NOTE

Fuzzy Markovian Segmentation in Application of Magnetic Resonance Images

Su Ruan, Bruno Moretti, Jalal Fadili, and Daniel Bloyet
GREYC-ISMRA, CNRS UMR 6072, F-14050 Cedex Caen, France
E-mail: su.ruan@greyc.ismra.fr

Received July 6, 2001; revised March 25, 2002

In this paper, we present a fuzzy Markovian method for brain tissue segmentation from magnetic resonance images. Generally, there are three main brain tissues in a brain dataset: gray matter, white matter, and cerebrospinal fluid. However, due to the limited resolution of the acquisition system, many voxels may be composed of multiple tissue types (partial volume effects). The proposed method aims at calculating a fuzzy membership in each voxel to indicate the partial volume degree, which is statistically modeled. Since our method is unsupervised, it first estimates the parameters of the fuzzy Markovian random field model using a stochastic gradient algorithm. The fuzzy Markovian segmentation is then performed automatically. The accuracy of the proposed method is quantitatively assessed on a digital phantom using an absolute average error and qualitatively tested on real MRI brain data. A comparison with the widely used fuzzy C-means algorithm is carried out to show numerous advantages of our method.

Key Words: fuzzy segmentation; Markovian random fields; brain tissue; partial volume effects.

INTRODUCTION

Image segmentation is a classical problem in computer vision and is of paramount importance to medical imaging. The study of many brain disorders involves accurate tissue segmentation from magnetic resonance (MR) images of the brain. Manual delineation of the three brain tissues, white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF), in MR images by a human expert is too time-consuming for studies involving larger databases. In addition, the lack of clearly defined edges induces large intra- and interobserver variability, which deteriorates the significance of the analysis of the resulting segmentation, thus calling for automatic segmentation methods. Automated and reliable tissue classification is
Further complicated by the overlap of MR intensities of different tissue classes (partial volume effects) and by the presence of a spatially smoothly varying intensity inhomogeneity. Especially, the problem of partial volume effects, which provide a hard classification where each voxel is assigned to a single tissue, is often ignored by many proposed methods [1, 2].

Two main reasons lead to the problem of partial volume effects. On the one hand, due to the imaging resolution, the complexity of tissue boundaries causes many voxels to be composed of at least two or more tissues (Fig. 1). On the other hand, the constitution of a brain cannot be restricted to only three pure tissues (GM, WM, and CSF). Some internal brain structures should be distinguished from the pure tissues. In fact, voxels located in these subcortical regions, such as the putamen and thalamus, are known to have a high admixture of GM and WM [3]. Their intensities vary often between that of pure tissues; precisely, they extend between the gray matter and the white matter. A hard segmentation of the three tissues may split the structures, where one part is labeled gray matter and the other part white matter. Therefore, in order to obtain accurate segmentation, it is necessary to determine the degree to which voxels are similar to, or belong to, one or more tissue categories. Recently some approaches have been presented to deal with these problems [4–6]. An extensive review of this issue is given in [7]. The unsupervised segmentation based on the fuzzy C-means (FCM) algorithm [8] has been used with some success in the fuzzy segmentation of MR images [9], as well as for the estimation of partial volumes [10]. It clusters data by computing a measure of fuzzy membership for each voxel given a specified number of classes. Barra et al. [11] used the wavelet detail coefficients as spatial and textural information in the possibilistic fuzzy clustering algorithm to improve the performance of the classification. The approaches based on the fuzzy connectedness [12, 13] have demonstrated their effectiveness in many MR images, such as the detection of multiple sclerosis lesions [14]. The concept related to fuzzy connectedness makes it possible to handle global three-dimensional object characteristics in a fuzzy framework, allowing us to deal realistically with global object-related uncertainties [14]. The approach presented in this paper differs from these published methods in merging fuzzy and statistical aspects of methods. This leads to our fuzzy statistical segmentation.
Both fuzzy logic and statistics communities now agree that fuzzy and probabilistic approaches are complementary rather than competitive. The distinction between them is based on how uncertainty is captured by each concept. In randomness, the uncertainty is derived from the nondeterministic membership of a point to a well-defined region of the sample space, while in fuzziness, the uncertainty is derived from the partial membership of a point to an imprecisely defined region (fuzzy set) in the space of discourse. Thus, probability and membership degree are measures reflecting different manners of capturing uncertain information. The essence of our model is to exploit the complementarity of fuzzy and probabilistic approaches, combining their strengths using a fuzzy Markovian random field (MRF). In our segmentation scheme, the uncertainty about the observation field (image intensity) is modeled in a fully statistical framework. The proposed model takes into account random fluctuations introduced by the noise as well as intrinsic heterogeneity of the tissue itself due to other sources of variability (e.g., biological, instrumental). Contextual information is also incorporated in the statistical model owing to the flexibility of MRFs. Moreover, the uncertainty on the labels (the \textit{a posteriori} field assigned to each voxel to reflect the partial volume effect) is modeled by partitioning the continuous proportion degree interval $[0, 1]$ into equal segments. The result is expressed as a fuzzy membership degree reflecting the proportion of any “pure” tissue (that can be viewed as a fuzzy set) at any given voxel. In this context, the fuzzy membership degree can be interpreted as the partial order induced by the discrete membership function over the universe of discourse (all fuzzy sets).

The MRF segmentation has been successfully used in the brain segmentation, particularly when inhomogeneities are present in images [15, 16]. Other alternatives are possible, including field inhomogeneity correction using phase maps (e.g., a locally developed algorithm [17]) that allow the inhomogeneity to be corrected independent of the segmentation process. However, displacements caused by static field inhomogeneities are negligible by an appropriate Shimming procedure [17] particularly at fine resolution (1 mm$^3$). Therefore, they are not considered in our method.

In this paper, we consider a brain to be 3D fuzzy fields which consist of three main tissues (hard classes) and the mixtures of the pure brain tissues (fuzzy classes or mixclasses). Because of the point spread function in MRI (a sinc function), the partial volume effect, and especially the anatomic arrangement of the tissues, the MR intensities are spatially correlated. They can be modeled as fuzzy MRFs, which take into account the contextual information, the statistical information of image signals, and the anatomical information of the brain. As our method is unsupervised, the first step is to optimally estimate all parameters used in the proposed models: statistical parameters of signal distribution of the three hard classes and parameters of the Markovian distribution of the fuzzy field. The fuzzy Markovian segmentation is then carried out using the ICM (iterated conditional modes) algorithm [18] with parameters of the Markovian distribution estimated in the first step. The statistical distribution parameters of the hard classes are refined at each iteration of the regularization. Experiments using simulated and real data have been carried out. A comparison between the fuzzy MRF method and the FCM method shows that our method is more robust.

**FUZZY MRF MODEL**

The classical hard Markov fields are widely used in many applications, such as image restoration [19], motion estimation [20], image segmentation [21], and volumetric object
reconstruction [22]. The fuzzy Markov fields, recently introduced by Pieczynski [23], show their interest in the case of the presence of fuzzy regions in an image, dealing with the satellite image analysis.

The partial volume effect appears when more than one type of class or material occupies one voxel or pixel of an image: these voxels or pixels are usually called mixels. Thus, the fuzzy segmentation of images consists in allowing each pixel to belong simultaneously to numerous classes. The problem is to associate to each pixel \( s \) a vector \( a = (a_1, \ldots, a_k) \in [0, 1]^k \), with \( a_1 + \cdots + a_k = 1 \), where \( a_i \)'s represent the partial volume proportions of each “pure” tissue. We consider two random fields \( A = (A_s)_{s \in \mathcal{S}} \) and \( Y = (Y_s)_{s \in \mathcal{S}} \). The image datum to be segmented is a realization \( Y = y \) of \( Y \) and the desired result is a realization \( A = a \) of the field \( A \). The joint distribution of \( (A, Y) \) is defined by the prior distribution \( P_A(a), \) which is assumed to be stationary and Markovian, and by the posterior distribution \( P_{Y|A}(y|a) \):

\[
P_{A,Y}(a, y) = P_A(a)P_{Y|A}(y|a). \tag{1}
\]

**Distribution of the Observed Field (Y)**

Let us first discuss the observed field \( Y \). Some works have been carried out to describe the partial volume effects in MRI and to find the mixture density functions. The most widely used model represents the intensity \( y \) of a mixel \( s \) as [7, 24]

\[
y = \sum_{j=1}^{k} a_j I_j + \varepsilon, \quad \forall j \in [1, k], \tag{2}
\]

where \( \varepsilon \) is a white gaussian noise with standard deviation \( \sigma_\varepsilon \), \( k \) represents the number of tissues, and \( I_j \) is related to the intensity value of the pure class \( j \) and \( a_j \) to its proportion which is assumed to be a uniform random variable in the interval \([0, 1]\). This model supposes that the same noise \( N(0, \sigma^2) \) is added to each type of pure class and ignores the variability within each pure tissue. The study carried out in [25] shows that the variance of the three brain tissue classes (WM, GM, and CSF) is almost of the same level as that of the noise. Hence, we propose a model for mixel \( y \), which takes into account not only the acquisition noise, but also the heterogeneity of pure matters.

Let us denote by \( y^p_j \) a random variable representing the pure class \( j \). The heterogeneity of both the pure matter and the imaging noise are taken into account in its probability density \( p(y^p_j) \). The intensity value \( y \) of a mixel can then be represented by a weighted sum of \( k \) pure tissues,

\[
y = \sum_{j=1}^{k} a_j y^p_j \quad \text{with} \quad \sum_{j=1}^{k} a_j = 1 \quad \text{and} \quad a_j > 0, \tag{3}
\]

where \( a_j \) is the proportion of the pure class \( j \). In our case, the distribution of each pure class \( p(y^p_j) \) is assumed to be gaussian, confirmed in the studies carried out in [26]. From (3), it is obvious to see that the observation \( y \) is a variable whose fluctuations depend directly on the statistical properties of the random variables \( y^p_j \) (i.e., noise variance and tissue heterogeneity). In the case of two pure tissues \( \Omega = \{c_1, c_2\} \), the intensity value of a mixel
takes the following form:

\[ y = ay_1^p + (1 - a)y_2^p. \]  
\[ (4) \]

In this model, we take \( a \in [0, 1] \), where the values 0 and 1 correspond to the hard classes (0, class 2; 1, class 1), and [0, 1] stands for the fuzzy classes. Assume that the two pure classes are independent of \( s \) and normally distributed. Denoting by \( N(\mu_k, \sigma_k^2) \) the normal distribution of mean \( \mu_k \) and variance \( \sigma_k^2 \) \((k = 1, 2)\), the probability density function of \( y \) is the convolution of \( p(y_1^p) \) and \( p(y_2^p) \) with a weight \( a \). It can be written as follows for a given \( a \):

\[
P_{Y/A}(y/a) = \frac{1}{\sqrt{2\pi}\sqrt{a^2\sigma_1^2 + (1-a)^2\sigma_2^2}} \exp\left\{ -\frac{[y - (a\mu_1 + (1-a)\mu_2)]^2}{2[a^2\sigma_1^2 + (1-a)^2\sigma_2^2]} \right\}. \]
\[ (5) \]

Then, \( p_{Y/A}(y/a) \) is also a gaussian function. The parameters \( \mu_0, \mu_1, \sigma_0^2, \) and \( \sigma_1^2 \) of the pure classes define all distributions of \( Y \) conditional on \( A \).

**Distribution of the Mixture Field (A)**

Let us consider the *apriori* model \( P_A(a) \). For a hard class, we have \( a_j = 1, a_i = 0 \forall i \neq j \). If the field \( A \) is considered Markovian with respect to a neighborhood \( V \), in concordance with the Hammersley–Clifford theorem [19], its distribution is defined as a Gibbs distribution

\[
P_A(a) = \frac{1}{Z} \exp(-U_A(a)), \]
\[ (6) \]

where \( U_A \) stands for the energy function, defined on cliques within \( V \), and \( Z \) is the normalizing constant. In the fuzzy case, the distribution of \( A \) is defined on \( \Omega_j = [0, 1] \). If we consider that \( \frac{1}{2} \exp(-U_j(a)) \) is the density of \( A \) with respect to the mixels, \( U_j \) has the same set of cliques as the function \( U_A \). Then, it is possible to show exactly, as in the hard case, that \( A \) is Markovian with respect to \( V \) [23]. Since \( a_j \) is a real number in the fuzzy case, it can take an infinity of values in \([0, 1]\). However, in a practical case, an infinite value is not realistic. Therefore, we can discretise \( a \) in different ways according to the context. Given a precision, one method is to define a finite number of segments of equal lengths with connecting the hard classes. For example, if a fuzzy step 0.1 is chosen, the number of equal segments between two hard classes is then nine, corresponding to the number of fuzzy classes to be segmented. As a result, \( a \) can take values in \([0.1, 0.2, 0.3, \ldots, 0.9]\). If two hard classes \( k = 2 \) are considered, we have \( A_k = \{a, 1 - a\} \). If \( a = 0.2 \), it means that the voxel is a mix of 20% of class 1 and 80% of class 2. Since the distribution of each pure class is statistical, (4) cannot be solved analytically to obtain \( A \). The classification of a voxel into fuzzy and hard classes is described below.

**SEGMENTATION ALGORITHM**

We assume that a normal human brain consists of three types of tissue: WM, GM, and CSF. However, due to the partial volume effects, a voxel in a MR image can be made up of a single type of tissue or of a mixture of different types. As the probability of mixing more than two tissues in a voxel is very low and the mixture of WM and CSF is very poor in brain data,
we consider that there are three pure classes, CSF, GM, and WM, and two mix classes, CG (mixture of CSF and GM) and GW (mixture of GM and WM). The fuzzy segmentation consists in finding the three pure classes and the fuzzy membership values of the two mix classes (CG and GW). Since the intensities of the two mix classes do not overlap and each mix class is mixed of only two hard classes, the fuzzy models associated to the two hard classes, described in (3), can be used directly for each mix class. We can consider that there are two fuzzy fields to be found which are independent of each other. Therefore, the data energy term concerning Y is expressed as

\[ U_1(y/a) = \ln P_{y/A}(y/a), \]

where \( a \in [0, 1] \). From Eq. (4), \( U_1(y/a) \) is defined by the set of parameters \( \Phi_Y = \{\mu_{CSF}, \sigma_{CSF}^2, \mu_{WM}, \sigma_{WM}^2, \mu_{GM}, \sigma_{GM}^2\} \) which are unknown a priori.

The energy term corresponding to the a priori model is considered in the hard and fuzzy cases. As described in the above section, we have \( P_A(a) = \frac{1}{Z} e^{-U_2(a)} \), where \( U_2(a) = U_h(a) + U_f(a) \). \( U_h(a) \) is related to the hard classes and \( U_f(a) \) to the fuzzy classes. As the object to be segmented is three-dimensional (3D), the system neighborhood \( V \) is the spatial 18-connexity as shown in Fig. 2a. Two kinds of cliques are defined: horizontal neighbors and vertical neighbors. Inferior-posterior neighbors belong to the first kind. The other cliques such as south-west and north-west belong to the second one. These cliques are shown in Fig. 2b.

FIG. 2. (a) 18-connexity spatial neighborhood. (b) Two-site cliques for the second order neighborhood and associated parameters.
$U_h(a)$ is defined when both pixels are hard,

$$
U_h(a) = \left( \sum_{(i,j) \in V_1} \beta_{h_1}^h (1 - \delta(a_i, a_j)) + \sum_{(i,j) \in V_2} \beta_{h_2}^h (1 - \delta(a_i, a_j)) \right),
$$

(7)

where $V_1$ and $V_2$ denote the two kinds of cliques and $\delta(\cdot)$ is the Kronecker delta function.

$U_f(a)$ is defined when one pixel or both pixels are fuzzy:

$$
U_f(a) = \left( \sum_{(i,j) \in V_1} \beta_{fCG_1}^f |a_i - a_j| + \sum_{(i,j) \in V_2} \beta_{fCG_2}^f |a_i - a_j| 
+ \sum_{(i,j) \in V_1} \beta_{fGM_1}^f |a_i - a_j| + \sum_{(i,j) \in V_2} \beta_{fGM_2}^f |a_i - a_j| \right).
$$

(8)

The a priori model is then defined by a set of parameters:

$$
\Phi_a = \{ \beta_{h_1}^h, \beta_{h_2}^h, \beta_{fCG_1}^f, \beta_{fCG_2}^f, \beta_{fGM_1}^f, \beta_{fGM_2}^f \}.
$$

(9)

Using the Bayes rule, the fuzzy segmentation objective is in fact to search for $a$ such that

$$
\hat{a} = \arg \max_a (P_A(a) P_{Y/A}(y/a)).
$$

This maximization is equivalent to:

$$
\hat{a} = \arg \min_a (U_1(y/a) + U_2(a)).
$$

(10)

The deterministic relaxation ICM is used to minimize this global energy function with known parameters $\Phi_y$ and $\Phi_a$. Although the ICM usually converges to a local minimum of the energy function, this loss of optimality can be compensated for by an appropriated initial guess. The initialization can be well obtained by a maximum likelihood method using the appropriate parameters $\Phi_y$. The estimation of the unknown parameters is discussed in the next section.

Parameter Estimation

In the unsupervised Markovian segmentation case, we have to estimate the parameters used in the model (i.e., the parameters $\beta$ in $P_A(a)$ and means and variance of the gaussian distributions in $P_{Y/A}(y/a)$). The complexity of the estimation problem is due to the absence of the observation $A$. Most of the recently proposed estimation methods are iterative, such as the Markov chain Monte Carlo algorithms (MCMC) [27, 28] and the iterative conditional estimation (ICE) algorithm [23, 29]. These methods seem efficient in the blind case. However, these algorithms are very time-consuming. Since the number of classes is known in our case, the initialization of parameters $\Phi_y$ can be obtained by a fit of the image histogram with a mixture of different distributions. Therefore, the method described in [30] is adapted to our parameter estimation.

Let us denote the prior distribution $P(A/\Phi_a)$, which depends on a parameter vector $\Phi_a$, and the intensity distribution $P(Y/A, \Phi_y)$, which depends on the parameter vector $\Phi_y$. We consider the estimation procedures for complete data and incomplete data separately. For the complete data, the estimator of $\Phi_y$ is chosen as the empirical mean and variance. For
the incomplete data, the stochastic gradient method is used to estimate $\Phi_a$. The criteria to estimate the parameters proposed in [30] are expressed as follows:

$$\left(\hat{A}, \hat{\Phi}_a, \hat{\Phi}_y\right) = \arg\max_{\hat{A}, \hat{\Phi}_a, \hat{\Phi}_y} P(A, Y/\hat{\Phi}_a, \hat{\Phi}_y).$$

(11)

It is possible to successively estimate $\hat{A}$ and $(\hat{\Phi}_a, \hat{\Phi}_y)$ by using two suboptimal criteria,

$$\hat{A} = \arg\max_{\hat{A}} P(A, Y/\hat{\Phi}_a, \hat{\Phi}_y)$$

(12)

and

$$(\hat{\Phi}_a, \hat{\Phi}_y) = \arg\max_{\hat{\Phi}_a, \hat{\Phi}_y} P(\hat{A}, Y/\hat{\Phi}_a, \hat{\Phi}_y).$$

(13)

Equation (12) corresponds to the maximum a posteriori for a given estimate $(\hat{\Phi}_a, \hat{\Phi}_y)$. Equation (13) can also be written as $P(\hat{A}/\Phi_a, \Phi_y)P(Y/\hat{A}, \Phi_a, \Phi_y)$. Since $P(Y/\hat{A}, \Phi_a, \Phi_y)$ is attached to the image data, it does not depend on the parameters. Thus, we have $P(Y/\hat{A}, \Phi_a, \Phi_y) = P(Y/\hat{A})$. Finally, Eq. (13) is just the maximum likelihood of $(\Phi_a)$ knowing $\hat{A}$. Using this estimator, the parameter estimation procedure is outlined below:

(a) Initialize the parameters $\Phi^0 = (\Phi_a, \Phi_y)$, iteration $n = 0$.
(b) Segment the image data with the current parameter $\Phi^n$, denoted $\hat{A}$.
(c) Estimate $\hat{\Phi}_a^{n+1}$ from $Y$ using the segmented image data $\hat{A}$.
(d) Generate a Markov field using a Gibbs sampler according to the a priori law $P_A(a) \propto e^{-U_2(a)}$ (see [23]) with the current parameter value $(\Phi^n)$, denoted $\hat{A}^{n+1}$.
(e) Estimate $\hat{\Phi}_a^{n+1}$ using the stochastic gradient algorithm

$$\hat{\Phi}_a^{n+1} = \hat{\Phi}_a^n + \frac{1}{(n + 1)\tau} \left( \frac{\partial U_2}{\partial \hat{\Phi}_a} (\hat{A}^{n+1}, \hat{\Phi}_a^n) - \frac{\partial U_2}{\partial \hat{\Phi}_a} (\hat{A}, \hat{\Phi}_a^n) \right),$$

(14)

where $\tau$ is a constant.

(f) If $\|\hat{\Phi}_a^{n+1} - \hat{\Phi}_a^n\| > \epsilon$, $n = n + 1$, and go back to step (b), where $\epsilon$ is a threshold depending on the required accuracy.

Step (b) segments the data using the algorithm described in the solution of Eq. (10). The parameters used in Eq. (10) are changed during the estimation procedure. The new parameters $\hat{\Phi}_y^{n+1}$ are estimated using the segmented datum $\hat{A}$, which gives the localization of the pure classes. Steps (d) and (e) aim at estimating the parameters $\hat{\Phi}_a^{n+1}$ by maximizing Eq. (13). The method proposed in [31] is used with a slightly different stochastic gradient algorithm (14). Instead of using two Gibbs samplers as in [31], the segmentation result here is used during the iterative procedure. The constant value $\tau(>0)$ adjusts the convergence step to guarantee the convergence: in our case, $\tau = 1$. Normally, at the end of the algorithm, the segmentation is obtained at the same time. However, the Gibbs sampler is very time-consuming. To seed up the algorithm, the estimation algorithm is carried out in a cubic region whose volume is about 1/6 over the entire volume. The segmentation algorithm is then applied to the entire volume after the parameter estimation.
VALIDATION AND RESULTS

The method has been validated in two ways: (a) a simulation study focused on a quantitative assessment in white noise and stationary conditions; (b) its application to real MRI data was qualitatively validated by experts.

Parameter Initialization

The initial parameter values have a significant impact on the quality of the final estimates and the final results, particularly for the initialization of parameters $\Phi \Phi$. In our application, the brain is first extracted from the raw MR images [32]. Then, the segmentation is carried out on the extracted brain containing three pure tissue classes, GM, WM, and CSF, and the corresponding mixtures. As demonstrated in [25], if the three pure classes have normal distributions, their mixtures can be also modeled by gaussian functions under some conditions (relative to the ratio of the standard deviations and the distance between the mean of two hard classes). These conditions are generally verified in high resolution T1 anatomical MRI data. Hence, the histogram of the extracted brain can be considered the sum of five gaussian functions. The problem is then to estimate the parameters of the gaussian functions. We have used the Davidon–Fletcher–Powell method described in [33] to fit the histogram in order to obtain the initial means and standard deviations of the three hard classes. Figure 3 shows the results of the histogram fit on real MRI data.

Simulated Data

The digital brain phantom, available on the Web site BrainWeb [34], is used as a gold standard. This three-dimensional phantom defines the spatial distribution of different tissues (e.g., GM, WM, CSF), where the voxel intensity is proportional to the fraction of tissue within each voxel. The simulated MRI volumes with different noise levels are also available on this site. Each volume datum consists of $181 \times 217 \times 181$ voxels with a cubic resolution $1 \times 1 \times 1 \text{ mm}^3$. The noiseless volume was first segmented by software developed in our laboratory [32] to obtain the brain mask. The tissue classification was carried out within this mask.

FIG. 3. Histogram of real MRI data of the encephalon and estimated mixture probability density. (a) An axial slice of the MRI data; (b) mixture of five gaussian distributions.
The segmentation results are evaluated by measuring the absolute average error $\xi$,

$$\xi = \frac{\sum_{s \in S} |a_s - a'_s|}{\text{Card}(S)},$$

where $a'_s$ denotes the proportion of tissue at the voxel $s$ of the phantom and $a_s$ denotes the same quantity at each voxel of the segmented fuzzy volume. $\text{Card}(S)$ denotes the number of voxels in the reference image $S$. The average error $\xi$ can be directly related to the false positive and the false negative errors, in the sense that a difference of proportions for a given tissue is similar to a difference of partial volumes. For example, a 1-mm$^3$ voxel which has a proportion $a$ of gray matter means that $a$ mm$^3$ of gray matter is contained in this voxel. The maximum value for $|a_s - a'_s|$ is 1, corresponding to a misclassified voxel.

According to the fuzzy degree of the phantom, the fuzzy step is set to 0.2 for the phantom analysis. Consequently, there are 11 classes to segment: 3 hard classes and 8 fuzzy classes (4 classes for each mixture). At first, we propose to study the robustness of the segmentation method with respect to the parameters $\Phi_a$. An optimal fuzzy segmentation result is defined when $\xi$ is minimal. In this optimal case, the parameters are also considered optimal. Therefore, we can find the optimal segmentation by interactively modifying the parameters and eliminating the estimation step. In fact, this procedure means that the optimal parameters are found manually. A comparison between the segmentation result with manually found parameters and the segmentation result with estimated parameters was carried out and provided in Table 1. For this study, the phantom is added with a 5% level gaussian noise. The optimal result is obtained by changing $\Phi_a$ with an incremental step 0.1. It can be observed from Table 1 that the differences between the two results are very small. These results provide support to the performance of our parameter estimation algorithm.

We have also studied the impact of noise on the segmentation results using the phantom corrupted with different noise levels (available at the same BrainWeb site). Figure 4 shows one axial slice of the phantom corrupted by 3, 5, and 7% of the noise level. The different degrees of noise are visible in these three images. The segmentation results are quantified in Table 2 using both the proposed fuzzy MRF method and the FCM method. The error rates of the segmentation were calculated for the entire phantom volume. It can be observed that the proposed method is more robust than the FCM method, since the performance decays less rapidly as the noise level increases. In a practical situation, the noise level corresponds generally to 5%. It can be also observed from Table 2 that more errors are found in the gray matter than in the others; this is because the gray matter is implied in both GM–WM and GM–CSF mixtures.

### Table 1

**Comparison in Percentages of the Errors between the Optimal and the Obtained Segmentations with the Estimated Parameters, Using the Phantom (5% Noise Level)**

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Optimal error $\xi$ (%)</th>
<th>Obtained error $\xi$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>White matter</td>
<td>5.63</td>
<td>5.97</td>
</tr>
<tr>
<td>Gray matter</td>
<td>8.60</td>
<td>9.28</td>
</tr>
<tr>
<td>CSF</td>
<td>2.41</td>
<td>2.56</td>
</tr>
</tbody>
</table>

*Note.* The small difference shows the efficiency of the parameter estimation.
FIG. 4. One axial slice of the phantom corrupted with 3, 5, and 7% noise levels shown from left to right.

Real MRI Data

The MRI data sets which acquired using SPGR sequences (TE = 7 ms, TR = 30 ms, \( \alpha = 40^\circ \)) on a 1.5-T Signa scanner. Each data set consists of 256 \( \times \) 256 \( \times \) 124 voxels with a resolution of 1 \( \times \) 1 \( \times \) 1.2 mm\(^3\). The acquisition orientation is axial.

The algorithm was applied on ten MRI data sets. Consistent and stable results of the segmentation were observed. Axial slices of two data sets are used here to illustrate the performance of the proposed method. The obtained proportions of the three pure tissues are shown in Figs. 5 and 6. The voxel intensities correspond to proportion degrees of each pure tissue. Bright intensity represents high proportions. Here, the fuzzy step is taken to be 0.1. Because the intensities of the MR images change within an interval of less than 200 gray levels, it is not reasonable to take the fuzzy membership value smaller than 0.1. Due to its sensitivity to the noise, as shown in the quantitative study, the FCM method leads to many small spurious regions in the center of the brain, which is very noisy. However, the proposed method provides more compact regions. In fact, voxels located in these regions (subcortical locations) are known to have a high admixture of GM and WM (such as the putamen and thalamus). Figure 5 shows that voxels in these regions or along the cortical gray–white matter boundary have, as predicted, significant proportions of both respective tissue types. It should be noted that the subcortical locations, composed of some internal structures, would be wrongly classified if a hard segmentation was used, while the fuzzy segmentation gives fine details in these regions.

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Fuzzy MRF ( \xi ) (%)</th>
<th>FCM ( \xi ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Noise level 3%</td>
<td>Noise level 5%</td>
</tr>
<tr>
<td>White matter</td>
<td>4.70</td>
<td>5.97</td>
</tr>
<tr>
<td>Gray matter</td>
<td>7.23</td>
<td>9.28</td>
</tr>
<tr>
<td>CSF</td>
<td>2.18</td>
<td>2.56</td>
</tr>
</tbody>
</table>

TABLE 2

Comparison in Percentages of the Segmentation Errors between the Fuzzy MRF Method and the FCM Method Dealing with the Phantom Corrupted with 3, 5, and 7% Noise Levels (see Fig. 4)
FIG. 5. Fuzzy segmentation results of real MRI data using the fuzzy MRF method (left column) and FCM method (right column). The original slice is shown in Fig. 3a. The voxel intensities correspond to proportion degrees of each pure tissue. Bright intensity represents high proportion. From top to bottom, the proportions of the white matter, gray matter, and CSF are shown, respectively.

As the mixture of WM–CSF is ignored in our brain model, some errors occur along the boundaries of the ventricles. However, these misclassified voxels are very rare compared to the entire data set. Since these fuzzy voxels are generally along the boundary of the ventricles, it is not difficult to process them locally in a further step.

Our algorithm was also used to process some data in which tumors are present. From the point of view of image intensity, a tumor can be considered as a mixture of the three tissues, even if it is not physically true. Figure 7 shows the obtained results where the tumor can be
FIG. 6. Fuzzy segmentation results of another real MRI data using the fuzzy MRF method. The original slice is shown in (a). The voxel intensities correspond to proportion degrees of each pure tissue. Bright intensity represents high proportion. From (b) to (d), the proportions of the white matter, gray matter, and CSF are shown, respectively.

FIG. 7. Fuzzy segmentation results of pathological MRI data using the fuzzy MRF method. The original slice is shown in (a). The voxel intensities correspond to proportion degrees of each pure tissue. Bright intensity represents high proportion. From (b) to (d), the proportions of the white matter, gray matter, and CSF are shown, respectively.
viewed as a mixture of GM and CSF. In this way, a recognition step can then be developed to obtain the tumor contours. Some ongoing works in our laboratory are directed toward this topic.

While the obtained results were qualitatively evaluated by experts to validate the algorithm on real data sets, some quantitative indication of performance is still needed. For this purpose, reference segmentations obtained from several experts are crucial to this step. However, many difficulties are encountered such as intra- and interexpert variabilities that are inevitably encountered and that have to be accounted for to reliably assess the performances of our segmentation. Furthermore, it appears a quite difficult task for an expert to label the different voxels, especially when the boundaries between hard classes are fuzzy. Receiver operating characteristics methods would be a very interesting framework to explore [35] to model all these sources of variability (i.e., algorithm, expert observer, random noise). Validation work in this direction is currently under investigation.

DISCUSSION

The segmentation algorithm presented here was designed to be applied to standard clinical MR images, particularly single channel data. The advantage over the majority of segmentation methods requiring multichannel data is that we do not need the registration procedure, which leads to greater volume-averaging effects.

Three applications of the obtained fuzzy segmentation can be considered:

(a) Quantitative volumetric analysis of the tissue with good accuracy. In some clinical applications, the ratios between each tissue volume have to be calculated. It is clear that the segmentation dealing with the partial volume effects represented by the fuzzy membership leads to more accurate results than a hard segmentation for quantitative volumetric analysis.

(b) Analysis of internal brain structures. The obtained fuzzy fields can be used as the first step for segmenting internal brain structures. The fuzzy information of many small regions in the center of the brain (Fig. 5) can help us to assign them to the corresponding structures.

(c) Segmentation of cerebral tumors (Fig. 7) and analysis of the metabolic signal from MR spectroscopy (MRS) combined with the brain tissue information from MRI.

The last two applications are our current work. The first one is focused on the segmentation of the neuro-structures from the obtained fuzzy fields in association with the digital Talairach stereotaxic atlas. Some results have already been reported in [36]. The second one is undergoing with neurologists. The fuzzy regions obtained can be clustered using the fuzzy-connectedness principles.

Compared to the FCM segmentation, the proposed fuzzy MRF segmentation is more robust with respect to the noise. Indeed, the FCM algorithm does not take into account either contextual or statistical information. Furthermore, additional forms of information can be easily incorporated in the MRF model by means of appropriate energy terms, such as the inhomogeneity in the images. The advantage of the FCM algorithm is that it is easier to perform and more efficient in terms of computation time than the fuzzy MRF. However, if we deal with similar types of data, for example, T1-weighted MRI brain data, the parameters can be initialized a priori close to the optimal ones in order to get a quick convergence in the step of parameter estimation, therefore reducing the execution time of the algorithm.
Some assumptions are made in our method. The statistical distributions of the pure tissues are considered gaussian. This assumption is verified in [26]. However, other models can also be used, such as the Rayleigh distribution. The algorithm would be the same; only the expression of $U_1$ would be adjusted accordingly. For the brain model, only the two-tissue mixtures are considered here, because the three-tissue mixtures are very poor in the brain data.

**CONCLUSION**

We have described an unsupervised fuzzy segmentation method, based on a mixture model and a MRF model, which seems well adapted and efficient for MR image segmentation. Quantitative results indicate that the estimated parameters lead to a result very close to the optimal ones. The proposed segmentation method is more robust than the FCM algorithm. When the real data are fuzzy, such as MRI brain data, the use of fuzzy segmentation is always more effective than the use of the hard one. Compared to the hard segmentation method reported in [25], the quantitative errors obtained from the proportion volumes of the three brain tissues provided by the proposed method are smaller under the same conditions, e.g., using the same phantom. Since the fuzzy component of the current phantom is not rich enough, the fuzzy step has been chosen 0.2. Further quantitative validation on the accuracy and stability of the method is still necessary, using realistic phantoms and a large number of clinical scans.

**REFERENCES**


