Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm

Yongyue Zhang*, Michael Brady, and Stephen Smith

***Abstract*—The finite mixture (FM) model is the most commonly used model for statistical segmentation of brain magnetic resonance (MR) images because of its simple mathematical form and the piecewise constant nature of ideal brain MR images. However, being a histogram-based model, the FM has an intrinsic limitation—no spatial information is taken into account. This causes the FM model to work only on well-defined images with low levels of noise; unfortunately, this is often not the the case due to artifacts such as partial volume effect and bias field distortion. Under these conditions, FM model-based methods produce unreliable results. In this paper, we propose a novel hidden Markov random field (HMRF) model, which is a stochastic process generated by a MRF whose state sequence cannot be observed directly but which can be indirectly estimated through observations. Mathematically, it can be shown that the FM model is a degenerate version of the HMRF model. The advantage of the HMRF model derives from the way in which the spatial information is encoded through the mutual influences of neighboring sites. Although MRF modeling has been employed in MR image segmentation by other researchers, most reported methods are limited to using MRF as a general prior in an FM model-based approach. To fit the HMRF model, an EM algorithm is used. We show that by incorporating both the HMRF model and the EM algorithm into a HMRF-EM framework, an accurate and robust segmentation can be achieved. More importantly, the HMRF-EM framework can easily be combined with other techniques. As an example, we show how the bias field correction algorithm of Guillemaud and Brady (1997) can be incorporated into this framework to achieve a three-dimensional fully automated approach for brain MR image segmentation.**

***Index Terms*—Bias field correction, expectation-maximization, hidden Markov random field, MRI, segmentation.**

I. Introduction

Magnetic resonance imagine (MRI) is an advanced medical imaging technique providing rich information about the human soft tissue anatomy. It has several advantages over other imaging techniques enabling it to provide three-dimensional (3-D) data with high contrast between soft tissues. However, the amount of data is far too much for manual analysis/interpretation, and this has been one of the biggest obstacles in the effective use of MRI. For this reason, automatic or semi-automatic techniques of computer-aided image analysis are necessary. Segmentation of MR images into different tissue classes, especially gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF), is an important task.

Brain MR images have a number of features, especially the following: First, they are statistically simple: MR Images are theoretically piecewise constant with a small number of classes. Second, they can have relatively high contrast between different tissues. Unlike many other medical imaging modalities, the contrast in an MR image depends strongly upon the way the image is acquired. By altering radio-frequency (RF) and gradient pulses, and by carefully choosing relaxation timings, it is possible to highlight different components in the object being imaged and produce high-contrast images. These two features facilitate segmentation. On the other hand, ideal imaging conditions are never realized in practice. The piecewise-constant property is degraded considerably by electronic noise, the bias field (intensity inhomogeneities in the RF field) and the partial-volume effect (multiple tissue class occupation within a voxel), all of which cause classes to overlap in the image intensity histogram. Moreover, MR images are not always high-contrast. Many T_2-weighted and proton density images have low contrast between GM and WM. Therefore, it is important to take advantage of useful data while at the same time overcoming potential difficulties.

A wide variety of approaches have been proposed for brain MR image segmentation. Statistical approaches, especially parametric ones, are widely employed [7], [8]. This type of method labels pixels according to probability values, which are determined based on the intensity distribution of the image. With a suitable assumption about the distribution, statistical approaches attempt to solve the problem of estimating the associated class label, given only the intensity for each pixel. Such an estimation problem is necessarily formulated from an established criterion. Maximum *a posteriori* (MAP) or maximum likelihood (ML) principles are two such examples. But before those criteria can be assessed, the formula for the density function of the pixel intensity has to be chosen carefully [9]. *Finite mixture* (FM) models, in particular the *finite Gaussian mixture* (FGM) model when the Gaussian likelihood distribution is assumed [10], [6], is one of the most widely used models in segmentation. FM models have a number of elegant features and are mathematically simple. However,

Manuscript received February 17, 2000; revised September 20, 2000. This paper was supported by the UK MRC and EPSRC. The Associate Editor responsible for coordinating the review of this a paper and recommending its publication was L. P. Clarke. *Asterisk indicates corresponding author*.

Y. Zhang is with the FMRIB Centre, John Radcliffe Hospital, University of Oxford, OX3 9DU Oxford, U.K. (e-mail: yongyue@fmrib.ox.ac.uk). He is also with the Robotics Research Group, Department of Engineering Science, University of Oxford, OX1 3PJ Oxford, U.K.

M. Brady is with the Robotics Research Group, Department of Engineering Science, University of Oxford, OX1 3PJ Oxford, U.K.

S. Smith is with the FMRIB Centre, John Radcliffe Hospital, University of Oxford, OX3 9DU Oxford, U.K.

Publisher Item Identifier S 0278-0062(01)00800-X.

being a histogram-based model, the FM has an intrinsic limitation—spatial information is not taken into account because all the data points are considered to be independent samples drawn from a population. Such a limitation causes the FM model to work only on well-defined images with low levels of noise; unfortunately, this is often not the case with MR images due to artifacts such as the partial volume effect and bias field distortion. Under these conditions, FM model-based methods produce unreliable results.

In order to address this problem, we have developed a hidden Markov random field (HMRF) model, which is a stochastic process generated by a MRF whose state sequence cannot be observed directly but which can be observed through a field of observations. The importance of the HMRF model derives from MRF theory, in which the spatial information in an image is encoded through contextual constraints of neighboring pixels. By imposing such constraints, we expect neighboring pixels to have the same class labels (in the case of piecewise constant images) or similar intensities (in the case of piecewise continuous images). This is achieved through characterizing mutual influences among pixels using conditional MRF distributions.

Any model requires descriptive parameters and a model is only complete when all its parameters are known; therefore, a parameter estimation step is also essential to our HMRF model. In this paper an expectation-maximization (EM) algorithm for solving ML estimation of the model parameters is derived. We show that by incorporating both the HMRF model and the EM algorithm into a mathematically sound HMRF-EM framework, an accurate and robust segmentation approach can be achieved, which is demonstrated through experiments on both simulated images and real data, and comparison made with the FM-EM framework. Being a flexible approach, the HMRF-EM can be easily combined with other techniques to improve the segmentation performance. As an example, we show how the bias field correction algorithm of Guillemaud and Brady [6] is incorporated into it.

Although MRF modeling and its application in image segmentation have been investigated by many other researchers [11]–[13], only in recently years has MRF theory become popular in MR image segmentation. However, most reported methods use MRF only as a general prior in an FM model-based parametric approach to build the MAP estimation. They either lack a proper parameter estimation step to fit the FM model [1], [2] or the parameter estimation procedure they use, such as ML or EM [3]–[5], suffers from the limitation of the FM model mentioned above. In general, although an MRF prior can improve the performance, the FM assumption is still a big limitation.

As to the problem of brain MR image segmentation, we focus on segmenting normal brains without apparent diseases into three tissues: Gray Matter (GM), White Matter (WM) and CSF. The algorithm starts with an initial estimation step to obtain initial tissue parameters and classification. It is then followed by a three-step EM process which updates the class labels, tissue parameters and bias field iteratively. During the iterations, an MRF-MAP approach is used to estimate class labels, MAP is applied to estimate the bias field, and the tissue parameters are estimated by ML. Since we are not interested in the skull or scalp, we have removed these parts of the image from all data, before applying segmentation. The resulting output from the algorithm is segmented tissues, estimated bias field and the restored image (without bias field). The rest of this paper is organized as follows: Section II presents the mathematical details of the HMRF model. Section III gives simulation examples from both the HMRF model and the FM model. Section IV introduces the concept of MRF-MAP classification for estimating the class labels. The complete HMRF-EM framework for classification, as well as parameter estimation through the EM algorithm, is presented in Section V. An additional step for estimating bias field for brain MR images is discussed and incorporated into the framework in Section VII. Comparison with other methods and experimental results are also shown, followed by discussions and future work in the final section.

II. HMRF Model

Let $\mathcal{L}$ and $\mathcal{D}$ be two alphabets:

$$\mathcal{L} = \{1, 2, \ldots, l\}, \quad \mathcal{D} = \{1, 2, \ldots, d\}.$$

Let $\mathcal{S} = \{1, 2, \ldots, N\}$ be the set of indexes and $R = \{r_i, i \in \mathcal{S}\}$ denote any family of random variables indexed by $\mathcal{S}$, in which each random variable R_i takes a value r_i in its state space. Such a family r is called a random field. The joint event $(R_i = r_i, \ldots, R_N = r_N)$ is simplified to $R = r$ where $r = \{r_1, \ldots, r_N\}$ is a *configuration* of R, corresponding to a realization of this random field. Let X and Y be two such random fields whose state spaces are $\mathcal{L}$ and $\mathcal{D}$, respectively, so that for $\forall i \in \mathcal{S}$ we have $X_i \in \mathcal{L}$ and $Y_i \in \mathcal{D}$. Let $\mathbf{x}$ denote a configuration of X and $\mathcal{X}$ be the set of all possible configurations so that

$$\mathcal{X} = \{\mathbf{x} = (x_1, \ldots, x_N) \,|\, x_i \in \mathcal{L}, i \in \mathcal{S}\}.$$

Similarly, let $\mathbf{y}$ be a configuration of Y and $\mathcal{Y}$ be the set of all possible configurations so that

$$\mathcal{Y} = \{\mathbf{y} = (y_1, \ldots, y_N) \,|\, y_i \in \mathcal{D}, i \in \mathcal{S}\}.$$

Given $X_i = \ell, Y_i$ follows a conditional probability distribution

$$p(y_i \,|\, \ell) = f(y_i; \theta_\ell), \quad \forall \ell \in \mathcal{L} \tag{1}$$

where θ_ℓ is the set of parameters. For all ℓ, the function family $f(\cdot; \theta_\ell)$ has the same known analytic form. We also assume that (X, Y) is pairwise independent, meaning

$$P(\mathbf{y}, \mathbf{x}) = \prod_{i \in \mathcal{S}} P(y_i, x_i). \tag{2}$$

In order to develop the HMRF model, we first take the standard FM model as a comparison.

A. Finite Mixture Model

For every $\ell \in \mathcal{L}$ and $i \in S$

$$P(X_i = \ell) = \omega_\ell$$

is independent of the individual sites $i \in \mathcal{S}$ and called a *mixing parameter*. We take ϕ as the model parameter set with

$$\phi = \{\omega_\ell; \theta_\ell \,|\, \ell \in \mathcal{L}\}.$$

Consider two configurations $\mathbf{y} \in \mathcal{Y}$ and $\mathbf{x} \in \mathcal{X}$. From (1) and (2), we can compute the joint probability distribution of $\mathbf{x}$ and $\mathbf{y}$ dependent on the model parameters (ϕ is treated as a set of random variables), namely

$$p(\mathbf{x}, \mathbf{y} \,|\, \phi) = \prod_{i \in \mathcal{S}} p(y_i, x_i \,|\, \phi) = \prod_{i \in \mathcal{S}} \{\omega_{x_i} \cdot f(y_i; \theta_{x_i})\}. \quad (3)$$

We can compute the marginal distribution of $Y_i = y$, dependent on the parameter set ϕ

$$p(y \,|\, \phi) = \sum_{\ell \in \mathcal{L}} p(\ell, y \,|\, \phi) = \sum_{\ell \in \mathcal{L}} \omega_\ell \cdot f(y; \theta_\ell). \quad (4)$$

This is the so-called *finite mixture* (FM) model.

Although this mathematically simple model is widely used [14], it is not considered to be a complete model in practice because it only describes the data statistically—no spatial information about the data is utilized. In other words, the FM model is spatially independent and can, therefore, be specified fully by the histogram of the data. However, images with the same intensity distribution may have totally different structural properties. To overcome this drawback, certain spatial considerations need to be incorporated into the model. Under certain intensity distributions, we want the model to be "adaptive" to structural information or spatially dependent in order to fit the actual image better. This leads to the consideration of MRF theory and our HMRF model.

B. MRF Theory

The spatial property can be modeled through different aspects, amongst which the *contextual constraint* is a general and powerful one. MRF theory provides a convenient and consistent way to model context-dependent entities such as image pixels and correlated features. This is achieved by characterizing mutual influences among such entities using conditional MRF distributions.

In an MRF, the sites in $\mathcal{S}$ are related to one another via a *neighborhood system*, which is defined as $\mathcal{N} = \{\mathcal{N}_i, i \in \mathcal{S}\}$, where $\mathcal{N}_i$ is the set of sites neighboring i, $i \notin \mathcal{N}_i$ and $i \in \mathcal{N}_j \Leftrightarrow j \in \mathcal{N}_i$. A random field X is said to be an MRF on $\mathcal{S}$ with respect to a neighborhood system $\mathcal{N}$ if and only if

$$P(\mathbf{x}) > 0, \quad \forall \mathbf{x} \in \mathcal{X}$$
$$P\left(x_i \,\middle|\, x_{\mathcal{S}-\{i\}}\right) = P\left(x_i \,|\, x_{\mathcal{N}_i}\right).$$

Note, the neighborhood system can be multidimensional. According to the Hammersley–Clifford theorem [15], an MRF can equivalently be characterized by a Gibbs distribution. Thus

$$P(\mathbf{x}) = Z^{-1} \exp(-U(\mathbf{x})) \quad (5)$$

where Z is a normalizing constant called the *partition function*, and $U(\mathbf{x})$ is an *energy function* of the form

$$U(\mathbf{x}) = \sum_{c \in \mathcal{C}} V_c(\mathbf{x}) \quad (6)$$

which is a sum of *clique potentials* $V_c(\mathbf{x})$ over all possible cliques $\mathcal{C}$. A *clique* c is defined as a subset of sites in $\mathcal{S}$ in which every pair of distinct sites are neighbors, except for single-site cliques. The value of $V_c(\mathbf{x})$ depends on the local configuration of clique c. For more detail on MRF and Gibbs distribution see [11].

C. HMRF Model

The concept of a *hidden Markov random field* model is derived from *hidden Markov models* (HMM), which are defined as stochastic processes generated by a Markov chain whose state sequence cannot be observed directly, only through a sequence of observations. Each observation is assumed to be a stochastic function of the state sequence. The underlying Markov chain changes its state according to a $l \times l$ transition probability matrix, where l is the number of states. HMMs have been applied successfully to speech recognition [16], [17] and handwritten script recognition [18].

Since original HMMs were designed as one-dimensional (1-D) Markov chains with first-order neighborhood systems, it cannot directly be used in two-dimensional (2-D)/3-D problems such as image segmentation. Here, we consider a special case of an HMM in which the underlying stochastic process is a MRF instead of a Markov chain, therefore, not restricted to one dimension. We refer to this special case as a *hidden Markov random field*model. Mathematically, an HMRF model is characterized by the following:

- **Hidden Random Field (MRF)**
  The random field $X = \{X_i, i \in \mathcal{S}\}$ is an underlying MRF assuming values in a finite state space $\mathcal{L}$ with probability distribution (5). The state of X is unobservable.
- **Observable Random Field**
  $Y = \{Y_i, i \in \mathcal{S}\}$ is a random field with a finite state space $\mathcal{D}$. Given any particular configuration $\mathbf{x} \in \mathcal{X}$, every Y_i follows a known conditional probability distribution $p(y_i \,|\, x_i)$ of the same functional form $f(y_i; \theta_{x_i})$, where θ_{x_i} are the involved parameters. This distribution is called the *emission probability function* and Y is also referred to as the *emitted random field*.
- **Conditional Independence**
  For any $\mathbf{x} \in \mathcal{X}$, the random variables Y_i are conditional independent

$$P(\mathbf{y} \,|\, \mathbf{x}) = \prod_{i \in \mathcal{S}} P(y_i \,|\, x_i). \quad (7)$$

Based on the above, we can write the joint probability of (X, Y) as

$$P(\mathbf{y}, \mathbf{x}) = P(\mathbf{y} \,|\, \mathbf{x}) P(\mathbf{x}) = P(\mathbf{x}) \prod_{i \in \mathcal{S}} P(y_i \,|\, x_i).$$

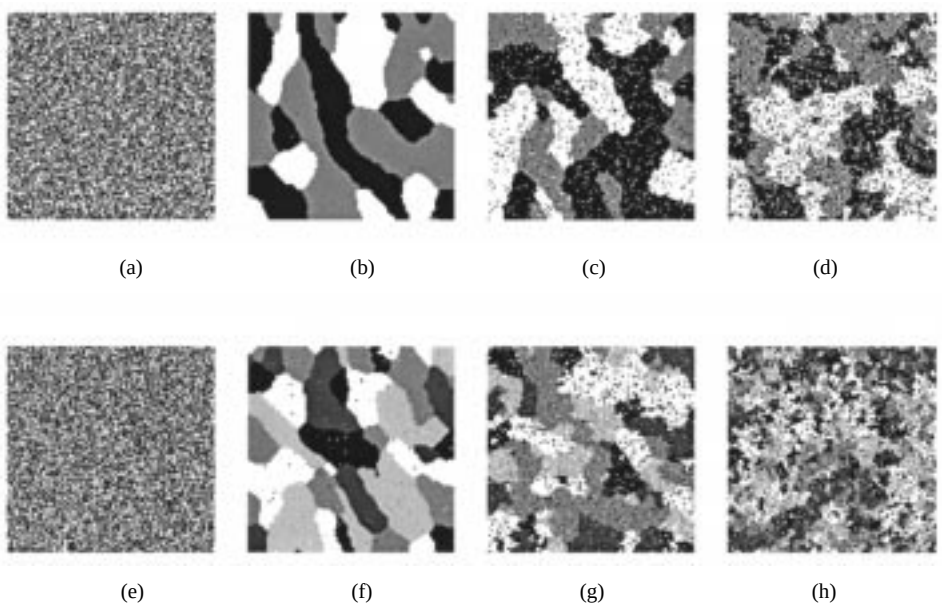


Fig. 1. Image simulation by the FGM model and the GHMRF model. The first row shows the 3-class case. (a) FGM model; (b)–(d) GHMRF model with standard deviation 0.23, 0.4, 0.5, respectively. The second row shows the five-class case. (e) FGM model; (f)–(h) GHMRF model with standard deviation 0.3, 0.47, 0.55, respectively.

According to the local characteristics of MRFs, the joint probability of any pair of (X_i, Y_i), given X_i's neighborhood configuration $X_{\mathcal{N}_i}$, is

$$P(y_i, x_i \mid x_{\mathcal{N}_i}) = P(y_i \mid x_i) P(x_i \mid x_{\mathcal{N}_i}). \tag{8}$$

Thus, we can compute the marginal probability distribution of Y_i dependent on the parameter set θ (in this case, we treat θ as a random variable) and $X_{\mathcal{N}_i}$

$$\begin{aligned} p(y_i \mid x_{\mathcal{N}_i}, \theta) &= \sum_{\ell \in \mathcal{L}} p(y_i, \ell \mid x_{\mathcal{N}_i}, \theta) \\ &= \sum_{\ell \in \mathcal{L}} f(y_i; \theta_\ell) p(\ell \mid x_{\mathcal{N}_i}) \end{aligned} \tag{9}$$

where $\theta = \{\theta_\ell, \ell \in \mathcal{L}\}$. We call this the *hidden Markov random field* model. Note, the concept of an HMRF is different from that of an MRF in the sense that the former is defined with respect to a pair of random variable families (X, Y) while the latter is only defined with respect to X.

If we assume the random variables X_i are independent of each other, which means that for $\forall \ell \in \mathcal{L}$ and $i \in \mathcal{S}$, we have

$$P(\ell \mid x_{\mathcal{N}_i}) = P(\ell) = \omega_\ell$$

then (9) reduces to

$$p(y \mid \theta) = \sum_{\ell \in \mathcal{L}} \omega_\ell \cdot f(y; \theta_\ell)$$

which is the definition of the finite mixture model. Therefore, an FM model is a degenerate special case of an HMRF model.

It is obvious from the above that the fundamental difference between the FM model and the HMRF model lies in their different spatial properties. The FM model is spatially independent whereas the HMRF model may be spatially dependent. Therefore, the HMRF model is more flexible for image modeling in the sense that it has the ability to encode both the statistical and spatial properties of an image.

With a Gaussian emission distribution, the FM model is usually known as the *finite Gaussian Mixture* (FGM) or *finite normal mixture* (FNM) model. More specifically, the observable random variables have the following density function:

$$p(y \mid \phi) = \sum_{\ell \in \mathcal{L}} \omega_\ell \cdot g(y; \theta_\ell) \tag{10}$$

where $\theta_\ell = (\mu_\ell, \sigma_\ell)^T$ and

$$g(y; \theta_\ell) = \frac{1}{\sqrt{2\pi\sigma_\ell^2}} \exp\left(-\frac{(y-\mu_\ell)^2}{2\sigma_\ell^2}\right). \tag{11}$$

Similarly, an HMRF model with a Gaussian emission distribution can be specified as

$$p(y_i \mid x_{\mathcal{N}_i}, \theta) = \sum_{\ell \in \mathcal{L}} g(y_i; \theta_\ell) p(\ell \mid X_{\mathcal{N}_i}) \tag{12}$$

where g and θ_ℓ are defined as in (11). We refer to this type of HMRF model as the *Gaussian hidden Markov random field* (GHMRF) model.

III. Model Simulation and Image Synthesis

Simulation is often used to verify statistical models. In this case, simulation is used to generate synthetic images by drawing random samples from the model distribution using stochastic sampling methods. Here, the Gibbs sampler proposed by Geman and Geman [11] is employed. Many different experiments have been carried out to compare the FGM model and the GHMRF

model. Fig. 1 shows two examples, in which the number of intensity levels was set equal to the number of classes and the Gaussian emission distributions have the same standard deviation for all classes. For the GHMRF model, a homogeneous and isotropic MRF model is employed to generate the prior distribution with clique potential $V_c(\mathbf{x}) = -\delta(x_i - x_j)$. The two rows in Fig. 1 correspond, respectively, to simulations with three and five classes. The first column is the sample drawn from the FGM model while the other three columns show samples drawn from the GHMRF model with different standard deviations. Apparently, the FGM model generates meaningless noise images whereas the GHMRF model generates images with controllable spatial structures—the smaller the standard deviation, the clearer the spatial structures.

IV. MRF-MAP Classification

The image classification problem we consider involves assigning to each pixel a class label taking a value from the set $\mathcal{L}$. The pixels are indexed by a 2-D or 3-D rectangular lattice $\mathcal{S}$ and each pixel is characterized by an intensity value y_i from the set $\mathcal{D}$. A labeling of $\mathcal{S}$ is denoted by $\mathbf{x}$, where $x_i, i \in \mathcal{S}$ is the corresponding class label of pixel i. We write $\mathbf{x}^*$ for the true but unknown labeling configuration and $\hat{\mathbf{x}}$ for an estimate of $\mathbf{x}^*$, both of which are interpreted as particular realizations of a random field X, which is an MRF with a specified distribution $P(\mathbf{x})$. The observable image itself is denoted by $\mathbf{y}$, which is a realization of a GHMRF as described in Section II. The problem of classification is the problem of recovering $\mathbf{x}^*$, given the observed image $\mathbf{y}$.

A. MRF-MAP Estimation

We seek a labeling $\hat{\mathbf{x}}$ of an image, which is an estimate of the true labeling $\mathbf{x}^*$, according to the MAP criterion

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}\in\mathcal{X}}\{P(\mathbf{y}\,|\,\mathbf{x})P(\mathbf{x})\}. \tag{13}$$

From (13), we need to compute the prior probability of the class and the likelihood probability of the observation. Since $\mathbf{x}$ is considered as a realization of an MRF, its prior probability can be derived from

$$P(\mathbf{x}) = \frac{1}{Z}\exp(-U(\mathbf{x})). \tag{14}$$

It is also assumed that the pixel intensity y_i follows a Gaussian distribution with parameters $\theta_i = \{\mu_\ell, \sigma_\ell\}$, given the class label $x_i = \ell$

$$p(y_i\,|\,x_i) = g(y_i;\theta_\ell) = \frac{1}{\sqrt{2\pi\sigma_\ell^2}}\exp\left(-\frac{(y_i-\mu_\ell)^2}{2\sigma_\ell^2}\right). \tag{15}$$

Based on the conditional independence assumption of $\mathbf{y}$ (7), we have the joint likelihood probability

$$\begin{aligned} P(\mathbf{y}\,|\,\mathbf{x}) &= \prod_{i\in\mathcal{S}} p(y_i\,|\,x_i) \\ &= \prod_{i\in\mathcal{S}}\left[\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(y_i-\mu_{x_i})^2}{2\sigma_{x_i}^2} - \log(\sigma_{x_i})\right)\right] \end{aligned}$$

which can be written as

$$P(\mathbf{y}\,|\,\mathbf{x}) = \frac{1}{Z'}\exp(-U(\mathbf{y}\,|\,\mathbf{x})) \tag{16}$$

with the *likelihood energy*

$$\begin{aligned} U(\mathbf{y}\,|\,\mathbf{x}) &= \sum_{i\in\mathcal{S}} U(y_i\,|\,x_i) \\ &= \sum_{i\in\mathcal{S}}\left[\frac{(y_i-\mu_{x_i})^2}{2\sigma_{x_i}^2} + \log(\sigma_{x_i})\right] \end{aligned} \tag{17}$$

and the constant normalization term $Z' = (2\pi)^{(N/2)}$. It is easy to show that

$$\log P(\mathbf{x}\,|\,\mathbf{y}) \propto -U(\mathbf{x}\,|\,\mathbf{y}), \tag{18}$$

where

$$U(\mathbf{x}\,|\,\mathbf{y}) = U(\mathbf{y}\,|\,\mathbf{x}) + U(\mathbf{x}) + \text{const} \tag{19}$$

is the *posterior energy*. The MAP estimation is equivalent to minimizing the posterior energy function

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}}\{U(\mathbf{y}\,|\,\mathbf{x}) + U(\mathbf{x})\}. \tag{20}$$

Although mathematically simple, this type of MAP estimation clearly presents a computationally infeasible problem. Therefore, optimal solutions are usually computed using some iterative optimization (minimization) techniques. In this paper, we adopt the *iterated conditional modes* (ICM) algorithm proposed by Besag [12], which uses the "greedy" strategy in the iterative local minimization and convergence is guaranteed after only a few iterations. Given the data $\mathbf{y}$ and the other labels $x_{\mathcal{S}-\{i\}}^{(k)}$, the algorithm sequentially updates each $x_i^{(k)}$ into $x_i^{(k+1)}$ by minimizing $U(x_i\,|\,\mathbf{y}, x_{\mathcal{S}-\{i\}})$, the conditional posterior probability, with respect to x_i.

V. Model Fitting Using the EM Algorithm

A statistical model is complete only if both its functional form and its parameters are determined. The procedure for estimating the unknown parameters is known as *model fitting*. For an HMRF model, the parameter set $\theta = \{\theta_\ell, \ell \in \mathcal{L}\}$ is what to be solved. If the Gaussian emission function is assumed for the observable random variable y, the mean and standard deviation of each Gaussian class are the parameters, so that $\theta_\ell = (\mu_\ell, \sigma_\ell)$.

Since both the class label and the parameters are unknown and they are strongly interdependent, the data set is said to be "incomplete" and the problem of parameter estimation is regarded as an "incomplete-data" problem. Many techniques have been proposed to solve this problem, among which the EM algorithm [19] is the one most widely used.

The strategy underlying the EM algorithm consists of the following: estimate the missing part as $\hat{\mathbf{x}}$ given the current θ estimate and then use it to form the complete data set $\{\hat{\mathbf{x}}, \mathbf{y}\}$; new θ can be estimated by maximizing the expectation of the complete-data log likelihood, $\mathcal{E}[\log P(\mathbf{x}, \mathbf{y}\,|\,\theta)]$. Mathematically, the EM algorithm can be described by the following.

Start An initial estimate $\theta^{(0)}$.

The E-step Calculate the conditional expectation

$$Q\left(\theta \,\middle|\, \theta^{(t)}\right) = \mathcal{E}\left[\log P(\mathbf{x},\mathbf{y} \,|\, \theta) \,\middle|\, \mathbf{y}, \theta^{(\mathbf{t})}\right] = \sum_{\mathbf{x}\in\mathcal{X}} p\left(\mathbf{x} \,\middle|\, \mathbf{y}, \theta^{(t)}\right) \cdot \log p(\mathbf{x},\mathbf{y} \,|\, \theta). \quad (21)$$

The M-step maximize $Q(\theta \,|\, \theta^{(t)})$ to obtain the next estimate

$$\theta^{(t+1)} = \arg\max_{\theta} Q\left(\theta \,\middle|\, \theta^{(t)}\right). \quad (22)$$

Let $\theta^{(t+1)} \rightarrow \theta^{(t)}$ and repeat from the E-step.

Under certain reasonable conditions, EM estimates converge locally to the ML estimates [20].

For the GHMRF field model, the intensity distribution function, dependent on the parameter set θ, is

$$p(y_i \,|\, \theta) = \sum_{\ell\in\mathcal{L}} \frac{1}{\sqrt{2\pi\sigma_\ell^2}} \exp\left(-\frac{(y-\mu_\ell)^2}{2\sigma_\ell^2}\right) \cdot p\left(\ell \,|\, x_{\mathcal{N}_i}\right) \quad (23)$$

where $p(\ell \,|\, x_{\mathcal{N}_i})$ is the locally dependent probability of $x_i = \ell$ and the parameter set $\theta = \{\mu_\ell, \sigma_\ell \,|\, \ell \in \mathcal{L}\}$.

The Q-function is then formulated as

$$Q = \sum_{i\in\mathcal{S}} \sum_{\ell\in\mathcal{L}} \left\{P^{(t)}(\ell \,|\, y_i) W + C\right\} \quad (24)$$

where

$$W = \log p^{(t)}\left(\ell \,|\, \mathbf{x}_{\mathcal{N}_i}\right) - \log\sigma_\ell - \frac{(y_i-\mu_\ell)^2}{2\sigma_\ell^2},$$

and $C = -0.5\log(2\pi)$.

Applying the EM algorithm, we obtain

$$\mu_\ell^{(t+1)} = \frac{\sum_{i\in\mathcal{S}} P^{(t)}(\ell \,|\, y_i) y_i}{\sum_{i\in\mathcal{S}} P^{(t)}(\ell \,|\, y_i)} \quad (25)$$

$$\left(\sigma_\ell^{(t+1)}\right)^2 = \frac{\sum_{i\in\mathcal{S}} P^{(t)}(\ell \,|\, y_i)(y_i-\mu_\ell)^2}{\sum_{i\in\mathcal{S}} P^{(t)}(\ell \,|\, y_i)} \quad (26)$$

which are the same update equations for the FGM model [9], except that

$$P^{(t)}(\ell \,|\, y_i) = \frac{g^{(t)}(y_i;\theta_\ell) \cdot P^{(t)}\left(\ell \,|\, x_{\mathcal{N}_i}\right)}{p(y_i)}. \quad (27)$$

The calculation of the conditional probability $P^{(t)}(\ell \,|\, x_{\mathcal{N}_i})$ involves estimation of the class labels, which are obtained through MRF-MAP estimation (20). We refer to this HMRF model-based EM algorithm as a HMRF-EM algorithm and the standard FM model-based EM algorithm as a FM-EM algorithm.

VI. Segmentation Using the HMRF-EM Framework

The EM algorithm presented in Section V not only provides an effective method for parameter estimation, but also a complete framework for unsupervised classification using iterative updating.

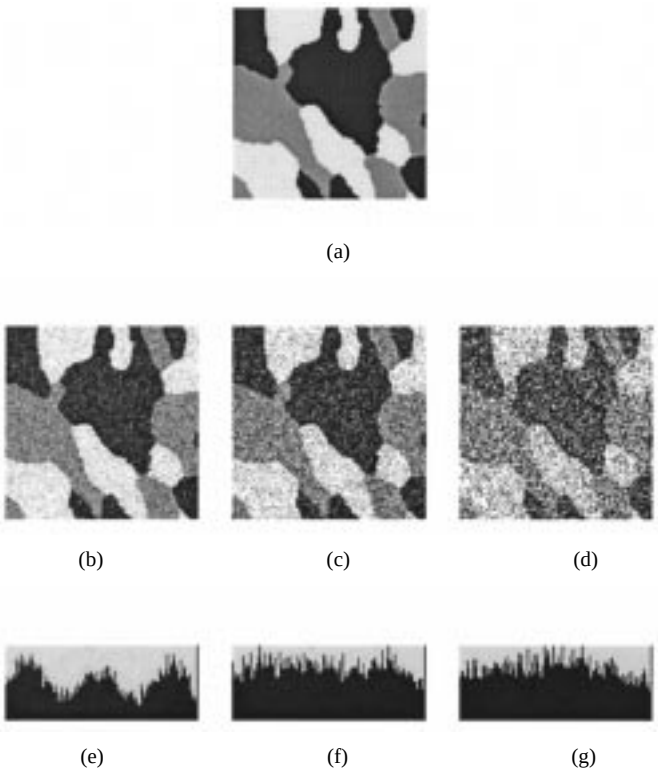


Fig. 2. Test images with three classes. Intensity means are 20, 125, 220, and proportions are 0.372, 0.299, and 0.329, respectively. (a) The original image. (b)–(d) Noisy images with SNR 3.4 ($\sigma = 28$), 2.0 ($\sigma = 47$), and 1.0 ($\sigma = 95$). (e)–(g) Histogram of (b)–(d).

A. Initial Parameter Estimation

Since both the EM model fitting algorithm and the ICM labeling algorithm converge locally, the choice of initial conditions, including the initial parameter set and the classification, is important.

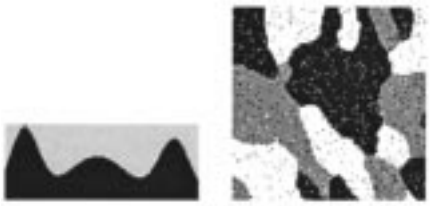
Without prior information, histogram analysis is widely used to estimate statistics such as means and variances of a distribution. From the standpoint of classification, we want the classes to be widely separated from each other while at the same time having relatively low intraclass variances. According to this, we carry out initial estimation using a discriminant measure-based thresholding method proposed by Otsu [21]. The basic idea is to find thresholds maximizing the interclass variances, thus, also minimizing the intraclass variances. According to theories of discriminant analysis, such thresholds are optimal solutions. Once the optimal thresholds have been determined, the mean μ and the standard deviation σ for each class type can then be used as the initial parameters for further estimation. The initial classification can also be obtained either directly through the thresholding, or through an ML estimation with those known parameters.

B. Experiments

We illustrate the performance of HMRF-EM segmentation with reference to a number of examples. First, we show a comparison between the standard FM-EM method and our HMRF-EM method for segmenting and parameter estimating piecewise-constant images with small numbers of classes. Fig. 2(a) shows a simulated three-class image sampled from an

TABLE I
THREE-CLASS PARAMETER ESTIMATION USING THE FM-EM ALGORITHM.

	μ	σ	ω
Class 1	25.6	19.8	0.309
Class 2	122.6	41.5	0.409
Class 3	224.5	21.2	0.282
$K = 47$			



(a) (b)

Fig. 3. Three-class segmentation for Fig. 2(b) using the standard FM-EM algorithm. (a) The reconstructed histogram and (b) the segmentation with MCR 5.82%.

MRF model using the Gibbs sampler. The intensities for the three classes are 30, 125 and 220 respectively. Fig. 2(b)–(d) shows the same images with added Gaussian noise with standard deviation of 28, 47, and 95. Because image contrast is what we are most interested in for examining qualities of an image, a measurement of the noise is more meaningful with image contrast being taken into account. Thus, we define the signal-to-noise ratio (SNR) as the following:

$$\text{SNR} = \frac{\text{mean interclass contrast}}{\text{standard deviation of the noise}}.$$

Thus, the SNRs of the four test images are 3.4, 2.0, and 1.0, respectively. Fig. 2(e)–(g) shows their intensity histograms. With high levels of noise, different classes in the histogram exhibit severe overlap. To measure the segmentation accuracy, we also define the misclassification ratio (MCR), which is

$$\text{MCR} = \frac{\text{number of mis-classified pixels}}{\text{total number of pixels}}.$$

The standard FM-EM algorithm and the HMRF-EM algorithm are then applied to the three test images. Without any limitation to the maximum number of iterations, the standard FM-EM algorithm only converges for the first image with the lowest level of noise (SNR $= 3.4$). In that case, the estimation results and the number of iterations K are shown in Table I. With the estimated parameters, we reconstruct the histogram and obtain the segmentation, as shown in Fig. 3. Note that, the parameter estimation is not accurate compared to the truth.

The HMRF-EM algorithm rapidly converges for all the three test images. Fig. 4 and Table II show the results. Clearly, the results of the HMRF-EM algorithm are much better than those of the standard FM-EM algorithm.

In the second example, we compare both algorithms on a more difficult problem, using a relatively large number of classes and lower image quality. In Fig. 5, we show a simulated 5-class image with three different levels of noise and their histograms. We then apply both methods; the FM-EM method failed for all the three cases whilst the HMRF-EM method gave reasonably good results, shown in Fig. 6 and Table III.

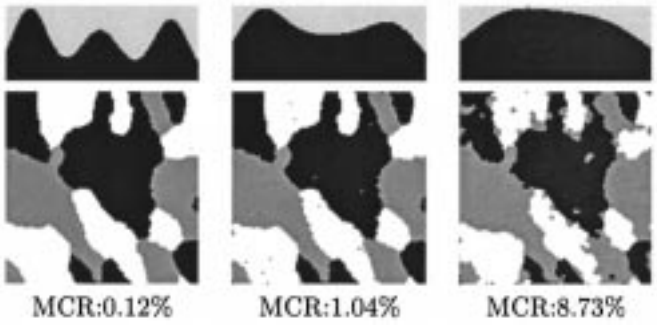


Fig. 4. Three-class segmentation for Fig. 2(b)–(d) using the HMRF-EM algorithm. Top row: the reconstructed histograms; bottom row: the segmentations.

TABLE II
THREE-CLASS PARAMETER ESTIMATION USING THE HMRF-EM ALGORITHM

class	Class 1			Class 2			Class 3			
parameter	μ_1	σ_1	ω_1	μ_2	σ_2	ω_2	μ_3	σ_3	ω_3	K
$SNR = 3.4$	32.0	24.6	0.378	124.9	27.8	0.300	219.2	24.8	0.322	7
$SNR = 2.0$	36.1	35.6	0.377	124.6	46.1	0.304	213.5	38.1	0.320	9
$SNR = 1.0$	52.1	61.7	0.355	127.7	80.5	0.363	203.6	62.1	0.281	31

From the above examples we can conclude that the FM-EM method is sensitive to noise and not robust in terms of convergence. It normally only performs well with limited numbers of classes and takes long to converge. In contrast to this, the HMRF-EM method can reasonably overcome all the drawbacks of the FM-EM method and, therefore, be considered as a superior approach.

VII. SEGMENTATION OF BRAIN MR IMAGES WITH BIAS FIELD CORRECTION

Since it is a complete approach to segmenting piecewise-constant images, the HMRF-EM framework can be applied to brain MR images. However, MR images are often corrupted by a low (spatial) frequency artifact known as the *bias field* arising from inhomogeneities in the RF field. Although such an artifact has little impact on visual diagnosis, the performance of most automatic image analysis techniques, especially intensity-based segmentation, can degrade dramatically. Therefore, a robust, automatic, and inexpensive way of correcting for this artifact is required.

A. Bias Field Correction Through Modified EM Algorithm

One of the most successful methods for dealing with the bias field problem was developed by Wells *et al.* [10], in which the bias field $B = (b_1, \ldots, b_N)$ is modeled as a multiplicative N-dimensional random vector with zero mean Gaussian prior probability density $p(B) = G_{\psi_B}(B)$, where ψ_B is the $N \times N$ covariance matrix. Let $I = (I_1, \ldots, I_N)$ and $I^* = (I_1^*, \ldots, I_N^*)$ be the observed and the ideal intensities of a given image respectively. The degradation effect of the bias field at pixel $i, 1 \leq i \leq N$ can be expressed as follows:

$$I_i = I_i^* \times b_i. \tag{28}$$

After logarithmic transformation of the intensities, the bias field effect can be treated as an additive artifact. Let $\mathbf{y}$ and $\mathbf{y}^*$ denote,

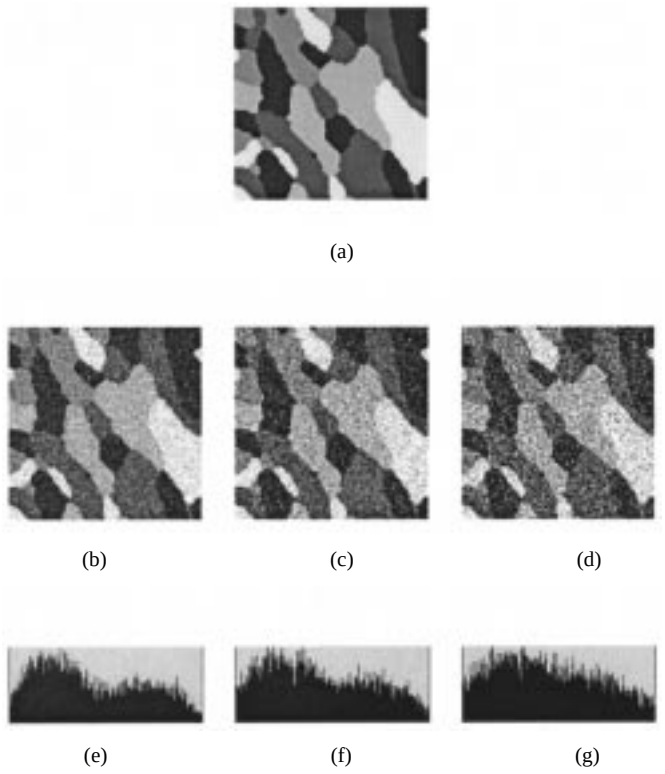


Fig. 5. Test images with five classes. Intensity means are 30, 77, 125, 172, and 220, and proportions are 0.28, 0.273, 0.113, 0.187, and 0.147, respectively. (a) The original image; (b)–(d) noisy images with SNR 2.0 ($\sigma = 23$), 1.4 ($\sigma = 33$), and 1.0 ($\sigma = 47$); (e)–(g) histogram of (b)–(d).

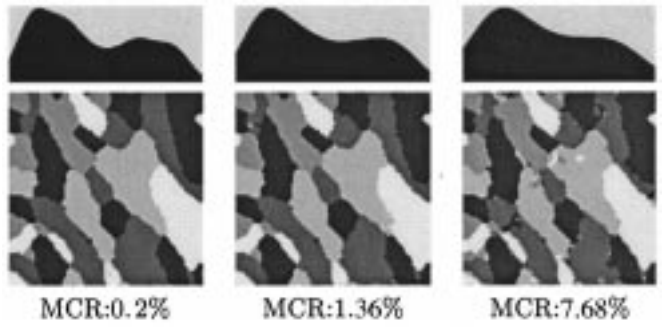


Fig. 6. Five-class segmentation for Fig. 5(b)–(d) using the HMRF-EM algorithm. Top row: the reconstructed histograms; bottom row: the segmentations.

respectively, the observed and the ideal log-transformed intensities: then $\mathbf{y} = \mathbf{y}^* + B$. Given the class labels $\mathbf{x}$, it is further assumed that the ideal intensity value at pixel i follows a Gaussian distribution with parameter $\theta(x_i) = (\mu_{x_i} \sigma_{x_i})$

$$p(y_i^* \mid x_i) = g(y_i^*; \theta(x_i)). \tag{29}$$

With the bias field b_i taken into account, the above distribution can be written in terms of the observed intensity y_i as

$$p(y_i \mid x_i, B) = g(y_i - b_i; \theta(x_i)). \tag{30}$$

Thus, the intensity distribution is modeled as a Gaussian mixture, given the bias field. It follows that

$$p(y_i \mid B) = \sum_{j \in \mathcal{L}} \{g(y_i - b_i; \theta(j)) P(j)\}. \tag{31}$$

The MAP principle is then employed to obtain the optimal estimate of the bias field, given the observed intensity values

$$\hat{B} = \arg\max_B p(\mathbf{y} \mid B) p(B), \tag{32}$$

A zero-gradient condition is then used to assess this maximum, which leads to (see [10] for detail)

$$W_{ij} = \frac{p(y_i \mid x_i, \beta) p(x_i = j)}{p(y_i \mid \beta)} \tag{33}$$

$$b_i = \frac{[FR]_i}{[F\psi^{-1}1]_i}, \quad \text{with } 1 = (1, 1, \ldots, 1)^T \tag{34}$$

where R is the *mean residual* for pixel i

$$R_i = \sum_{j \in \mathcal{L}} \frac{W_{ij}(y_i - \mu_j)}{\sigma_j^2}. \tag{35}$$

ψ is the *mean inverse covariance*

$$\psi_{ik}^{-1} = \begin{cases} \sum_{j \in \mathcal{L}} W_{ij} \sigma_j^{-2}, & \text{if } i = k \\ 0 & \text{otherwise} \end{cases} \tag{36}$$

and F is a low-pass filter. W_{ij} is the posterior probability that pixel i belongs to class j given the bias field estimate.

The EM algorithm is applied to (33) and (34). The E step assumes that the bias field is known and calculates the posterior tissue class probability W_{ij}. In the M step, the bias field B is estimated given the estimated W_{ij} in the E step. Once the bias field is obtained, the original intensity I^* is restored by dividing I by the inverse log of B. Initially, the bias field is assumed to be zero everywhere.

Wells *et al.*'s algorithm is found to be problematic when there are classes in an image that do not follow a Gaussian distribution. The variance of such a class tends to be very large and consequently the mean cannot be considered representative [6]. Such situations are commonly seen in the regions of CSF, pathologies and other nonbrain classes. Bias field estimation can be significantly affected by this type of problem. To overcome this problem, Guillemaud and Brady [6] unify all such classes into an outlier class, which is called "other", with uniform distribution. Let $\mathcal{L}_\mathcal{G}$ denote the set of labels for Gaussian classes and l_o the class label for the "other" class. The intensity distribution of the image is still a finite mixture except for an additional non-Gaussian class

$$p(y_i \mid b_i) = \sum_{j \in \mathcal{L}_\mathcal{G}} \{g(y_i - b_i; \theta(j)) P(j)\} + \lambda P(l_o) \tag{37}$$

where λ is the density of the uniform distribution. Due to the large variance of the uniform distribution, the bias field is only

TABLE III
FIVE-CLASS PARAMETER ESTIMATION USING THE HMRF-EM ALGORITHM.

class	Class 1			Class 2			Class 3			Class 4			Class 5			
parameter	μ_1	σ_1	ω_1	μ_2	σ_2	ω_2	μ_3	σ_3	ω_3	μ_4	σ_4	ω_4	μ_5	σ_5	ω_5	K
$SNR = 2.0$	31	22	0.28	76	23	0.27	124	23.5	0.11	171	24	0.19	219	22	0.14	7
$SNR = 1.4$	33	38	0.29	77	32.5	0.26	123	32.6	0.11	171	33	0.19	217	28.5	0.15	10
$SNR = 1.0$	40	37	0.33	81	44	0.25	125	44	0.08	174	45	0.21	216	37	0.12	9

1. Perform the initial parameter estimation and segmentation.
2. Estimate the bias field

$$b_i^{(t)} = \frac{[FR]_i}{[F\psi^{-1}1]_i}, \text{ with } 1 = (1, 1, \cdots, 1)^T$$

3. Calculate the likelihood distribution

$$p^{(t)}(y_i|x_i, B) = g^{(t)}(y_i - b_i; \theta(x_i))$$

4. Estimate the class labels by MRF-MAP estimation

$$\mathbf{x}^{(t)} = \arg\max_{\mathbf{x}\in\mathcal{X}}\{P(\mathbf{y}|\mathbf{x}, \theta^{(t)}) + P(\mathbf{x})\}$$

5. Calculate the posterior distribution

$$P^{(t)}(\ell|y_i) = \frac{g^{(t)}(y_i; \theta_\ell) \cdot p^{(t)}(\ell|x_{\mathcal{N}_i})}{p(y_i)}$$

6. Update parameters by

$$\mu_\ell^{(t+1)} = \frac{\sum_{i\in\mathcal{S}} P^{(t)}(\ell|y_i) y_i}{\sum_{i\in\mathcal{S}} P^{(t)}(\ell|y_i)}$$

$$(\sigma_\ell^{(t+1)})^2 = \frac{\sum_{i\in\mathcal{S}} P^{(t)}(\ell|y_i)(y_i - \mu_\ell)^2}{\sum_{i\in\mathcal{S}} P^{(t)}(\ell|y_i)}$$

7. $t \leftarrow t+1$ and repeat from 2 until enough iterations have been performed.

Fig. 7. HMRF-EM algorithm for brain MR image segmentation and bias field correction.

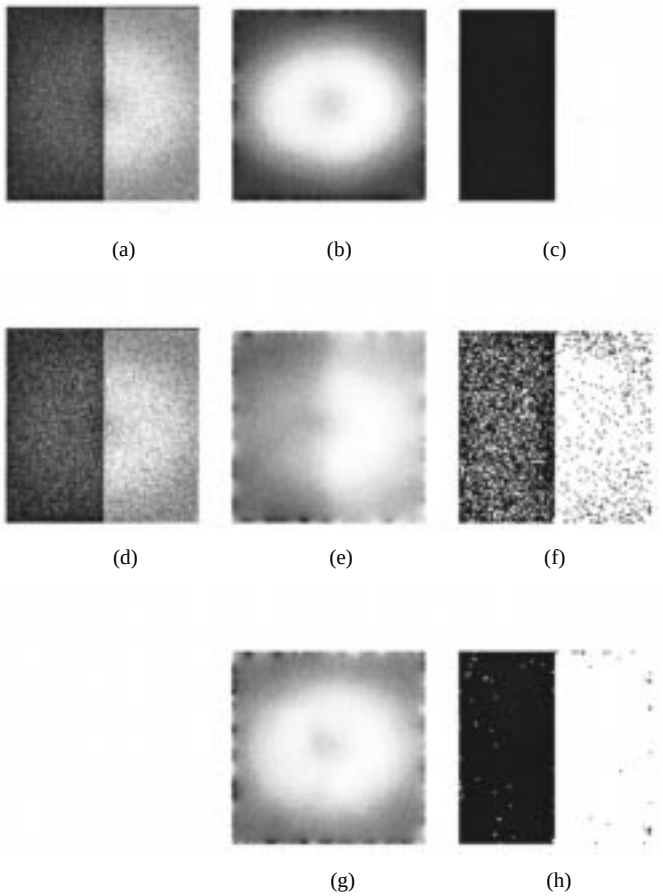


Fig. 8. Comparison of the MEM and the HMRF-EM algorithm on simulated 2-D images. (a) The original image with 3% noise. (b) Bias field estimation for (a) by both the algorithms. (c) Segmentation for (a) by both the algorithms. (d) The original image with 15% noise. (e) Bias field estimation for (d) by the MEM algorithm. (f) Segmentation for (d) by the MEM algorithm. (g) Bias field estimation for (d) by HMRF-EM algorithm. (h) Segmentation for (d) by the HMRF-EM algorithm.

estimated with respect to the Gaussian classes. The same iterative EM method can be applied, except for a slight modification to the formulation of mean residual R_i (35)

$$R_i = \sum_{j\in\mathcal{L}_\mathcal{G}} \frac{W_{ij}(y_i - \mu_j)}{\sigma_j^2}. \tag{38}$$

With such a modification, the performance of the EM algorithm can be significantly improved in certain situations. This approach is referred to as the modified EM (MEM) algorithm.

B. HMRF-EM Framework for Brain MR Image Segmentation

As has been stated in the previous section, FM model-based segmentation methods do not utilize any spatial information and, therefore, are not robust in many cases. The MEM algorithm for brain MR image segmentation suffers from the same problem. But as an effective way to remove bias field, the MEM algorithm is worth improving by overcoming this drawback. We show in this section how the HMRF-EM framework can be easily extended to incorporate an additional bias field correction step. More specifically, we seek an EM solution for three dependent unknowns: the bias field, the image classification and the model parameters. In the E step, we calculate the MAP estimate of the bias field and the class labels to form the Q-function. In the M step, we calculate the ML estimate of the parameters using the estimated bias field and the class labels in the E step.

- **E step**

$$\text{MAP} \quad B^{(t)} = \arg\max_B p\left(B \,\middle|\, \mathbf{y}, \mathbf{x}^{(\mathbf{t-1})}, \theta^{(\mathbf{t})}\right) \tag{39}$$

$$\text{MAP} \quad \mathbf{x}^{(t)} = \arg\min_{\mathbf{x}\in\mathcal{X}} U\left(\mathbf{x} \,\middle|\, \mathbf{y}, B^{(t)}, \theta^{(t)}\right). \tag{40}$$

- **M step**

$$\text{ML} \quad \theta^{(t+1)} = \arg\max_\theta P\left(\mathbf{y} \,\middle|\, \theta, \mathbf{x}^{(\mathbf{t})}, B^{(t)}\right). \tag{41}$$

The complete algorithm is described in Fig. 7.

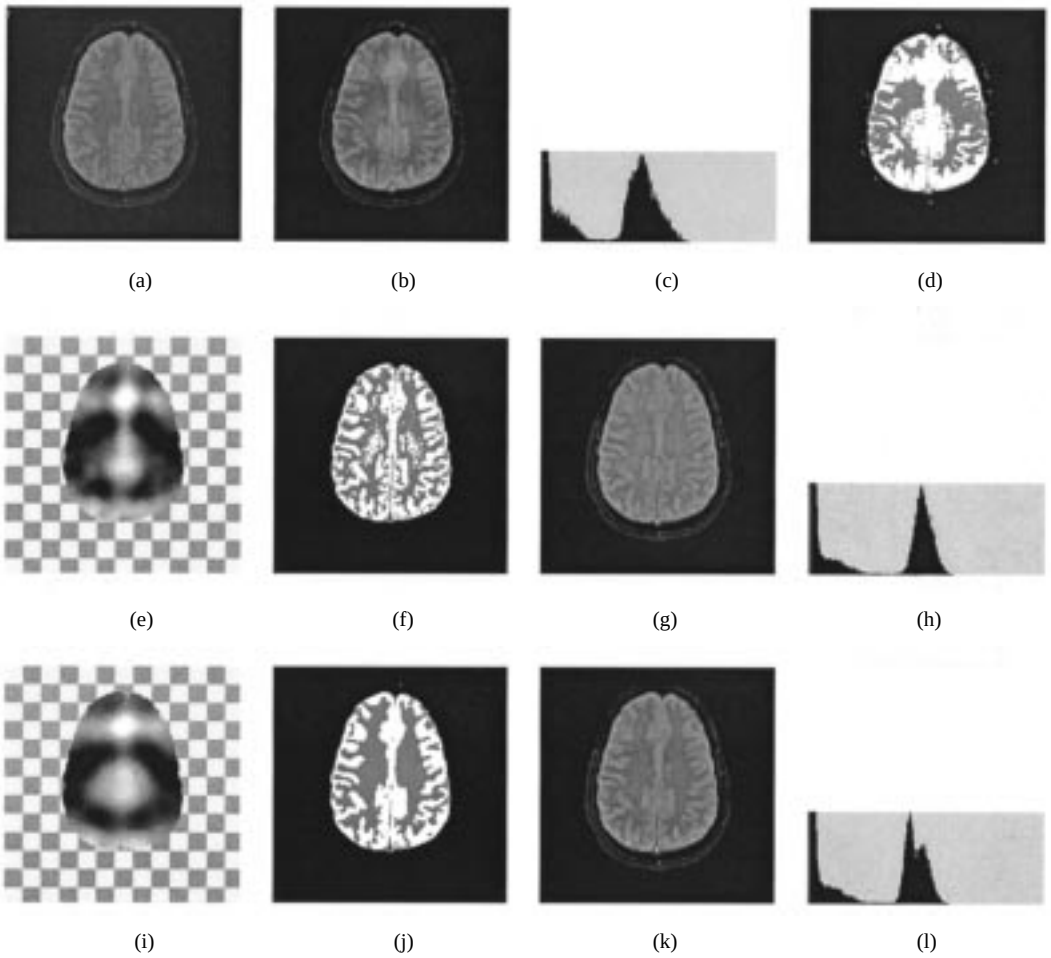


Fig. 9. Comparison of the MEM and the HMRF-EM algorithm on real 2-D MR images with simulated bias field. (a) The original image; (b) the image with simulated bias field; (c) histogram of (b); (d) best thresholding on (b); (e)–(h) the results from the MEM algorithm; (i)–(l) the results from HMRF-EM algorithm. For the last two rows, from left to right: the estimated bias field (the checkerboard is used to represent the background which is assumed to have no bias field), the segmentation, the restored image and the histogram of the restored image.

VIII. Experiments

Various experiments have been carried out on real and simulated data, in both two and three dimensions. For the MEM algorithm, parameters are manually estimated since it does not deal with parameter estimation itself. For the HMRF-EM algorithm, parameters are estimated automatically.

The first experiment shown here tests the noise sensitivity of the two algorithms. Two images consisting of two constant regions with the same simulated bias field but with different levels of Gaussian noise were generated [Fig. 8(a), (d)]. Two Gaussian classes, corresponding to the two regions, are used. For Fig. 8(a), both algorithms give perfect estimates, as shown in Fig. 8(b) and (c). However, for Fig. 8(d), the estimate from the MEM algorithm gives a poor result.

The second experiment tests the performance of the two algorithms on real 2-D MR images but with a simulated bias field. Fig. 9(a) shows one slice of a proton density image and Fig. 9(b) shows the image with a simulated circular bias field. Fig. 9(c) is the histogram of Fig. 9(b), from which a substantial intensity overlap between WM and GM can be seen. Fig. 9(d) shows the best result that can be obtained from Fig. 9(b) using global thresholding. When applying both algorithms two Gaussian distributions are used for the two tissue classes (WM and GM) and a uniform distribution (density$=0.3$) is used for the rest. The second row of Fig. 9 shows the result from the MEM algorithm and the last row shows that from the HMRF-EM algorithm.

Although the above two experiments are only simulations they can still give us some ideas about the two algorithms. Along with those experiments on general images presented in Section V, we can conclude in general that the HMRF-EM algorithm has the advantage over the FM-EM algorithm when the image noise is not negligible; this is always the case in brain MRI. More importantly, as the bias field estimation and the segmentation are mutually influenced by each other, the performance of the segmentation algorithm will have direct influence on the bias field estimation, as well as the overall performance. Thus, we can also conclude that the HMRF-EM segmentation/bias field correction framework is superior to the FM-EM framework.

In the following, we show several experiments applying our HMRF-EM algorithm on real 3-D images taken by different scanners. The first one is on a T1 scan obtained from the Montreal Neurological Institute, McGill University (courtesy of D. Arnold). The original volume has 50 256×256 slices with voxel size $0.977 \times 0.977 \times 3.0$ mm. Fig. 10 shows the segmentation and bias field estimation results of four different slices.

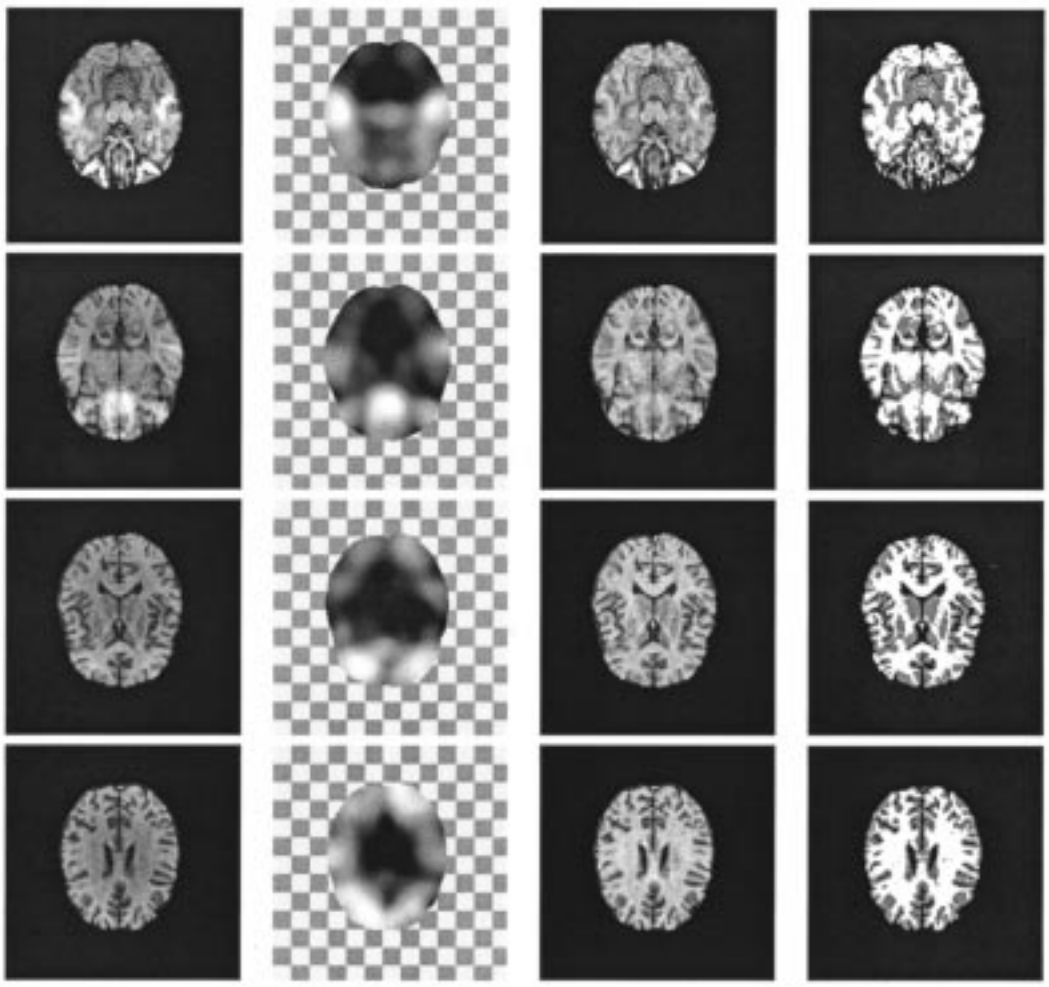


Fig. 10. Four slices of a 3-D MR volume image with real bias field. In each row, from left to right: the original slice, the estimated bias field, the restored slice, and the segmentation.

We also obtained several coronal scans along with their manually guided segmentations from the Internet Brain Segmentation Repository (IBSR) of the Masachusetts General Hospital. Two examples with their original images, bias field estimations, restorations, and both the manually guided segmentation and our segmentations are shown in three views (coronal, transverse, and sagittal) in Figs. 11 and 12.

We can see the above two data sets have very different bias field patterns and image qualities. In both cases, our algorithm performs well visually. Although a quantitative comparison with the provided manually guided segmentation could be carried out, this was not done due to the poor quality of the manually guided segmentation. The transverse and the sagittal views of the image show various small structures are missing from the manual segmentation; also various CSF and nonbrain regions are misclassified as gray or WM.

IX. Discussions and Conclusion

A practical issue has to be addressed in the 3-D implementation of the HMRF-EM algorithm. Theoretically the MRF neighborhood system should be three-dimensionally isotropic. However, the slice thickness of a 3-D volume is often larger than the intraslice voxel dimensions. In such a situation, an isotropic neighborhood system may cause problems. Therefore, an anisotropic 3-D neighborhood system is used with a smaller weight across slices.

Although the HMRF-EM framework itself is theoretically sound, the initial estimation based on thresholding is rather heuristic. Due to the high variability of brain MR images in terms of their intensity ranges and contrasts between brain tissues, it is not guaranteed that the thresholding procedure will produce perfect results. In most cases, however, the final segmentation results are stable even with slightly different initial estimates. This is largely attributable to the robust HMRF-EM algorithm. However, as a local minimization method, the EM algorithm can be trapped in a local minimum. In some cases, where the image is poorly defined, the thresholding procedure may fail to find the right thresholds for brain tissues, especially the threshold for GM and WM. With an initial condition far from normal, the EM procedure is likely to give a wrong final segmentation. In general, the initial estimation is a difficult problem and it will certainly be an important issue to be studied in future work.

We found many scanners produce images with strongly varying bias field across slices. Thus, the very top or bottom

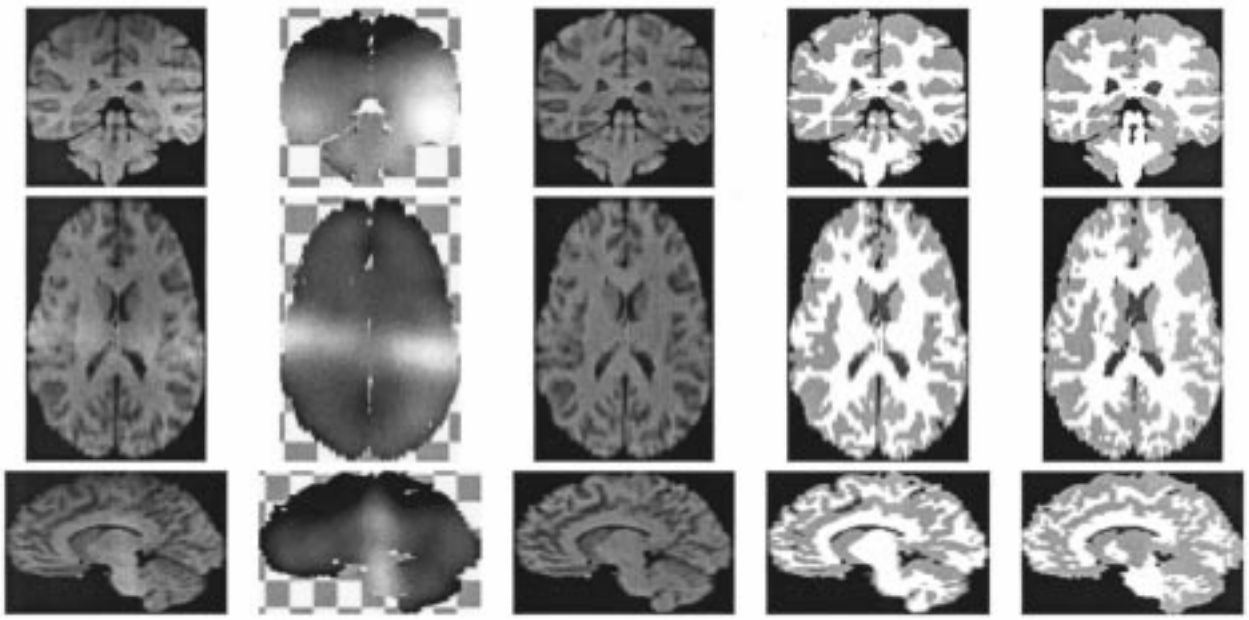


Fig. 11. Data set I from the IBSR. The three rows show the image in coronal, transverse, and sagittal view respectively. In each row, from left to right: the original slice, the estimated bias field, the restored slice, our segmentation, and the manually guided segmentation.

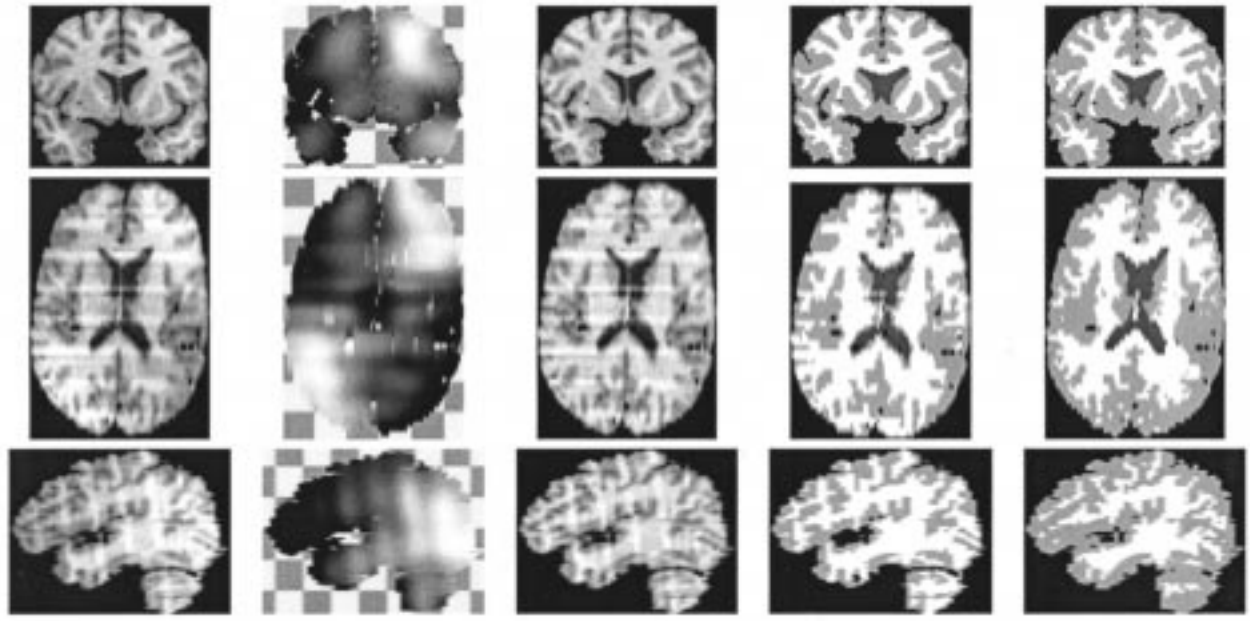


Fig. 12. Data set II from the IBSR. The three rows show the image in coronal, transverse, and sagittal view respectively. In each row, from left to right: the original slice, the estimated bias field, the restored slice, our segmentation, and the manually guided segmentation.

slices can have very different intensities to those in the middle. This can cause the initial parameter estimation to be totally wrong as it is done in 3-D and all the pixels in the volume are considered as a whole. Attempting to carry out the initial estimation two-dimensionally in a slice-by-slice manner is still far from ideal as there are not enough pixels to compute statistics in top or bottom slices. It would appear that to effectively solve this problem certain intensity normalization across slices has to be performed so that for different slices, the mean intensities of a given tissue is similar. However, without knowing the bias field pattern *a-priori*, this normalization is a very difficult problem.

With respect to the computational load, the whole algorithm is slightly slower than the original FM model-based MEM algorithm due to the additional MRF-MAP classification and the EM fitting procedure. However, by employing the fast deterministic ICM method and certain optimizations to the program, it runs reasonably quickly. Currently, it takes 10 seconds for a 256×256 2-D image and less than 10 minutes for a 3-D image with 40 256×256 slices in an Intel PII 400 MHZ-based system.

To conclude, we have presented a fully automatic approach for segmenting brain MR images. The method is based on an HMRF-EM framework, which is a combination of the HMRF model and the associated MRF-MAP estimation and the EM fitting procedures. The HMRF model is proposed in this paper as a substitute for the widely used FM model, which is sensitive to noise and, therefore, not robust. As a very general method, the HMRF-EM framework could be applied to many different image segmentation problems.

We also show that the framework can easily be extended by incorporating other techniques in order to improve its performance on certain problems. As an example, we demonstrated how the bias field correction algorithm by Guillemaud and Brady [6] can be incorporated into this framework. As a result, a 3-D fully automatic approach for brain MR image segmentation is achieved and significant improvements have been observed in terms of both the bias field estimation and the tissue classification.

ACKNOWLEDGMENT

The authors would like to thank A. Blake, P. Matthews, M. Jenkinson, C. Beckmann, M. Woolrich, and P. Bannister for helpful discussions. They would also like to thank the anonymous reviewers for their constructive comments.

REFERENCES

[1] K. Held, E. R. Kops, B. J. Krause, W. M. Wells, and R. Kikinis, "Markov random field segmentation of brain MR images," *IEEE Trans. Med. Imag.*, vol. 16, pp. 878–886, Dec. 1997.

[2] T. Kapur, E. L. Grimson, R. Kikinis, and W. M. Wells, "Enhanced spatial priors for segmentation of magnetic resonance imagery," *Lect. Notes Comput. Sci.*, vol. 1496, pp. 148–157, 1998.

[3] J. Zhang, "The mean field theory in EM procedures for blind Markov random field image restoration," *IEEE Trans. Image Processing*, vol. 2, pp. 27–40, Jan. 1993.

[4] J. C. Rajapakse and F. Kruggel, "Segmentation of MR images with intensity inhomogeneities," *Image Vis. Computing*, vol. 16, no. 3, pp. 165–180, 1998.

[5] K. V. Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Suetens, "Automated segmentation of MS lesions from multi-channel MR images," *Lect. Notes Comput. Sci.*, vol. 1679, pp. 11–21, 1999.

[6] R. Guillemaud and J. M. Brady, "Estimating the bias field of MR images," *IEEE Trans. Med. Imag.*, vol. 16, pp. 238–251, June 1997.

[7] M. E. Brummer, "Optimized intensity thresholds for volumetric analysis of magnetic resonance imaging data," *Proc. SPIE*, vol. 1808, pp. 299–310, 1992.

[8] A. Kundu, "Local segmentation of biomedical images," *Comput. Med. Imag. Graph.*, vol. 14, pp. 173–183, 1990.

[9] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Univ., 1995.

[10] W. M. Wells, E. L. Grimson, R. Kikinis, and F. A. Jolesz, "Adaptive segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 15, pp. 429–442, Aug. 1996.

[11] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, no. PAMI–6, pp. 721–741, June 1984.

[12] J. Besag, "On the statistical analysis of dirty pictures (with discussion)," *J. of Royal Statist. Soc.*, ser. B, vol. 48, no. 3, pp. 259–302, 1986.

[13] S. Z. Li, *Markov Random Field Modeling in Computer Vision*. Berlin, Germany: Springer-Verlag, 1995.

[14] T. Lei and W. Sewchand, "Statistical approach to X-ray CT imaging and its applications in image analysis—Part II: A new stochastic model-based image segmentation technique for X-ray CT image," *IEEE Trans. Med. Imag.*, vol. 11, no. 1, pp. 62–69, 1992.

[15] J. Besag, "Spatial interaction and the statistical analysis of lattice systems (with discussion)," *J. Roy. Stat. Soc.*, ser. B, vol. 36, no. 2, pp. 192–326, 1974.

[16] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, pp. 257–286, Feb. 1989.

[17] J. A. Vlontzos and S. Y. Kung, "Hidden Markov models for character recognition," *IEEE Trans. Image Processing*, vol. 1, pp. 539–543, Apr. 1992.

[18] A. Kundu, Y. He, and P. Bahl, "Recognition of handwritten word: First and second-order hidden Markov model based approach," *Pattern Recogn.*, vol. 22, no. 3, pp. 283–297, 1989.

[19] A. P. Dempster, N. M. Laird, and D. B. Bubin, "Maximum likelihood from incomplete data via EM algorithm," *J. Roy. Stat. Soc.*, ser. B, vol. 39, no. 1, pp. 1–38, 1977.

[20] C. F. J. Wu, "On the convergence properties of the EM algorithm," *Ann. Stat.*, vol. 11, pp. 95–103, 1983.

[21] N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Trans. Syst. Man Cybern.*, vol. 9, pp. 62–66, Jan. 1979.