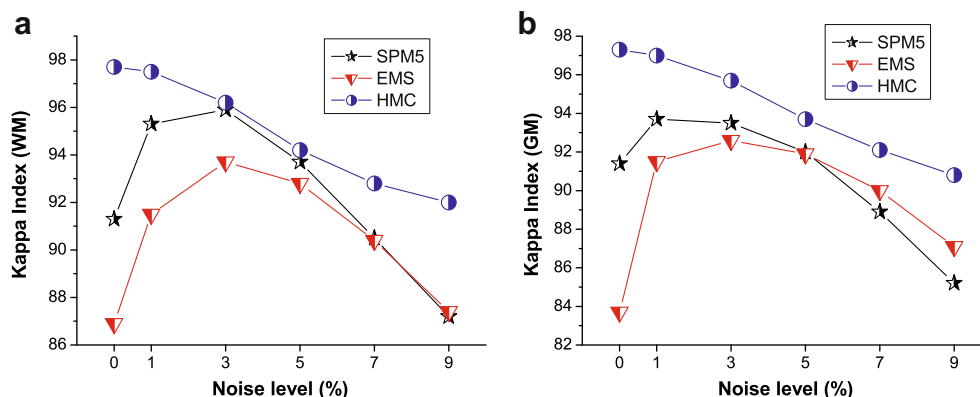


Fig. 10. Comparison of the results obtained on the T1 Brainweb database with 20% inhomogeneity and with three different methods: SPM5, EMS, and the HMC proposed method. (a) and (b) correspond, respectively, to the Kappa index obtained for white matter (WM) and gray matter (GM).

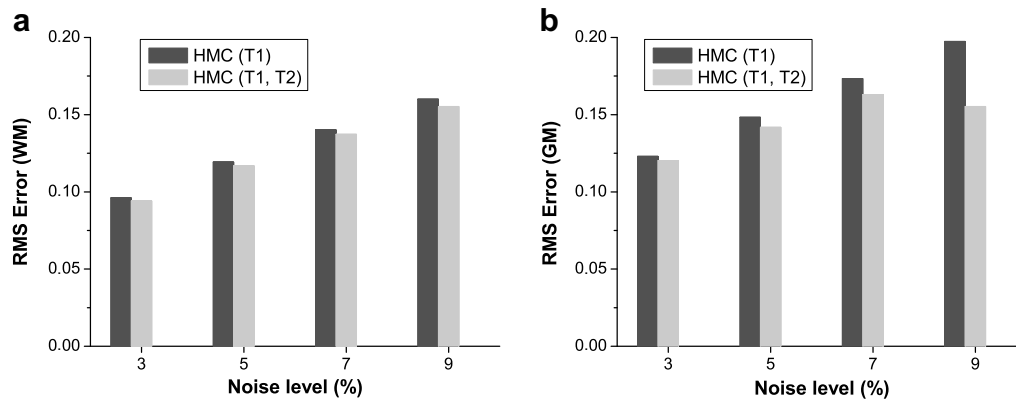


Fig. 11. RMS error obtained on the Brainweb database with 20% inhomogeneity for both monomodal and multimodal images. (a) and (b) correspond, respectively, to the results obtained for white matter (WM) and gray matter (GM).

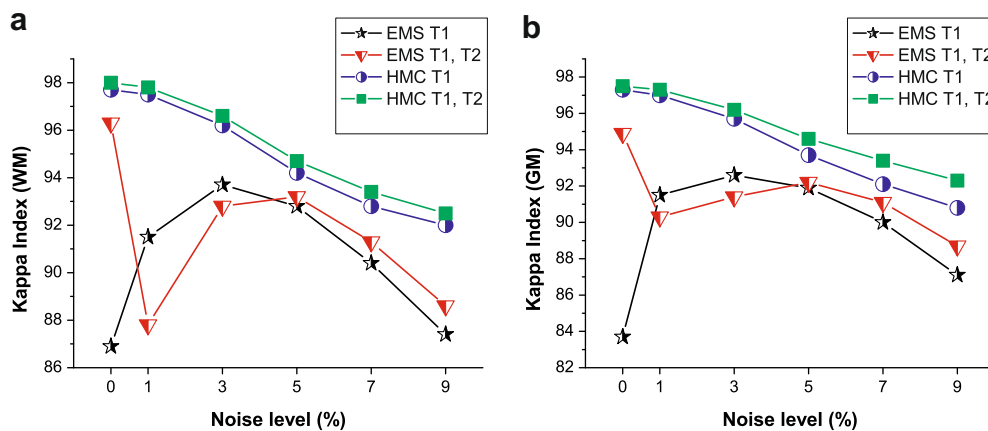


Fig. 12. Comparison of the results obtained on the T1 Brainweb database with 20% inhomogeneity and with different methods: EMS and the HMC proposed method both on monomodal (T1) and multimodal (T1, T2) images. (a) and (b) correspond, respectively, to the Kappa index obtained for white matter (WM) and gray matter (GM).

the RMS error obtained for both monomodal and multimodal images. The use of multimodal images brings more information and decreases the RMS error for both white and gray matter. Multimodality improves results from 2% to 3% for WM and for 2.7% to 7.3% for GM.

4.2.6. Hard segmentation

After the mixture-based segmentation step, we assign each voxel to the most representative class to obtain a hard segmentation into three classes. Comparisons on this dataset with EMS were carried out. The graphs in Fig. 12 emphasize that multimodality improves the results. For white and gray matter, HMC Kappa index on T1 and T2 images is higher than index on T1 images and than multimodal EMS.

5. Validation on real data

The 18 normal MR brain datasets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at <http://www.cma.mgh.harvard.edu/ibsr/>. Only the “hard” segmentation provided by experts is available in the Internet Brain Segmentation Repository (IBSR). The dimensions for the images were $256 \times 256 \times 128$. We use the evaluation criteria presented in Section 4.2.1. To use these criteria, we need a segmentation into three classes. Therefore, after applying the proposed algorithm (cf. Fig. 6),

we perform a hard segmentation, by assigning each voxel to the most representative class.

Dealing with real data, we are faced to problems using BET (Smith, 2002) in order to separate the brain from non-brain tissue. Indeed some non-brain tissue still appear in the images, degrading the segmentation results. To carry out this issue, the atlas was used to discard non-brain tissues: voxels where the atlas indicates a probability of zero of being white matter, gray matter, or cerebrospinal fluid were removed.

Each case of the IBSR database was segmented with the three different methods: SPM5, EMS, and our proposed method. An example of the results obtained on IBSR data is presented in Fig. 13. The different coefficients were computed for both white matter and gray matter segmentation and the average of the results obtained on the 18 IBSR images is reported in Table 2. Fig. 14 compares the performance of the different methods on each image of the IBSR dataset. As expert segmentation only includes internal fluid (ventricles) and our method also segments external fluid, we do not report results for cerebrospinal fluid.

White matter sensitivity is higher with the proposed method, meaning that we detect white matter as white matter: our method produces a WM sensitivity of 90% whereas SPM5 and EMS produce a WM sensitivity of 84.7 and 81, respectively. Gray matter sensitivity is lower for the proposed method. However, gray matter specificity is higher. The proposed method detects less gray matter compared with the two other methods, but it does not classify

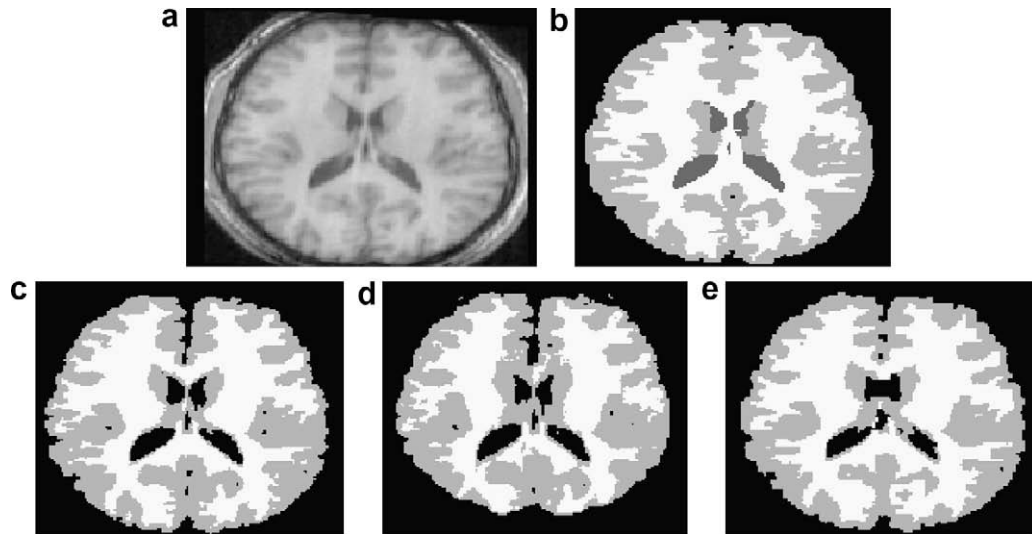


Fig. 13. Results obtained on the image IBSR 02. (a) corresponds to the data to be segmented and (b) to the expert segmentation. (c), (d), and (e) correspond, respectively, to the results obtained with SPM5, EMS, and the proposed method. Kappa index obtained for this image with the different methods is presented in Fig. 14.

Table 2
Results obtained on 18 Brains from the IBSR database with three different methods

| Algorithm | Criterion | White matter (%) | | Gray matter (%) | |
|-----------------|-------------|------------------|--------------------|-----------------|--------------------|
| | | Mean | Standard deviation | Mean | Standard deviation |
| SPM5 | Sensitivity | 84.72 | 5.52 | 79.97 | 8.81 |
| | Specificity | 99.38 | 0.18 | 97.56 | 4.48 |
| | Kappa index | 85.27 | 4.13 | 78.7 | 13.98 |
| EMS | Sensitivity | 81.09 | 4.92 | 72.86 | 7.62 |
| | Specificity | 99.67 | 0.15 | 99.03 | 0.26 |
| | Kappa index | 85.87 | 2.27 | 78.94 | 5.68 |
| Proposed method | Sensitivity | 90.64 | 2.53 | 71.78 | 6.85 |
| | Specificity | 99.16 | 0.24 | 99.35 | 0.26 |
| | Kappa index | 86.53 | 1.73 | 79.94 | 5.57 |

Averaged results on the 18 images are presented.

other substances as gray matter. Furthermore, the proposed method obtained higher Kappa index than SPM5 and EMS for both white and gray matter. Nevertheless, the segmentation of two images of the IBSR database obtained with SPM5 was completely false

(Fig. 14), particularly for gray matter which explains the high standard deviation. When we do not include these two images in the average for SPM5, the results for white matter are closer than from the HMC and higher for gray matter. Fig. 14 presents the Kappa index obtained for each image of the IBSR dataset with the three different methods. In some cases, the proposed framework obtains the most accurate white matter segmentation. The white matter Kappa index lies between 82 and 90. The average from the 18 images is 86.5 with a low standard deviation. Concerning gray matter, the Kappa index lies between 70 and 90, with a mean of 79.9, higher than the average obtained with EMS and SPM5 and a lower standard deviation.

6. Discussion

In this paper, we propose a unifying framework for unsupervised segmentation of multimodal brain MR images using Hidden Markov Chains. Table 3 shows the comparison of the performance of different methods we tested. Tests on both synthetic and real data have shown that the global framework that brings neighborhood information using HMC-based regularization scheme, outperforms SPM5, EMS and the method proposed by Shattuck. The HMC model has already shown advantages compared to the

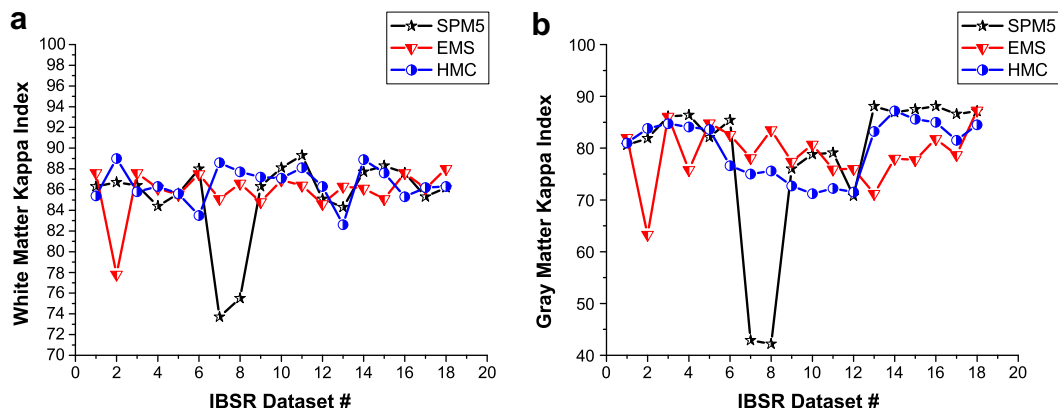


Fig. 14. Comparison of the results obtained on the IBSR dataset with the different methods: SPM5, EMS, and the proposed method. Kappa index is reported for: (a) white matter and (b) gray matter.

Table 3
Comparison of different segmentation methods

| | SPM5 | EMS | Shattuck | HMC |
|----------------|------------------|------------------------|------------------------|-----------|
| Bias field | Yes | Yes | Yes | Yes |
| PVE | Yes ^a | No | Yes | Yes |
| Neighborhood | No | Yes ^b (MRF) | Yes (MRF) | Yes (HMC) |
| Unsupervised | Yes | Yes | Partially ^c | Yes |
| Atlas | Yes | Yes | No | Yes |
| Multimodal | No | Yes | No | Yes |
| Computing time | 20–35' | 20–35' | 5' | 20–40' |

^a Proportions of each tissue in each voxel are not computed.

^b Parameter estimation is based on a limited subset of voxels.

^c The β term is fixed.

MRF model for 2D SAR images (Fjortoft et al., 2003) and has been applied for the first time to 3D brain MR images in this paper using the 3D Hilbert-Peano scan to vectorize the images. The influence of the Hilbert-Peano scan on the parameter estimation was studied and the tests have shown that segmentation results do not depend on any particular scan amongst several Hilbert-Peano paths generated by modifying the starting point. The proposed method also includes bias field correction by adapting the method proposed by Van Leemput et al. (1999a) to the HMC model. Furthermore, information provided by probability maps is used as a complementary sensor in an original way. The information brought by this atlas is used at each iteration to estimate model parameters whereas other methods only use the atlas to initialize the segmentation process. SPM5 also uses information brought by probability maps but does not include neither neighborhood information nor multimodal data. To cope with partial volume effects, they simply assume that the intensity distribution of each class may not be Gaussian and assign belonging probabilities according to these non-Gaussian distributions. They only provide maps of the posterior probabilities. Nevertheless, they do not compute the relative amount of each pure tissue in each voxel. Another contribution of this paper lies in the approach proposed to estimate the tissue fraction of WM, GM and CSF. For each voxel n , we compute the vector \mathbf{A}_n containing the proportion a_n^i of each pure class i . Using the HMC model, a Markovian prior is introduced to model the spatial correlation between voxels. Shattuck et al. (2001) first classify the brain into different PV labels using a Potts model and then assign a tissue fraction of WM, GM, or CSF to each voxel. In this step, spatial properties of the brain are not included whereas our method proposes a mixture-based HMC segmentation which assigns directly to each voxel a vector containing the amounts of each pure tissue using neighborhood information. Moreover, in their Potts model the user has to chose β term controlling the strength of the prior. Thus, their method is not totally unsupervised. Comparisons with Shattuck et al. (2001) have been performed on Brainweb images. Our method provides a more accurate estimation of the tissue fraction in each voxel. Comparisons with EMS and SPM5 were also carried out for the segmentation in “hard” classes on both synthetic and real images, demonstrating the advantages of the method which we propose. Concerning the computing time, between 20 and 40 min are required in here-proposed method to segment each image and to compute the relative amount of each tissue in each voxel. Segmentation using SPM5 or EMS requires between 20 and 35 min. However, SPM5 does not include neighborhood information and EMS estimates the parameters on a limited subset of voxels to speed up the algorithm, whereas our method carry out the estimation of the mixture parameters on the entire set of voxels. All computation times are on a PC with a Pentium IV 2.4 GHz processor running Linux with 1 Gigabit RAM.

To conclude, we have presented a fully automated algorithm to segment multimodal MR images. We incorporate in the HMC model *a priori* information given by probability maps from a brain atlas, including bias field inhomogeneities and partial volume effects. Both simulated and real data were used to evaluate the performance of the algorithm, yielding accurate results. To validate this approach, we have applied our method to several healthy brains from a database. Comparisons with other methods show that our method outperformed several published methods. Future work will consist in extending this model to detect lesions in the brain. Another possible extension could be to include topology constraints: some recent models (Huang, 2004, 2005) combine MRFs and deformable models in a graphical model-based framework.

Acknowledgement

The authors would like to acknowledge the region Alsace and CNRS for support of this research project.

Appendix A

In this section, we derive the equations of the *forward* and *backward* probabilities with atlas information. We have for the *forward* probability:

$$\alpha_n \propto P(X_n, \mathbf{Y}_{\leq n}, B_{\leq n}) \quad (\text{A.1})$$

This leads to

$$\begin{aligned} P(X_n = \omega_k, \mathbf{Y}_{\leq n}, B_{\leq n}) &= \sum_{\omega_l} P(X_{n-1} = \omega_l, X_n = \omega_k, \mathbf{Y}_{\leq n-1}, \mathbf{Y}_n, B_{\leq n-1}, B_n) \\ &= \sum_{\omega_l} P(X_{n-1} = \omega_l, \mathbf{Y}_{\leq n-1}, B_{\leq n-1}) P(X_n = \omega_k | X_{n-1} = \omega_l) \\ &\quad \times P(B_n | X_n = \omega_k) P(\mathbf{Y}_n | X_n = \omega_k) \\ &= \sum_{\omega_l} \alpha_{n-1}(l) P(X_n = \omega_k | X_{n-1} = \omega_l) P(B_n | X_n = \omega_k) P(\mathbf{Y}_n | X_n = \omega_k) \end{aligned} \quad (\text{A.2})$$

We have for the *backward* probability:

$$\beta_n \propto P(\mathbf{Y}_{>n}, B_{>n} | X_n) \quad (\text{A.3})$$

This leads to

$$\begin{aligned} P(\mathbf{Y}_{>n}, B_{>n} | X_n = \omega_k) &= \sum_{\omega_l} P(\mathbf{Y}_{>n}, B_{>n}, X_{n+1} = \omega_l | X_n = \omega_k) \\ &= \sum_{\omega_l} [P(\mathbf{Y}_{>n+1}, B_{>n+1} | X_{n+1} = \omega_l) P(B_{n+1} | X_{n+1} = \omega_l) \\ &\quad \times P(\mathbf{Y}_{n+1} | X_{n+1} = \omega_l) P(X_{n+1} = \omega_l | X_n = \omega_k)] \\ &= \sum_{\omega_l} [\beta_{n+1}(l) P(B_{n+1} | X_{n+1} = \omega_l) P(\mathbf{Y}_{n+1} | X_{n+1} = \omega_l) P(X_{n+1} = \omega_l | X_n = \omega_k)] \end{aligned} \quad (\text{A.4})$$

Forward and *backward* probabilities can be computed as follows:

A.1. Computation of α

- For $n = 1$

$$\alpha_1(k) = \pi_k f_k(\mathbf{y}_1) b_1(k) \quad \forall k = 1, \dots, K \quad (\text{A.5})$$

Probabilities are then normalized by dividing $\alpha_1(k) \forall k = 1, \dots, K$ by $N_1 = \sum_{k=1}^K \alpha_1(k)$.

- For $n > 1$

$$\alpha_n(k) = \sum_{l=1}^K \alpha_{n-1}(l) t_{lk} f_k(\mathbf{y}_n) b_n(k) \quad \forall k = 1, \dots, K \quad (\text{A.6})$$

where $b_n(k)$ denotes the probability of voxel n belonging to tissue k given by the atlas. Probabilities are then normalized by dividing $\alpha_n(k) \forall k = 1, \dots, K$ by $N_n = \sum_{k=1}^K \alpha_n(k)$.

A.2. Computation of β

- For $n = N$

$$\beta_N(k) = 1 \quad \forall k = 1, \dots, K \quad (\text{A.7})$$

- For $n < N$

$$\beta_n(k) = \frac{1}{N_{n+1}} \sum_{l=1}^K \beta_{n+1}(l) t_{kl} f_l(\mathbf{y}_{n+1}) b_{n+1}(l) \quad \forall k = 1, \dots, K \quad (\text{A.8})$$

Appendix B

In this section, we derive the GEM algorithm of Section 3.2 with bias field estimation step in the monomodal case. z_n denotes the log-transformed data and $\Phi = \{\Phi_x, \Phi_y\}$ contains the *a priori* parameters $\Phi_x = \{\pi_k, t_{kl}\}$, and the *data-driven* parameters Φ_y

$$Q(\Phi, \Phi^{[q]}) = E[\log p(x, z|\Phi) | z, \Phi^{[q]}] \quad (\text{B.1})$$

$$= \sum_i \gamma_i^{[q]}(i) \log \pi_i + \sum_{n \neq 1} \sum_i \sum_j \xi_n^{[q]}(i, j) \log t_{ij} + \sum_n \sum_i \gamma_n^{[q]}(i) \log f_i(z_n) \quad (\text{B.2})$$

with:

- *a posteriori* marginal probability

$$\gamma_n^{[q]}(i) = p(x_n = \omega_i | z, \Phi^{[q]}) = \frac{\alpha_n(i) \beta_n(i)}{\sum_j \alpha_n(j) \beta_n(j)} \quad (\text{B.3})$$

- *a posteriori* joint probability

$$\xi_n^{[q]}(i, j) = p(x_{n-1} = \omega_j, x_n = \omega_i | z, \Phi^{[q]}) = \alpha_{n-1}(j) t_{ji} f_i(z_n) \beta_n(i) \quad (\text{B.4})$$

- initial distribution

$$\pi_i = p(x_1 = \omega_i) \quad (\text{B.5})$$

- transition matrix

$$t_{ij}^n = p(x_{n+1} = \omega_j | x_n = \omega_i) \quad (\text{B.6})$$

We maximize this quantity under the following constraints: $\sum_i \pi_i = 1$ and $\sum_j t_{ij} = 1$. Using Lagrangian multipliers, we obtain

$$\pi_i^{[q+1]} = \gamma_1^{[q]}(i) \quad (\text{B.7})$$

$$t_{ij}^{[q+1]} = \frac{\sum_{n=2}^N \xi_n^{[q]}(i, j)}{\sum_{n=1}^{N-1} \gamma_n^{[q]}(i)} \quad (\text{B.8})$$

$$\sum_n \sum_i \gamma_n^{[q]}(i) \log f_i(z_n) = \sum_n \sum_i \gamma_n^{[q]}(i) \left[-\log \sigma_i - \frac{1}{2\sigma_i^2} \left(z_n - \mu_i - \sum_k c_k \psi_k(t_n) \right)^2 \right] \quad (\text{B.9})$$

Setting the partial derivative for μ_i and σ_i of $\sum_n \sum_i \gamma_n^{[q]}(i) \log f_i(z_n)$ to zero leads to

$$\mu_i^{[q+1]} = \frac{\sum_n \gamma_n^{[q]}(i) [z_n - \sum_k c_k \psi_k(t_n)]}{\sum_n \gamma_n^{[q]}(i)} \quad (\text{B.10})$$

$$\sigma_i^{2[q+1]} = \frac{\sum_n \gamma_n^{[q]}(i) [z_n - \mu_i - \sum_k c_k \psi_k(t_n)]^2}{\sum_n \gamma_n^{[q]}(i)} \quad (\text{B.11})$$

Setting the partial derivative for c_k to zero leads to Van Leemput et al. (1999a)

$$\sum_n \psi_k(x_n) \sum_i \frac{\gamma_n^{[q]}(i)}{\sigma_i^2} [z_n - \mu_i - \sum_l c_l \psi_l(t_n)] = 0 \quad \forall k \quad (\text{B.12})$$

This yields to Van Leemput et al. (1999a)

$$C = (D^T W D)^{-1} D^T W R \quad (\text{B.13})$$

$$C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_K \end{bmatrix}, \quad R = \begin{bmatrix} y_1 - \bar{y}_1 \\ y_2 - \bar{y}_2 \\ \vdots \end{bmatrix}, \quad D = \begin{bmatrix} \psi_1(t_1) & \psi_2(t_1) & \dots \\ \psi_1(t_2) & \psi_2(t_2) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad (\text{B.14})$$

with $w_{ni} = \frac{\gamma_n(i)}{\sigma_i^2}$, $w_n = \sum_i w_{ni}$, $W = \text{diag}(w_n)$, and the estimated bias corrected data $\bar{y}_n = \frac{\sum_i w_{ni} \mu_i}{\sum_i w_{ni}}$.

References

- Al-Zubi, S., Toennies, K., Bodammer, N., Hinrichs, H., 2002. Fusing Markov random fields with anatomical knowledge and shape based analysis to segment multiple sclerosis white matter lesions in magnetic resonance images of the brain. In: *Bildverarbeitung für die Medizin*. Springer, pp. 185–188.
- Ashburner, J., Friston, K., 1997. Multimodal image coregistration and partitioning – a unified framework. *NeuroImage* 6 (3), 209–217.
- Ashburner, J., Friston, K., 2000. Voxel-based morphometry – the methods. *NeuroImage* 11 (6), 805–821.
- Ashburner, J., Friston, K., 2005. Unified segmentation. *NeuroImage* 26 (3), 839–851.
- Bandoh, Y., Kamata, A., 1999. An address generator for a 3-dimensional pseudo-Hilbert scan in a cuboid region. In: *IEEE International Conference on Image Processing (ICIP'99)*, pp. 1:496–500.
- Baum, L., 1972. An inequality and associated maximization technique in statistical estimation for probabilistic function of Markov processes. *Inequalities* 3, 1–8.
- Belaroussi, B., Milles, J., Carme, S., Zhu, Y.M., Benoit-Cattin, H., 2006. Intensity non-uniformity correction in MRI: existing method and their validation. *Medical Image Analysis* 10 (2), 234–246.
- Bosc, M., Heitz, F., Armspach, J.-P., 2003. Statistical atlas-based sub-voxel segmentation of 3D brain MRI. In: *IEEE International Conference on Image Processing (ICIP'03)*, pp. 1077–1080.
- Bovik, A., 2000. *Handbook of Image and Video Processing*. Academic Press, San Diego.
- Bricq, S., Collet, C., Armspach, J.-P., 2006. Triplet markov chain for 3d mri brain segmentation using a probabilistic atlas. In: *IEEE International Symposium on Biomedical Imaging (ISBI'06)*, pp. 386–389.
- Choi, H.S., Haynor, D.R., Kim, Y.M., 1991. Partial volume tissue classification of multichannel magnetic resonance images – a mixel model. *IEEE Transactions on Medical Images* 10, 395–407.
- Cocosco, C., Kollokian, V., Kwan, R.-S., Evans, A., 1997. BrainWeb: online interface to a 3D MRI simulated brain database. In: *Proceedings of 3rd International Conference on Functional Mapping of the Human Brain*, vol. 5, p. 425.
- Collins, D., Zijdenbos, A., Kollokian, V., Sled, J., Kabani, N., Holmes, C., Evan, A., 1998. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging* 17 (3), 463–468.
- Dempster, A., Laird, N., Rubin, D., 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B* 39, 1–38.
- Devijver, P.A., 1985. Baum's forward-backward algorithm revisited. *Pattern Recognition Letters* 3 (6), 369–373.
- Dugas-Phocion, G., González Ballester, M.A., Malandain, G., Lebrun, C., Ayache, N., 2004. Improved EM-based tissue segmentation and partial volume effect quantification in multi-sequence brain MRI. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'04)*. Lecture Notes in Computer Science, vol. 3216, pp. 26–33.
- Fjortoft, R., Delignon, Y., Pieczynski, W., Sigelle, M., Tupin, F., 2003. Unsupervised classification of radar images using Hidden Markov Chains and Hidden Markov Random Fields. *IEEE Transactions on Geoscience and Remote Sensing* 41 (3), 675–686.
- Forney, G.D., 1973. The Viterbi algorithm. *Proceedings of the IEEE* 61 (3), 268–278.
- Gelman, A., Carlin, J., Stern, H., Rubin, D., 2005. *Bayesian Data Analysis*. Chapman and Hall, New York.
- Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6 (6), 721–741.
- Graffigne, C., Heitz, F., Pérez, P., Prêteux, F., Sigelle, M., Zerubia, J., 1995. Hierarchical Markov random field models applied to image analysis: a review. In: *SPIE Neural Morphological and Stochastic Methods in Image and Signal Processing*, vol. 2568, pp. 2–17.
- Guillemaud, R., Brady, M., 1997. Estimating the bias field of MR images. *IEEE Transactions on Medical Imaging* 16 (3), 238–251.
- Huang, R., Pavlovic, V., Metaxas, D.N., 2004. A graphical model framework for coupling MRFs and deformable models. In: *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'04)*, pp. 739–747.
- Huang, R., Pavlovic, V., Metaxas, D.N., 2005. A hybrid framework for image segmentation using probabilistic integration of heterogeneous constraints. In: *Computer Vision for Biomedical Image Application: Current Techniques and Future Trends*. Springer, pp. 82–92.
- Kwan, R.-S., Evans, A., Pike, G., 1996. An extensible MRI simulator for post-processing evaluation. *Proceedings of the International Conference on*

- Visualization in Biomedical Computing, VBC'96, vol. 1131. Springer-Verlag, pp. 135–140.
- Lanchantin, P., Pieczynski, W., 2004. Unsupervised non stationary image segmentation using Triplet Markov Chains. In: *Advanced Concepts for Intelligent Vision Systems (ACIVS 04)*, pp. 60–66.
- Marroquin, J., Vemuri, B., Botelo, S., et al., 2002. An accurate and efficient Bayesian method for automatic segmentation of brain MRI. *IEEE Transactions on Medical Imaging* 21 (8), 934–945.
- Masson, P., Pieczynski, W., 1993. SEM algorithm and unsupervised statistical segmentation of satellite images. *IEEE Transactions on Geoscience and Remote Sensing* 31 (3), 618–633.
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., Teller, E., 1953. Equation of state calculations by fast computing machines. *Journal of Chemical Physics* 21, 1087–1091.
- Noblet, V., Heinrich, C., Heitz, F., Armspach, J., 2005. 3-D deformable image registration: a topology preservation scheme based on hierarchical deformation models and interval analysis optimization. *IEEE Transactions on Image Processing* 14 (5), 553–566.
- Pérez, P., Chardin, A., Laferté, J.-M., 2000. Noniterative manipulation of discrete energy-based models for image analysis. *Pattern Recognition* 33 (4), 573–586.
- Pieczynski, W., 1992. Statistical image segmentation. *Machine Graphics and Vision* 1 (2), 261–268.
- Pieczynski, W., 2003. Pairwise Markov Chains. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25 (5), 634–639.
- Pieczynski, W., Hulard, C., Veit, T., 2002. Triplet Markov Chains in hidden signal restoration. In: *SPIE International Symposium on Remote Sensing*, pp. 58–68.
- Provost, J., Collet, C., Rostaing, P., Pérez, P., Bouthemy, P., 2003. Hierarchical Markovian segmentation of multispectral images for the reconstruction of water depth maps. *Computer Vision and Image Understanding* 93 (2), 155–174.
- Romberg, J., Choi, H., Baraniuk, R., 2001. Bayesian tree-structured image modeling using wavelet domain Hidden Markov Models. *IEEE Transactions on Image Processing* 10 (7), 1056–1068.
- Ruan, S., Jaggi, C., Fadili, J., Bloyet, D., 2000. Brain tissue classification of magnetic resonance images using partial volume modeling. *IEEE Transactions on Medical Imaging* 19 (12), 1179–1187.
- Ruan, S., Moretti, B., Fadili, J., Bloyet, D., 2002. Fuzzy Markovian segmentation in application of magnetic resonance images. *Computer Vision and Image Understanding* 85, 54–69.
- Shattuck, D.W., Sandor-Leahy, S.R., Schaper, K.A., Rottenberg, D.A., Leahy, R.M., 2001. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage* 13 (5), 856–876.
- Sled, J.G., Pike, B.G., 1998. Understanding intensity non-uniformity in MRI. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'98)*. Lecture Notes in Computer Science, vol. 1496. Springer Verlag, pp. 614–622.
- Sled, J.G., Zijdenbos, A.P., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging* 17 (1), 87–97.
- Smith, S., 2002. Fast robust automated brain extraction. *Human Brain Mapping* 17 (3), 143–155.
- Stark, H., Woods, J.W., 1986. *Probability, Random Processes, and Estimation Theory for Engineers*. Prentice-Hall, Englewood Cliffs, NJ.
- Styner, M., Brechbühler, C., Székely, G., Gerig, G., 2000. Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Transactions on Medical Imaging* 19 (3), 153–165.
- Tanner, M., 1993. *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. Springer-Verlag.
- Tohka, J., Zijdenbos, A., Evans, A., 2004. Fast and robust parameter estimation for statistical partial volume models in brain MRI. *NeuroImage* 23 (1), 84–97.
- Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P., 1999a. Automated model-based bias field correction of MR images of the brain. *IEEE Transactions on Medical Imaging* 18 (10), 885–896.
- Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P., 1999b. Automated model-based tissue classification of MR images of the brain. *IEEE Transactions on Medical Imaging* 18 (10), 897–908.
- Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P., 2003. A unifying framework for partial volume segmentation of brain MR images. *IEEE Transactions on Medical Imaging* 22 (1), 10–113.
- Wells, W.M., Grimson, W.E.L., Kikinis, R., Jolesz, F.A., 1996. Adaptive segmentation of MRI data. *IEEE Transactions on Medical Imaging* 15 (4), 429–442.