ROBUST UNSUPERVISED TISSUE CLASSIFICATION IN MR IMAGES

Dzung L. Pham

Dept. of Radiology & Radiological Science
Johns Hopkins University
Baltimore, MD 21287

Jerry L. Prince

Dept. of Electrical & Computer Engineering
Johns Hopkins University
Baltimore, MD 21218

ABSTRACT

A general framework for performing robust, unsupervised tissue classification in magnetic resonance images is presented. Tissue classification is formulated as an estimation problem based on an imaging model. Prior models are used within the estimation problem to compensate for noise and intensity inhomogeneity artifacts. From this framework, approaches based on K-means clustering, clustering via the expectation-maximization algorithm, and fuzzy clustering can be derived. The performance of the different types of approaches are evaluated using both simulated and real neuroimaging data.

1. INTRODUCTION

Accurate tissue classification in magnetic resonance (MR) images is an important objective in medical imaging applications. Segmentation methods for performing tissue classification are hindered, however, by multiple imaging artifacts such as noise, intensity inhomogeneities, and partial volume effects. Noise in MR images can induce segmented regions to become disconnected or possess holes. Intensity inhomogeneities, which are caused by nonuniformities in the RF field and other factors [1], produce a shading artifact over the image that can confound intensity-based segmentation methods. Partial volume effects occur where multiple tissues contribute to a single voxel, resulting in a blurring of intensities at tissue boundaries.

In this paper, a general framework for unsupervised (i.e., requiring no training data) tissue classification is presented. Methods derived from the proposed imaging model are robust to both noise and intensity inhomogeneities. Under this framework, we describe algorithms based on K-means clustering, the expectation-maximization (EM) algorithm, and fuzzy clustering. Unlike the K-means algorithm, the latter two approaches are capable of producing soft segmentations that can better model partial volume effects. The performance of the various algorithms is compared using both real and simulated MR images of the brain. This paper can be considered both a unification and generalization of the authors' work presented in [2, 3, 4] with a considerable amount of additional experimental validation.

2. METHODS

The true image is assumed to be composed of K tissue classes, with each class having a distinct, unknown, intensity value called the class mean, denoted $v_k, k = 1, \ldots, K$. Because of intensity inhomogeneities or variability in the tissue itself, the class mean may be spatially varying in the image domain Ω. Two cases are considered in regard to this variation. In the more general case, each mean can be considered to vary independently. In the second case, each mean varies according to a multiplicative gain field, g_j. In the first case, the observed image $\mathbf{y}$ can be expressed as

$$y_j = \sum_{k=1}^{K} z_{jk} v_{jk} + \eta_j, \; j \in \Omega, \tag{1}$$

whereas in the gain field case, we substitute $v_{jk} = g_j v_k$. Under both models, y_j is the observed intensity at pixel j, and η_j is white Gaussian noise with unknown variance σ^2. We define z_{jk} to be an indicator function that satisfies

$$z_{jk} = \begin{cases} 1, & \text{if pixel } j \text{ is in tissue class } k \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

and $\sum_{k=1}^{K} z_{jk} = 1$ (i.e., classes do not overlap). Thus, a segmentation can be obtained by estimating the indicator functions. The likelihood function under the mean-varying model is

$$f(\mathbf{y}; \mathbf{z}, \theta) = \prod_{j \in \Omega} \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-(y_j - \sum_{i=1}^{K} z_{jk} v_{jk})^2 / 2\sigma^2), \tag{3}$$

where bold characters represent stacked vectors, and θ is the parameter vector consisting of the means and variance. The likelihood function in the gain field case is analogous.

A Markov random field prior model is used for the indicator functions z_{jk} and is given by

$$f(\mathbf{z}) = \frac{1}{Z_z} \prod_{j \in \Omega} \exp(-\beta \sum_{i \in N_j} \mathbf{z}_j^T \mathbf{V} \mathbf{z}_i), \tag{4}$$

0-7803-8388-5/04/$20.00 ©2004 IEEE

where β is a parameter controlling the strength of the prior and is determined empirically, N_j is the set of the four nearest neighbors of j for 2-D images and six nearest neighbors for 3-D images, and Z_z is a normalizing constant. The matrix $\mathbf{V}$ is $K \times K$ and is used to penalize the classification of pixel j based on its neighbors. In this work, $\mathbf{V} = [1,\ldots,1][1,\ldots,1]^T - I$, where I is the $K \times K$ identity matrix. Thus, a pixel belonging to the same tissue class as its neighbors is favored over configurations consisting of different classes.

The brightness variation of the class means or gain field is assumed to be slowly varying and smooth. These properties can be enforced by representing the means as the summation of smooth basis functions (e.g. splines, polynomials). Here, the variation is modeled by assuming that v possesses the following prior probability

$$f(\mathbf{v}) = \frac{1}{Z_v} \prod_{j \in \Omega} \exp\left(-\lambda \sum_{k=1}^{K} \sum_{r=1}^{R} (D_r * v)_{jk}^2 \right), \qquad (5)$$

where Z_v is a normalizing constant, R is the dimensionality of the image (i.e., $R = 2$ for 2-D images and $R = 3$ for 3-D images), and D_r are finite difference operators along the dimensions of the image. These finite difference operators act like derivatives in the discrete domain. The parameter λ is assumed to be known and controls the amount of smoothness and the overall variation of the means. Second order differences [4] may be used for greater robustness at the expense of computational complexity.

2.1. K-means approach

One approach for estimating an optimal segmentation under the given model is closely related to a K-means algorithm. Maximization of the joint posterior probability function is equivalent to minimizing the following sum of three energy functions:

$$E = \frac{1}{2\sigma^2} E_1 + \lambda E_2 + \beta E_3, \qquad (6)$$

where $E_1 = \sum_{j \in \Omega} \sum_{k=1}^{K} z_{jk} (y_j - v_{jk})^2$, $E_2 = \sum_{j \in \Omega} \sum_{k=1}^{K} \sum_{r=1}^{R} (D_r * v)_{jk}^2$, and $E_3 = \sum_{j \in \Omega} \sum_{i \in N_j} \mathbf{z}_j^T \mathbf{V} \mathbf{z}_i$. The equation for E_1 follows from the properties that $z_{jk}^2 = z_{jk}$ and $\sum_{k=1}^{K} z_{jk} = 1$. Except for the spatially varying means, E_1 is exactly the K-means clustering algorithm objective function. Minimization of E is therefore equivalent to a penalized K-means algorithm and can be solved using an iterative approach. This is a generalization of the approach proposed by Unser [5]. Without loss of generality, we assume $\sigma^2 = 1/2$, since its value merely influences the balance of the penalty terms and can already be controlled by λ and β.

To estimate the indicator functions, the following computation is performed for each pixel $j \in \Omega$:

$$\mathbf{z}_j = \arg\min_{\mathbf{z}_j} (f_j - g_j \sum_{k=1}^{K} z_{jk} v_k)^2 + \lambda \sum_{i \in N_j} \mathbf{z}_j^T \mathbf{V} \mathbf{z}_i. \qquad (7)$$

This step is similar to the *iterated conditional modes* algorithm. The parameter estimation step can be derived by solving for a zero gradient of (6) with respect to each parameter. In the spatially varying mean case, this leads to solving the following difference equation for $k = 1,\ldots,K$:

$$y_j z_{jk} = z_{jk} v_{jk} + 2\beta_1 (H_1 * v)_{jk}, \qquad (8)$$

where the convolution kernel H_1 is given by

$$H_1 = \sum_{r=1}^{R} (D_r * \check{D}_r)_j. \qquad (9)$$

For the gain field case, the means are computed using the equation:

$$v_k = \frac{\sum_{j \in \Omega} z_{jk} g_j y_j}{\sum_{j \in \Omega} z_{jk} (g_j^2)}. \qquad (10)$$

The gain field is computed by solving for g in the following equation.

$$y_j \sum_{k=1}^{K} z_{jk} v_k = g_j \sum_{k=1}^{K} z_{jk} v_k^2 + 2\beta_1 (H_1 * g)_j. \qquad (11)$$

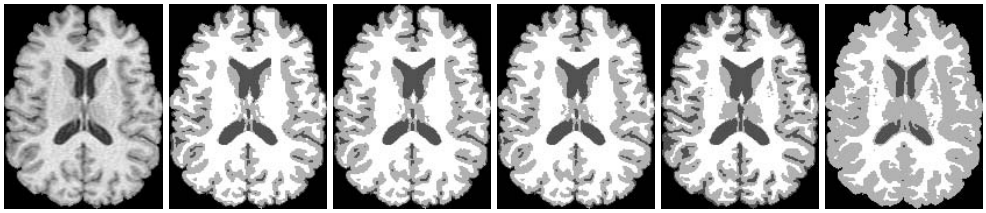
Equations (8) and (11) can both be solved efficiently using a multigrid algorithm as described in [4].

2.2. EM Approach

An alternative approach to optimization of (6) is to use an EM algorithm [6]. Using the EM approach results in several differences from the K-means approach. The full derivation is similar to that presented in [2] and is omitted here. First, the classification step is replaced by the so called E-step. Because of the Markov random field assumption, a mean-field approximation [7] is used to estimate the posterior probabilities w_{jk}. The resulting equation for the spatially varying mean case is

$$w_{jk} = \frac{\exp((y_j - v_{jk})^2/\sigma^2 + \beta \sum_{i \in N_j} \sum_{m \neq k} w_{jm})}{\sum_{l=1}^{K} \exp((y_j - v_{jl})^2/\sigma^2 + \beta \sum_{i \in N_j} \sum_{m \neq k} w_{jm})}. \qquad (12)$$

Because the variance parameter now plays an important role in the E-step, it must be estimated during the parameter estimation step [2]. The equations for estimating the class means and/or gain field are nearly identical to the K-means equations. It can be shown that these parameter estimation steps satisfy the conditions for a generalized EM algorithm [2]. It is also possible to allow the priors and variance to vary across tissue classes, equivalent to the standard finite normal mixture model assumption [6].



(a) (b) (c) (d) (e) (f)

Fig. 1. Application of different algorithms on a T1-weighted data set: (a) slice from MR image of the brain, (b) GKM classification, (c) GFC classification, (d) GEM classification, (e) GEMV classification, (f) GEMPV classification.

2.3. Fuzzy clustering approach

Another alternative to extending the K-means algorithm to provide soft segmentations is to use a fuzzy clustering approach. Although fuzzy clustering does not have an explicit statistical interpretation, it can be derived through a direct generalization of the energy equation (6) by allowing the indicator function to be continuously valued between 0 and 1, and placing an exponent parameter within the objective function [8]. In this case, the following fuzzy clustering replacement for E_1 is obtained:

$$E_1^{(\mathrm{FC})} = \sum_{j\in\Omega}\sum_{i=1}^{K} u_{jk}^{q}(y_j - g_j v_k)^2, \tag{13}$$

where u_{jk} are now called *membership functions*. When $q = 1$, it has been shown that the membership functions will behave like binary indicator functions [8]. As q increases, the membership functions become increasingly "fuzzy".

To enforce spatial continuity within the membership functions, the neighborhood interaction energy can be generalized in a similar manner [3]:

$$E_3^{(\mathrm{FC})} = \sum_{j\in\Omega}\sum_{i\in N_j} (\mathbf{u}_j^q)^T \mathbf{V}\mathbf{u}_i^q,$$

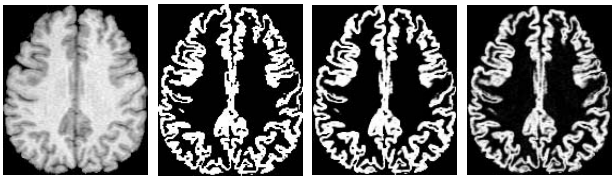
where $\mathbf{u}_j^q = [u_{j1},\ldots,u_{jK}]^T$. Using this new energy function, the classification step now becomes:

$$u_{jk} = \frac{\left((y_j - v_{jk})^2 + 2\beta\sum_{i\in N_j}\sum_{m\neq k} u_{im}^q\right)^{-1/(q-1)}}{\sum_{l=1}^{K}\left((y_j - v_{lk})^2 + 2\beta\sum_{i\in N_j}\sum_{m\neq i} u_{im}^q\right)^{-1/(q-1)}}. \tag{14}$$

Note the similarities between (14) and (12). The mean/gain field estimation for the fuzzy clustering approach can be obtained simply by substituting u_{jk}^q for z_{jk} within the K-means equations.

3. RESULTS

In this section, the different approaches are compared using simulated and real MR images of the brain. All brain images were preprocessed to remove extracranial tissue. Three classes were assumed, corresponding to gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). Three-dimensional data were used in all the experiments. Parameter settings were determined empirically, but were fixed for each experiment (except where noted). The following abbreviations are used in this section to designate the different algorithm. A 'G' denotes gain field methods, while 'M' denotes methods using independently varying means. 'KM' is K-means,'EM' is expectation-maximization, and 'FC' is fuzzy clustering. For all soft segmentation algorithms, a hard classification was obtained by assigning each pixel to the class of highest probability or membership. For GFC and MFC, the q parameter was set to 2. See [4] for the initialization techniques used.



(a) (b) (c) (d)

Fig. 2. Soft segmentation results using different methods: (a) slice from MR image of the brain, (b) GKM gray matter estimate, (c) GEM gray matter estimate, (d) GFC gray matter estimate.

3.1. Comparisons on a single data set

Figure 1 shows the results of the hard classification obtained by the various algorithms to a T1-weighted, high resolution ($0.9375\times0.9375\times1.5$mm) MR image. The spatially varying mean results are nearly identical to the gain field results and are therefore not shown. The GKM, GFC, and GEM are almost extremely similar to one another and seem to underestimate subcortical GM. Also shown are the results of using a GEM approach with class-varying variances (GEMV), and class-varying priors and variances (GEMPV). GEMV appears to overestimate sulcal CSF, while GEMPV overestimates cortical GM. The GKM, GFC, and GEM methods perform well in cortical regions but somewhat underestimate subcortical GM.

Figure 2 shows the GM soft segmentation produced by the various methods. The GKM result is included as a ref-

Table 1. Simulated MR errors

Method	*Data set* 3%N,0%I	3%N,20%I	3%N,40%I	5%N,20%I	7%N,20%I
RKM-MCR	4.07%	4.31%	4.51%	5.34%	7.02%
REM-MCR	4.02%	4.13%	4.52%	5.31%	6.69%
RFC-MCR	3.90%	4.11%	4.78%	5.20%	6.89%
FCM-MCR	3.98%	5.45%	9.04%	7.60%	10.51%
RKM-RMS	0.206	0.231	0.235	0.242	0.260
REM-RMS	0.221	0.229	0.256	0.225	0.229
RFC-RMS	0.124	0.126	0.133	0.142	0.169
FCM-RMS	0.139	0.165	0.227	0.208	0.259

erence point of a true binary result. The figure demonstrates that the EM results are much more binarized than the fuzzy clustering results.

3.2. Simulated MR results

Table 1 shows the classification obtained using the various algorithms when applied to the Brainweb simulated MR image database (*http://www.bic.mni.mcgill.ca/brainweb/*). T1-weighted MR images were simulated with 1mm resolution and different amounts of noise (3%-7%) and inhomogeneity (0%-40-%). The misclassification rate, defined as the number of pixels misclassified divided by the total number of pixels within the brain volume, was computed for each algorithm. Very similar results were obtained with the GKM, GFC, and GEM methods. Most of the differences can be attributed to slight differences in the selection of λ and β. Note that the Brainweb data simulates inhomogeneities as a gain field, thus the mean-varying methods will perform slightly worse.

Although the hard segmentations yield similar results between methods, the soft segmentation errors given in Table 1 reveal a distinct advantage to the fuzzy clustering methods. The errors of the fuzzy clustering are also lower than those given in [9], particularly as noise levels increase.

3.3. IBSR results

The algorithms were also applied to the 20 Internet Brain Segmentation Repository (IBSR) data sets (*http://neuro-www.mgh.harvard.edu/cma/ibsr/*), which include manually segmented truth models. To be consistent with the database, an overlap measure was used, defined as the ratio of the intersection between the computed and true tissue class to the union. Four data sets had severe artifacts and required custom settings of λ and β. In general, the results in Fig. 3 are competitive with or superior to previously published results [9]. The mean-varying methods seem to provide little advantage over the gain field methods. It should be noted that fuzzy clustering methods and EM methods yielded very low standards of deviation, indicating that there are likely

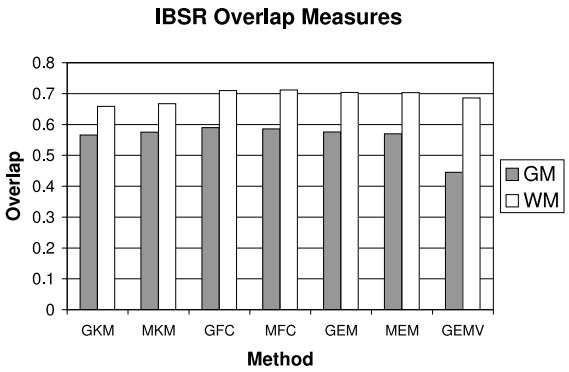


Fig. 3. Average overlap measures for different methods using the IBSR data

some systematic differences between the computed segmentations and the truth model. In particular, the truth models contain no sulcal CSF– these regions are classified as all gray matter.

4. REFERENCES

[1] J.G. Sled and G.B. Pike, "Standing-wave and RF penetration artifacts caused by elliptic geometry: an electrodynamic analysis of MRI," *IEEE Trans. on Medical Imaging*, vol. 17, pp. 653–662, 1998.

[2] D.L. Pham and J.L. Prince, "A generalized EM algorithm for robust segmentation of magentic resonance images," in *Proceedings of the 33rd Annual Conference on Information Sciences and Systems (CISS99)*. Johns Hopkins University, 1999.

[3] D. L. Pham, "Spatial models for fuzzy clustering," *Computer Vision and Image Understanding*, vol. 84, no. 2, pp. 285–297, 2001.

[4] D.L. Pham and J.L. Prince, "Adaptive fuzzy segmentation of magnetic resonance images," *IEEE Trans. on Medical Imaging*, vol. 18, no. 9, pp. 737–752, 1999.

[5] M. Unser, "Multigrid adaptive image processing," in *Proc. of the IEEE Conference on Image Processing (ICIP95)*, 1995, vol. I, pp. 49–52.

[6] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society Series B*, vol. 39, pp. 1–38, 1977.

[7] J. Zhang, "The mean field theory in EM procedures for Markov random fields," *IEEE Trans. on Signal Processing*, vol. 40, pp. 2570–2583, 1992.

[8] J.C. Bezdek, "A convergence theorem for the fuzzy ISODATA clustering algorithms," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. PAMI-2, pp. 1–8, 1980.

[9] D.W. Shattuck, S.R. Sandor-Leahy, K.A. Schaper, D.A. Rottenberg, and R.M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *Neuroimage*, vol. 13, pp. 856–876, 2001.