Segmentation of Brain MR Images with Bias Field Correction

Seung-Gu Kim
Sangji University
Department of Applied Statistics
Wonju, Gangwon, 220-702, Korea
sgukim@mail.sangji.ac.kr

Shu-Kay Ng
University of Queensland
Department of Mathematics
Brisbane QLD 4072, Australia
skn@maths.uq.edu.au

Geoffrey J. McLachlan
University of Queensland
Department of Mathematics
Brisbane QLD 4072, Australia
gjm@maths.uq.edu.au

Deming Wang
University of Queensland
Centre for Magnetic Resonance
Brisbane QLD 4072, Australia
deming.wang@cmr.uq.edu.au

Abstract

We consider a statistical model-based approach to the segmentation of magnetic resonance (MR) images with bias field correction. The proposed method of penalized maximum likelihood is implemented via the expectation-conditional maximization (ECM) algorithm, using an approximation to the E-step based on a fractional weight version of the iterated conditional modes (ICM) algorithm. A Markov random field (MRF) is adopted to model the spatial dependence between neighouring voxels. The approach is illustrated using some simulated and real MR data.

1. Introduction

Medical magnetic resonance imaging (MRI) has the advantages of being able to penetrate bony and air-filled structures with negligible attenuation and artifact. The modality has proven to be a very useful noninvasive medical imaging technique because of the ability to render high anatomical resolution of soft tissues with imaging in any arbitrary plane. Tissue-segmentation of magnetic resonance (MR) images of the human brain has a large potential to facilitate an imaging-based medical diagnosis, providing an aid to surgery and treatment planning [10]. Accurate estimation of the tissue parameters, including their volume sizes, will help to monitor changes in brain haemodynamics and metabolism resulting from neuronal activity [4], and so will assist in the diagnosis and treatment of neurogenerative disease such as Alzheimer's disease. MRI is also useful in providing anatomical information about the location of potential discontinuities in the Positron Emission Tomography (PET) image [17] and an opportunity to monitor the human brain activation effects to stimuli at relatively high spatial resolution [15]. Such tissue segmentation of MR images is often achieved by applying statistical classification techniques to the signal intensities [5, 21, 29], in conjunction with post-processing operations to remove acquisition artifacts [7, 12, 18]. A comprehensive review on MR image segmentation methods is provided by [6].

We consider here a statistical-based approach whereby the intensities on each voxel is modelled by a mixture of a finite number, say g, of normal distribution [19, 21]. In the latter, the expectation-maximization (EM) algorithm [8] is adopted to segment MR images and estimate the tissue parameters. An approximation to the E-step of the EM algorithm is employed based on a fractional weight version of Besag's iterated conditional modes (ICM) algorithm [2]. The prior (spatial) distribution of different tissue types is modelled by a hidden Markov Random Field (MRF) so as to incorporate spatial continuity constraints on the tissue segmentation. We refer to this model as GMM-HMRF (Gaussian Mixture Model with Hidden Markov Random Field). However, the intensity inhomogeneity of MR images due to acquisition equipments, severely degrades intensity-based segmentation of MR images [14, 30]. This low (spatial) frequency artifact known as the *bias field* arises from inhomogeneities in the radio-frequency (RF) field. Let $\boldsymbol{Y} = (y_1, \ldots, y_n)^T$ and $\boldsymbol{Y}^* = (y_1^*, \ldots, y_n^*)^T$ be the observed and the ideal log-transformed intensities of a given image of n voxels, respectively, where the superscript T denotes vector transpose. The degradation effect of the bias field at

the jth voxel can be expressed by an additive model as

$$y_j = y_j^* + b_j \qquad (j = 1, \ldots, n), \tag{1}$$

where b_j is the bias field at the jth voxel. It is noted that (1) implies that the observed MRI signal intensity is modelled as a product of the ideal intensity and a spatially varying factor (exponential of b_j).

In this paper, we extend the GMM-HMRF model by allowing the segmentation of MR images with bias field correction. Based on a penalized likelihood approach, we show how the estimation of the bias field and tissue parameters, and segmentation of the MR images can be obtained simultaneously via a partial version of an expectation-conditional maximization (ECM) algorithm [24].

2. Segmentation of MR Images for Gaussian Mixture Model with HMRF

Suppose that a continuous MR image is partitioned into a set of disjoint voxels labelled 1 to n, and that each voxel is assumed to belong to one of g distinct tissue types. This assumption is tenable because MR images have a spatial resolution at the range of the voxel size [19]. For notational convenience, we consider univariate intensity of each voxel, where the observed log-transformed intensities are denoted by a one-dimensional (1D) array $\boldsymbol{Y} = (y_1, \ldots, y_n)^T$. Also, we let the g groups $G_1, \ldots, G_g$ represent the g possible tissue types. Further, we let $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$ denote the unobservable group-indicator vectors, where the ith element z_{ij} of $\boldsymbol{z}_j$ is taken to be one or zero according as to whether the jth voxel does or does not belong to the ith group. We put $\boldsymbol{z} = (\boldsymbol{z}_1^T, \ldots, \boldsymbol{z}_n^T)^T$.

A parametric mixture model approach [22] is adopted to represent the marginal distribution of $\boldsymbol{Y}$ of a given image of n voxel, where the bias field $b_1, \ldots, b_n$ are considered as unknown parameters. We assume that, unconditionally with respect to the group of origin, y_j $(j = 1, \ldots, n)$ has the finite mixture form of

$$p(y_j|b_j, \boldsymbol{\theta}, \beta) = \sum_{i=1}^{g} p(z_{ij} = 1|\beta) p(y_j|z_{ij} = 1, b_j, \boldsymbol{\theta}) \tag{2}$$

where β is the parameter in the prior probability function of $\boldsymbol{Z}$ and $\boldsymbol{\theta}$ is the vector containing the tissue parameters. Suppose that the ideal log-transformed intensity of a voxel belonging to the ith group is normally distributed around a certain mean μ_i, with a variance σ_i^2. Then, we have

$$p(y_j|z_{ij} = 1, b_j, \boldsymbol{\theta}) = \phi_{\sigma_i}(y_j - \mu_i - b_j), \tag{3}$$

where $\phi_{\sigma_i}(\cdot)$ denotes a zero-mean normal distribution with variance σ_i^2 [28, 30] and $\boldsymbol{\theta} = (\mu_1, \sigma_1^2, \ldots, \mu_g, \sigma_g^2)^T$ is the vector containing the tissue parameters.

For the segmentation of MR images, the problem of inferring the vector $\boldsymbol{z}$ can be viewed as an incomplete-data problem, which can be approached by the application of the EM algorithm. This line of approach was undertaken by Kay and Titterington [16], who related some of the relaxation algorithms for image analysis with methods in the literature on the statistical analysis of incomplete data. The complete-data vector is given by $(\boldsymbol{y}^T, \boldsymbol{z}^T)^T$.

We consider a penalized complete-data log likelihood as

$$\begin{aligned} \log L_{Pc}(\boldsymbol{b}, \boldsymbol{\theta}, \beta) &= \log L_c(\boldsymbol{b}, \boldsymbol{\theta}, \beta) - \tfrac{1}{2}\boldsymbol{b}^T \Sigma_b^{-1} \boldsymbol{b} \\ &= \log p(\boldsymbol{y}|\boldsymbol{z}, \boldsymbol{b}, \boldsymbol{\theta}) + \log p(\boldsymbol{z}|\beta) \\ &\quad - \tfrac{1}{2}\boldsymbol{b}^T \Sigma_b^{-1} \boldsymbol{b}, \end{aligned} \tag{4}$$

where $\boldsymbol{b} = (b_1, \ldots, b_n)^T$, $\log L_c$ is the complete-data log likelihood, and $\Sigma_b = LL^T$, where L is a pre-defined low-pass filter [30]. In (4), The latter term can be viewed as a penalty term to stabilise the ML solution and to promote piecewise smoothness in the resulting segmentation and the bias field estimation. Thus, (4) can be regarded as a regularization method based on a penalized likelihood approach [13, 25]. It will be seen in Section 4 that this penalty term improves the segmentation and the bias field estimation.

Markov random fields are commonly employed in image processing problems to reflect the extent to which spatially neighbouring voxels belong to the same group [2, 11]. The incorporation of such spatial information on the images plays an important role in the estimation of $\boldsymbol{z}$ [5, 19]. The Hammersley-Clifford theorem states that the MRF prior can be specified using a Gibbs distribution [2, 11]

$$p(\boldsymbol{z}|\beta) = \exp\{-U(\boldsymbol{z}|\beta)\}/C(\beta), \tag{5}$$

where $C(\beta)$ is a normalizing constant known as the partition function and $U(\boldsymbol{z}|\beta)$ is the energy function specified by the neighbourhood system for the image. Because of the existence of the term $C(\beta)$ on the right-hand side of (5), there will be a stumbling block with the M-step with respect to β. Although the parameter β of a fairly general MRF can be estimated using Besag's pseudo-likelihood method [2], good estimates of β do not necessarily result in good segmentation [1]. We assume henceforth that β is specified *a priori*; see also the discussion in [3]. Besag [2] settled on $\beta = 1.5$ empirically. For the segmentation of MR images, the usage of such large value of β might, however, fail to detect small patches of voxels of one group surrounded by voxels of another group.

It will be seen that the E-step requires the calculation of $\tau_{ij}^{(k)} = E\{Z_{ij} \mid \boldsymbol{y}, \boldsymbol{\Psi}^{(k)}\}$, which unfortunately cannot be computed exactly under a HMRF mixture model [27]. McLachlan et al. [21] considered an approximation to the E-step based on a fractional weight version of the ICM algorithm. In the next section, we extend their GMM-HMRF

model to simultaneously estimate the bias field and the tissue parameters, via an ECM algorithm.

3. An ECM Algorithm for Penalized ML Estimation

Concerning the probability density function of $\boldsymbol{Y}$ given $\boldsymbol{z}$ and $\boldsymbol{b}$, a common assumption in image analysis is to take $\boldsymbol{Y}_j$ to be independently distributed given the group membership and bias field. Thus, with (3), the first term of (4) can be expressed as

$$\log p(\boldsymbol{y}|\boldsymbol{z},\boldsymbol{b},\boldsymbol{\theta}) = \sum_{i=1}^{g}\sum_{j=1}^{n} z_{ij}\log\phi_{\sigma_i}(y_j-\mu_i-b_j). \quad (6)$$

Let $\boldsymbol{\Psi}$ denote the vector of all the unknown parameters in the elements of $\boldsymbol{b}$ and $\boldsymbol{\theta}$. On the $(k+1)$th iteration of the EM algorithm, the E-step requires the calculation of

$$\begin{aligned} & Q_P(\boldsymbol{\Psi};\boldsymbol{\Psi}^{(k)}) \\ &= E\{\log L_{Pc}(\boldsymbol{b},\boldsymbol{\theta},\beta) \mid \boldsymbol{y},\boldsymbol{\Psi}^{(k)}\} \\ &= \sum_{i=1}^{g}\sum_{j=1}^{n} E\{Z_{ij} \mid \boldsymbol{y},\boldsymbol{\Psi}^{(k)}\}\log\phi_{\sigma_i}(y_j-\mu_i-b_j) \\ &\quad + E\{\log p(\boldsymbol{z}|\beta) \mid \boldsymbol{y},\boldsymbol{\Psi}^{(k)}\} - \tfrac{1}{2}\boldsymbol{b}^T\Sigma_b^{-1}\boldsymbol{b}, \end{aligned} \quad (7)$$

which is the conditional expectation of the penalized complete-data log likelihood given $\boldsymbol{Y}=\boldsymbol{y}$, using the current fit $\boldsymbol{\Psi}^{(k)}$ for $\boldsymbol{\Psi}$. On the M-step of the $(k+1)$th iteration, the intent is to find the value of $\boldsymbol{\Psi}$ that maximizes $Q_P(\boldsymbol{\Psi};\boldsymbol{\Psi}^{(k)})$, which gives $\boldsymbol{\Psi}^{(k+1)}$. The E- and M-steps are then alternated repeatedly until the penalized log likelihood changes by an arbitrary small amount, assuming convergence of the sequence of the penalized likelihood values.

E-step: We follow the approximation to the E-step considered in [21] by specifying the current conditional expectation of Z_{ij} given $\boldsymbol{y}$ and $\boldsymbol{\Psi}^{(k)}$ as

$$\begin{aligned} \tau_{ij}^{(k)} &= E\{Z_{ij} \mid \boldsymbol{y},\boldsymbol{\Psi}^{(k)}\} \\ &\approx E\{Z_{ij} \mid \boldsymbol{y},\boldsymbol{\Psi}^{(k)}, z_{N_j}=\hat{z}_{N_j}^{(k-1)}\} \\ &= \mathrm{pr}\{Z_{ij}=1 \mid \boldsymbol{y},\boldsymbol{\Psi}^{(k)}, z_{N_j}=\hat{z}_{N_j}^{(k-1)}\} \\ &= \frac{\pi_{ij}^{(k)}\phi_{\sigma_i^{(k)}}(y_j-\mu_i^{(k)}-b_j^{(k)})}{\sum_{h=1}^{g}\pi_{hj}^{(k)}\phi_{\sigma_h^{(k)}}(y_j-\mu_h^{(k)}-b_j^{(k)})}, \end{aligned} \quad (8)$$

where N_j is some specified neighbourhood of the jth voxel, containing s voxels, labelled $j_1,\ldots,j_s$, and $z_{N_j} = (z_{j_1}^T,\ldots,z_{j_s}^T)^T$ is the vector containing the group labels of these s voxels in N_j. In (8), $\pi_{ij}^{(k)} = \mathrm{pr}\{Z_{ij}=1 \mid z_{N_j} = \hat{z}_{N_j}^{(k-1)}\}$ is the probability that the jth voxel belongs to the ith group G_i given the group membership of its specified neighbours as implied by $\hat{z}_{N_j}^{(k-1)}$. As in [21], we adopt the MRF model

$$\log\pi_{ij}^{(k)} \propto \beta(\gamma_1\sum_m \hat{z}_{im}^{(k-1)} + \gamma_2\sum_m \hat{z}_{im}^{(k-1)} + \gamma_3\sum_m \hat{z}_{im}^{(k-1)}), \quad (9)$$

where the summations in (9) are, respectively, over the prescribed first-, second-, and third-neighbours of the jth voxel. The parameters γ_1,γ_2, and γ_3 control the spatial relatedness between neighbouring voxels [5]. In the third-order model adopted by [19] for 3D MR images, γ_1,γ_2, and γ_3 are set equal to 1, $1/\sqrt{2}$, and $1/\sqrt{3}$, respectively. In the calculation of $\pi_{ij}^{(k)}$ in (9), we followed McLachlan et al. [21] and replaced $\hat{z}_{im}^{(k-1)}$, which is zero or one, by $\tau_{im}^{(k-1)}$. This modification avoids the discretization in counting the neighbours of the jth voxel and effectively avoids premature classification of the voxel with insufficient neighbourhood information. It can be viewed as a fractional weight version of the ICM algorithm [26]. Initially, we calculated $\log\pi_{ij}^{(0)}$ by using $\hat{z}_{im}^{(0)}$ in the right-hand side of (9).

M-step: The M-step involves the maximization of $Q_P(\boldsymbol{\Psi};\boldsymbol{\Psi}^{(k)})$ with respect to $\boldsymbol{b}$ and $\boldsymbol{\theta}$. This maximization is implemented using a conditional approach, and the resulting algorithm can be viewed as an ECM algorithm. With the application of the ECM algorithm here, the M-step is replaced by two conditional maximization (CM) steps. The first involves the calculation of $\boldsymbol{b}^{(k+1)}$ by maximization (7) with $\boldsymbol{\theta}$ fixed at $\boldsymbol{\theta}^{(k)}$. The second CM step calculates $\boldsymbol{\theta}^{(k+1)}$ by maximization (7) with $\boldsymbol{b}$ fixed at $\boldsymbol{b}^{(k+1)}$.

On the $(k+1)$th iteration, the first CM-step yields

$$\begin{aligned} & \boldsymbol{b}^{(k+1)} \\ &= \arg\max_{\boldsymbol{b}} Q_P(\boldsymbol{b},\boldsymbol{\theta}^{(k)};\boldsymbol{b}^{(k)},\boldsymbol{\theta}^{(k)}) \\ &= \arg\max_{\boldsymbol{b}} \Bigg\{\sum_{i=1}^{g}\sum_{j=1}^{n}\tau_{ij}^{(k)}\log\phi_{\sigma_i^{(k)}}(y_j-\mu_i^{(k)}-b_j) \\ &\qquad -\tfrac{1}{2}\boldsymbol{b}^T\Sigma_b^{-1}\boldsymbol{b}\Bigg\} \\ &= \arg\max_{\boldsymbol{b}} \left\{\boldsymbol{b}^T\boldsymbol{r}^{(k)} - \tfrac{1}{2}\boldsymbol{b}^T\boldsymbol{D}^{(k)}\boldsymbol{b} - \tfrac{1}{2}\boldsymbol{b}^T\Sigma_b^{-1}\boldsymbol{b}\right\}, \end{aligned} \quad (10)$$

where, for $j=1,\ldots,n$,

$$r_j^{(k)} = (\boldsymbol{r}^{(k)})_j = \sum_{i=1}^{g}\tau_{ij}^{(k)}\left(\frac{y_j-\mu_i^{(k)}}{\sigma_i^{2(k)}}\right) \quad (11)$$

and $\boldsymbol{D}^{(k)}$ is a diagonal matrix with the diagonal elements

$$d_j^{(k)} = (\boldsymbol{d}^{(k)})_j = \sum_{i=1}^{g}\tau_{ij}^{(k)}\left(1/\sigma_i^{2(k)}\right). \quad (12)$$

Letting $\boldsymbol{I}$ denote the $n\times n$ identity matrix, it follows from (10) that we have

$$\boldsymbol{b}^{(k+1)} = (\boldsymbol{D}^{(k)}+\Sigma_b^{-1})^{-1}\boldsymbol{r}^{(k)}$$

$$= (\Sigma_b \boldsymbol{D}^{(k)} + \boldsymbol{I})^{-1} \Sigma_b \boldsymbol{r}^{(k)}$$
$$= (LL^T \boldsymbol{D}^{(k)} + \boldsymbol{I})^{-1} LL^T \boldsymbol{r}^{(k)}, \quad (13)$$

which coincides with result of Wells et al. [30]. To speed up the estimation of the bias field, Wells et al. [30] presented an efficient filtering method for (13) by replacing the matrix $\boldsymbol{D}^{(k)}$ with a vector $\boldsymbol{D}^{(k)}\mathbf{1}$ and the identity matrix $\boldsymbol{I}$ by the vector $\mathbf{1}$ as follows:

$$b_j^{(k+1)} = \frac{(L\boldsymbol{r}^{(k)})_j}{(L\boldsymbol{D}^{(k)}\mathbf{1})_j + 1} = \frac{(L\boldsymbol{r}^{(k)})_j}{(L\boldsymbol{d}^{(k)})_j + 1} \approx \frac{(L\boldsymbol{r}^{(k)})_j}{(L\boldsymbol{d}^{(k)})_j} \quad (14)$$

for $j = 1, \ldots, n$, where $\mathbf{1} = (1, 1, \ldots, 1)^T$. The linear transformation $L\boldsymbol{d}$ in (14) is implemented by convolution of the point spread function (psf) ℓ and the 3D image matrix A corresponding to L and $\boldsymbol{d}$, respectively. The convolution $\ell \odot A$ is operated repeatedly by 20–30 times in order to increase its lowpass filtering effect. An alternative filtering operation using Gaussian convolution may also be adopted [9].

With $\boldsymbol{b}$ fixed at $\boldsymbol{b}^{(k+1)}$ in (7), the second CM-step yields

$$\boldsymbol{\theta}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} Q_P(\boldsymbol{b}^{(k+1)}, \boldsymbol{\theta}; \boldsymbol{b}^{(k+1)}, \boldsymbol{\theta}^{(k)}) = \arg\max_{\boldsymbol{\theta}} \left\{ \sum_{i=1}^{g} \sum_{j=1}^{n} \tau_{ij}^{(k)} \log \phi_{\sigma_i}(y_j - \mu_i - b_j^{(k+1)}) \right\}, \quad (15)$$

which can be carried out in closed form as

$$\mu_i^{(k+1)} = \sum_{j=1}^{n} \tau_{ij}^{(k)} (y_j - b_j^{(k+1)}) / \sum_{j=1}^{n} \tau_{ij}^{(k)} \quad (16)$$

and

$$\sigma_i^{2(k+1)} = \sum_{j=1}^{n} \tau_{ij}^{(k)} (y_j - \mu_i^{(k+1)} - b_j^{(k+1)})^2 / \sum_{j=1}^{n} \tau_{ij}^{(k)} \quad (17)$$

for $i = 1, \ldots, g$.

The ECM algorithm preserves the appealing convergence properties of the EM algorithm. It thus has reliable global convergence in that it monotonely increases the penalized likelihood after each iteration, no matter what starting value is used. A detailed account of the convergence properties of the EM and ECM algorithms can be found in [20, 23]. In our proposed algorithm, we do not update the bias field (low frequency degrade) estimate $\boldsymbol{b}$ in every CM cycle. This approach speeds up the algorithm as the computational cost using an iterative lowpass filtering for a 3D MR image can be enormous. In addition, as the bias field is estimated using a more accurate estimate of $\boldsymbol{z}$ and $\boldsymbol{\theta}$ by more frequent update, it avoids the oscillations of the bias field estimate and hence improves the final segmentation and the bias field estimation, as demonstrated in Section 4. Our algorithm is summarized as follows:

1. *Obtain initial estimates of $\boldsymbol{b}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, and $\boldsymbol{z}^{(0)}$.*
2. *E-step: Calculate the posterior probabilities $\tau_{ij}^{(k)}$ based on (8).*
3. *CM-Step 1: Estimate the bias field $\boldsymbol{b}^{(k+1)}$ based on (14), given $\tau_{ij}^{(k)}$ and $\boldsymbol{\theta}^{(k)}$.*
4. *Do T cycles with $\boldsymbol{b}$ fixed at $\boldsymbol{b}^{(k+1)}$:*
   - *4.1 E-step: Calculate $\tau_{ij}^{(k+t/T)}$ based on (8),*
   - *4.2 CM-Step 2: Calculate $\boldsymbol{\theta}^{(k+t/T)}$ based on (16) and (17),*

   where the $\tau_{ij}^{(k+t/T)}$ and $\boldsymbol{\theta}^{(k+t/T)}$ denote the posterior probabilities and the value of $\boldsymbol{\theta}$, respectively, after the tth cycle on the $(k+1)$th iteration ($t = 1, \ldots, T$).
5. *Repeat from 2 until parameter sequences converge.*

As not all the CM-steps were performed in every CM cycle, we refer to our algorithm as a "partial" ECM algorithm. As an E-step is performed before each CM-step, the algorithm corresponds to a "multi-cycle ECM", where a cycle is defined by one E-step followed by one CM-step [24]. It can be seen that (10) and (15) imply

$$Q_P(\boldsymbol{b}^{(k+1)}, \boldsymbol{\theta}^{(k)}; \boldsymbol{b}^{(k)}, \boldsymbol{\theta}^{(k)}) \geq Q_P(\boldsymbol{b}^{(k)}, \boldsymbol{\theta}^{(k)}; \boldsymbol{b}^{(k)}, \boldsymbol{\theta}^{(k)})$$

and

$$Q_P(\boldsymbol{b}^{(k+1)}, \boldsymbol{\theta}^{(k+1)}; \boldsymbol{b}^{(k+1)}, \boldsymbol{\theta}^{(k)}) \geq Q_P(\boldsymbol{b}^{(k+1)}, \boldsymbol{\theta}^{(k)}; \boldsymbol{b}^{(k+1)}, \boldsymbol{\theta}^{(k)}).$$

It follows that the penalized log-likelihood

$$L(\boldsymbol{y}; \boldsymbol{b}, \boldsymbol{\theta}) - \tfrac{1}{2} \boldsymbol{b}^T \Sigma_b^{-1} \boldsymbol{b}$$

increases at each cycle and thus increases at each iteration monotonically if an exact E-step were used.

As in [14, 28], we set initially the bias field $\boldsymbol{b}^{(0)}$ to be zero. The initial estimates $\boldsymbol{\theta}^{(0)}$ and $\boldsymbol{z}^{(0)}$ can be obtained by performing the "noncontextual" segmentation of the MR image [21]. That is, the segmentation of the voxels is proceeded by ignoring all the spatial characteristics ($\beta = 0$) and the bias field estimation step (14). Van Leemput et al. [28] refer to this method as the independent model.

4. Experimental Results

We now demonstrate the use of the partial ECM algorithm for the fitting of the GMM-HMRF model with bias field correction. The first example is a simulated three-class

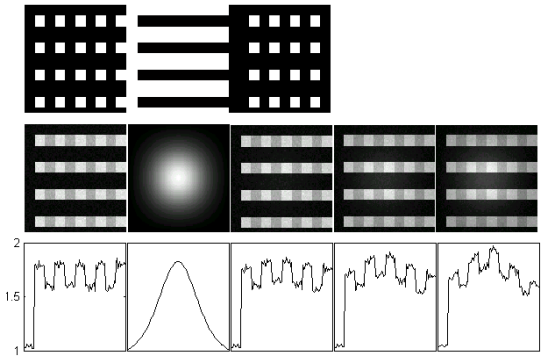

Figure 1. Simulated data: (From left to right) Top: the three groups; Middle: the true image, the true bias field, and the combined images with d=0.05, 0.12, and 0.20; Bottom: the intensity profile for each image

image obtained by adding the true image $\boldsymbol{y}^*$ and true bias field as

$$y_j = (1-d)y_j^* + db_j,$$

where $0 < d < 1$ is a constant governing the amount of bias field contamination. The true image simulates 256 gray-leveled brain MR image. The intensities for the white matter, gray matter, and the CSF are 150, 120, and 10, respectively. Gaussian noise with variance of 20, 14, and 12 is added to the three groups, respectively, before scaling by 256. The bias field simulates Gaussian-distributed function varied from zero (dark) to one (bright). Figure 1 presents the simulated data and the contaminated images. It can be seen that the image is heavily contaminated with bias field when $d = 0.20$. In this simulation experiment, we considered $g = 3$, $\beta = 0.9$, and $T = 3$. The results from fitting the GMM-HMRF model via our ECM algorithm are displayed in Figure 2 for $d = 0.20$.

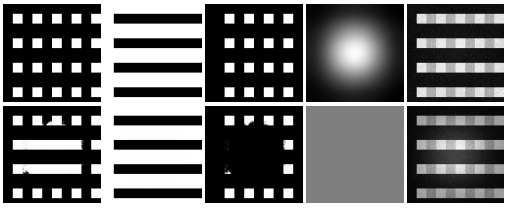

Figure 2. From left to right: the three groups segmented, estimated bias field, and the restored image. Top: GMM-HMRF model with bias field correction; Bottom: GMM-HMRF model without bias field correction

For comparison, we present also in Figure 2 the results obtained by the GMM-HMRF model without bias field correction. It can be seen that the GMM-HMRF algorithm fails to segment correctly. This result indicates how the segmentation is affected by the contamination of bias field. With the modification of GMM-HMRF model with bias field correction, it can be seen that almost perfect segmentation is obtained.

The second example is a real MR image of the human brain. The data set was a slice of a 3D T_1-weighted MR image, where the acquisition matrix was $256 \times 256 \times 256$. In this example, we considered $g = 3$ corresponding to the white matter, gray matter, and CSF. We adopted $\beta = 0.3$ and $T = 3$. The results are displayed in Figure 3. It can be seen that the three tissue types, brain-white matter, brain-gray matter, and CSF are well separated using the proposed GMM-HMRF model with bias field correction. For comparison, the segmented image from the GMM-HMRF algorithm without bias field correction is also given in Figure 3. It can be seen that white matter at the upper half of the brain is misclassified as gray matter, whereas the result is much better when the bias field correction is included in the algorithm.

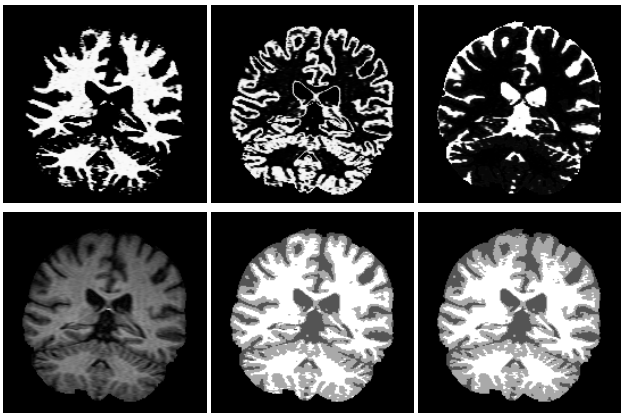

Figure 3. From left to right: Top (GMM-HMRF with bias field correction): Segmented white matter, Segmented gray matter, Segmented CSF; Bottom: Original image, Segmented image (with bias field correction), Segmented image (without bias field correction)

5. Discussion

For the segmentation of MR images with bias field correction, Wells et al. [30] have developed a mixture model-based approach via the EM algorithm to estimate the bias field and segment the images. Guillemaud and Brady [14] further refined this technique by introducing the extra tissue

class "other" and initializing the EM algorithm automatically for a given number of classes. Both methods assume statistical independence of the voxel intensities (that is, the noise in the MR signal is spatially white). Moreover, the parameters of each tissue class are required to be pre-defined or estimated in advance of applying the algorithm. Recently, Zhang et al. [31] proposed a hidden MRF model to allow for the spatial continuity of image intensities and the bias field correction simultaneously. This model adopts the ICM algorithm [2] to sequentially update each z_{ij}, which are zero or one, by local minimization of the conditional posterior probability. A maximum *a posteriori* (MAP) approach is applied to estimate the bias field, and the tissue parameters are estimated by maximum likelihood. However, this algorithm fails to segment correctly when the bias field contamination is heavy.

The EM algorithm is a popular tool in statistics for carrying out ML estimation because of its simplicity of implementation and reliable global convergence [20]. Here we have been able to develop an extension of the EM algorithm which can handle penalized ML estimation for the present problem, while still preserving the desirable properties of the EM algorithm. The extension of the GMM-HMRF model with bias field correction is justified using some simulated and real MR data, as shown in Section 4.

References

[1] R. G. Aykroyd and P. J. Green. Global and local priors, and the location of lesions using gamma-camera imagery. *Philos. Trans. Roy. Soc. A*, 337(1647):323–342, 1991.

[2] J. E. Besag. On the statistical analysis of dirty pictures (with discussion). *J. Roy. Stat. Soc., ser. B*, 48(3):259–302, 1986.

[3] J. E. Besag, J. York, and A. Mollié. Bayesian image restoration, with two applications in spatial statistics (with discussion). *Ann. I. Stat. Math.*, 43(1):1–59, 1991.

[4] G. D. Cascino, C. R. Jack, J. E. Parisi, et al. Magnetic resonance imaging-based volume studies in temporal lobe epilepsy: Pathological correlations. *Ann. Neurol.*, 30(1):31–36, 1991.

[5] H. S. Choi, D. R. Haynor, and Y. Kim. Multivariate tissue classification of MRI images for 3D volume reconstruction – a statistical approach. *Proceedings SPIE Medical Imaging III: Image Processing*, 1092:183–193, 1989.

[6] L. P. Clarke, R. P. Velthuizen, M. A. Camacho, et al. MRI segmentation: Methods and applications. *Magn. Reson. Imag.*, 13(3):343–368, 1995.

[7] B. M. Dawant, A. P. Zijdenbos, and R. A. Margolin. Correction of intensity variations in MR images for computer-aided tissue classification. *IEEE Trans. Med. Imag.*, 12(4):770–781, 1993.

[8] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. Roy. Stat. Soc., ser. B*, 39(1):1–38, 1977.

[9] R. Deriche. Recursively implementing the Gaussian and its derivatives. *Technical Report 1893*. INRIA Sophia-Antipolis, France, 1993.

[10] M. Freund, S. Hahnel, M. Thomsen, and K. Sartor. Treatment planning in severe scoliosis: the role of MRI. *Neuroradiology*, 43(6):481–484, 2001.

[11] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, 6(6):721–741, 1984.

[12] G. Gerig, O. Kübler, R. Kikinis, and F. A. Jolesz. Nonlinear anisotropic filtering of MRI data. *IEEE Trans. Med. Imag.*, 11(2):221–232, 1992.

[13] P. J. Green. On use of the EM algorithm for penalized likelihood estimation. *J. Roy. Stat. Soc., ser. B*, 52(3):443–452, 1990.

[14] R. Guillemaud and M. Brady. Estimating the bias field of MR images. *IEEE Trans. Med. Imag.*, 16(3):238–251, 1997.

[15] L. Jancke, T. W. Buchanan, K. Lutz, and N. J. Shah. Focused and nonfocused attention in verbal and emotional dichotic listening: An FMRI study. *Brain Lang.*, 78(3):349–363, 2001.

[16] J. Kay and D. M. Titterington. Image labelling and the statistical analysis of incomplete data. In *Proc. 2nd Int. Conf. Image Processing and Applications*. Institute of Electrical Engineers, London, 1986, pp. 44–48.

[17] R. Leahy and X. Yan. Incorporation of anatomical MR data for improved functional imaging with PET. *Lect. Notes Comput. Sc.*, 511:105–120, 1991.

[18] S. K. Lee and M. W. Vannier. Post-acquisition correction of MR inhomogeneities. *Magn. Reson. Med.*, 36(2):275–286, 1996.

[19] Z. Liang, J. R. MacFall, and D. P. Harrington. Parameter estimation and tissue segmentation from multispectral MR images. *IEEE Trans. Med. Imag.*, 13(3):441–449, 1994.

[20] G. J. McLachlan and T. Krishnan. *The EM Algorithm and Extensions*. Wiley, New York, 1997.

[21] G. J. McLachlan, S. K. Ng, G. Galloway, and D. Wang. Clustering of magnetic resonance images. In *Proc. American Statistical Assoc. (Statistical Computing Section)*. American Statistical Assoc., Alexandria, Virginia, 1996, pp. 12–17.

[22] G. J. McLachlan and D. Peel. *Finite Mixture Models*. Wiley, New York, 2000.

[23] X. L. Meng. On the rate of convergence of the ECM algorithm. *Ann. Stat.*, 22(1):326–339, 1994.

[24] X. L. Meng and D. B. Rubin. Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika*, 80(2):267–278, 1993.

[25] J. A. O'Sullivan. Roughness penalties on finite domains. *IEEE Trans. Image Processing*, 4(9):1258–1268, 1995.

[26] A. Owen. Contribution to the discussion of paper by B.D. Ripley. *Canad. J. Statist.*, 14(1):106–110.

[27] W. Qian and D. M. Titterington. Estimation of parameters in hidden Markov models. *Philos. Trans. Roy. Soc. A*, 337(1647):407–428, 1991.

[28] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classification of MR images of the brain. *IEEE Trans. Med. Imag.*, 18(10):897–908, 1999.

[29] M. W. Vannier, R. L. Butterfield, D. Jordon, et al. Multispectral analysis of magnetic resonance images. *Radiology*, 154(1):221–224, 1985.

[30] W. M. Wells, III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz. Adaptive segmentation of MRI data. *IEEE Trans. Med. Imag.*, 15(4):429–442, 1996.

[31] Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imag.*, 20(1):45–57, 2001.