Generalized fuzzy clustering for segmentation of multi-spectral magnetic resonance images

Renjie He*, Sushmita Datta¹, Balasrinivasa Rao Sajja², Ponnada A. Narayana³

Department of Diagnostic and Interventional Imaging, University of Texas Medical School at Houston, 6431 Fannin, Houston, TX 77030, United States

Received 3 January 2007; received in revised form 18 February 2008; accepted 20 February 2008

Abstract

An integrated approach for multi-spectral segmentation of MR images is presented. This method is based on the fuzzy c-means (FCM) and includes bias field correction and contextual constraints over spatial intensity distribution and accounts for the non-spherical cluster’s shape in the feature space. The bias field is modeled as a linear combination of smooth polynomial basis functions for fast computation in the clustering iterations. Regularization terms for the neighborhood continuity of intensity are added into the FCM cost functions. To reduce the computational complexity, the contextual regularizations are separated from the clustering iterations. Since the feature space is not isotropic, distance measure adopted in Gustafson–Kessel (G–K) algorithm is used instead of the Euclidean distance, to account for the non-spherical shape of the clusters in the feature space. These algorithms are quantitatively evaluated on MR brain images using the similarity measures.

© 2008 Elsevier Ltd. All rights reserved.

Keywords: Adaptive FCM; Contextual constraints; G–K algorithm; Inhomogeneity field; Multi-spectral segmentation; MRI; Similarity measures

1. Introduction

The superb soft tissue contrast seen on magnetic resonance imaging (MRI) is ideally suited for tissue classification and volumetry. This has significant implications in understanding the neural basis for many neurological disorders [19]. For instance, in a number of neurological disorders, such as multiple sclerosis (MS) and Alzheimer’s disease, the volume changes in total brain, gray matter (GM), and white matter (WM) provide important information about the neuronal and axonal loss [4,25,26]. In addition, MRI-derived tissue volumetry is increasingly employed as a secondary end point in many clinical trials [22]. Accurate and robust tissue classification or segmentation is critical for detecting changes in tissue volumes in healthy and diseased brain. Commonly used techniques for segmentation have been recently reviewed [24].

A unique feature of MRI is its multi-model nature that allows acquisition of images with different tissue contrasts (T1-, T2-, density-weighting, etc.) It is possible to improve the quality of segmentation by combining information from images with multiple contrasts [3,13,16,32]. Feature map-based classification techniques for MR image segmentation have attracted considerable attention because they are fast, simple to implement, and allow expert’s input in tissue classification. However, in practice, this multi-spectral segmentation is prone to false tissue classifications and requires significant manual intervention and pre-processing since the distribution of intensities in the feature space is distorted by various factors that include image intensity inhomogeneity arising from the radio frequency receiver and transmitter coil profiles, partial volume averaging effects from the limited resolution, image noise, and spatial misalignment of images. While there are a few methods to overcome some of these problems, we focus on fuzzy c-means (FCM)-based methods [1,8,10,23,27,28,31,33] because of their many desirable features in tissue classification.

Conventional FCM-based methods do not correct the intensity inhomogeneity and do not exploit contextual information. The adaptive FCM (AFCM) incorporates the intensity inhomogeneity correction, contextual constraints to overcome the noise problems and fuzzy membership to address the...
partial volume averaging effect, and automatic clustering 
\[1,20,23,27,28,30,34\].

One essential problem with AFCM (and a number of other clustering algorithms) is that they are totally based on the objective cost function. Therefore, the performance of the algorithms is greatly dependent on the way the objective cost function is constructed. When using the Euclidean distance in the objective cost function, as in FCM, the algorithm has a tendency to generate equal cluster volumes with spherical occupancy in the feature space \([7,8,9,14,21]\). This could have a significant effect on the MRI segmentation results. To deal with this problem, a few methods have been proposed in MR image segmentation in which some pre-selected seeds are included \([5,6]\). Methods aimed at automating the selection of seeds are proposed by Suckling et al. \([30]\). However, these methods are cumbersome to implement. In order to automatically produce reasonable clusters, more sophisticated distance measure is included in the Gustafson–Kessel (G–K) algorithm \([15,17]\) in which a positive definite, symmetric scatter matrix (or covariance matrix) is used instead of the Euclidean distance to define the Mahalanobis distance to form an ellipsoidal cluster in the feature space. In the G–K algorithm each cluster is characterized by a symmetric and positive definite matrix for automatic adaptation of each individual cluster volume.

The fundamentals of fuzzy clustering in medical image segmentation are well described in Sutton et al. \([31]\). The main formalism presented in this paper is based on the grouped coordinate descent method (also called alternating optimization (AO)) \([31]\). Important details, such as singularity, initialization, rules of thumb for the parameters, methods to determine the number of clusters are not described in this paper due to space limitations \([31]\). Throughout this paper, we assumed that the total number of clusters is known and that proper initialization is available.

In the current studies, we extended AFCM to multi-spectral segmentation that included efficient intensity non-uniformity (or bias field) correction and contextual constraints over neighborhood spatial intensity distribution. The non-spherical occupancy of the feature space was accounted for utilizing cluster scatter measures to define the Mahalanobis distance. This method was applied to segment GM, WM and CSF (cerebrospinal fluid) of MR images acquired with fast spin echo (FSE) pulse sequence by Suckling et al. \([30]\). However, these methods are cumbersome to implement. In order to automatically produce reasonable clusters, more sophisticated distance measure is included in the objective function as suggested by others \([1,20,27]\). With the inclusion of the regularization terms, Eq. (1) can be written as

\[
J = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^p ||x_k - v_i||^2
\]

subject to

\[
\sum_{i=1}^{c} u_{ik} = 1, \quad u_{ik} \in [0, 1], \quad 0 < \sum_{k=1}^{N} u_{ik} < \sum_{k=1}^{N} u_{ik} < N
\]

where \(u_{ik}\) is the membership of \(k\)-th voxel belonging to class-\(i\), \(v_i\) is the cluster center of class-\(i\), and \(p\) is a preset weighting exponent or fuzzifier. To include the influence of immediate neighborhood for forcing the solution towards piecewise-homogeneous labeling, regularization terms are introduced into the objective function as suggested by others \([1,20,27]\). With the inclusion of the regularization terms, Eq. (1) can be written as

\[
J_m = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^p ||x_k - v_i||^2 + \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^p \left( \sum_{x_r \in N_k} \left( \frac{\alpha}{N_k} ||x_r - v_i||^2 \right) \right)
\]

where \(N_k\) represents the neighbors of current voxel, and \(N_k\) is the cardinality of \(N_k\). The regularization term can be adjusted by setting the value of \(\alpha\) corresponding to neighborhood constraints in Eq. (3) to compromise between the sharp segmentation and the pulse noise \([1,27]\).

In general, the distribution of MR image intensities in the feature space is distorted due to the presence of bias field that needs to be corrected for proper tissue classification. The intensity of a voxel located at the spatial position \(k (k = 1, \ldots, N)\), where \(N\) is the number of voxels) can be represented as \([28]\).

\[
o_k = G_k t_k + \text{noise}(k)
\]

where \(o_k\) is the observed intensity, \(t_k\) is the true intensity, \(G_k\) is the diagonal matrix representing the gain field, and \(\text{noise}(k)\) is the noise. Since we are mainly interested in multi-spectral case, all the above variables, except \(G_k\), are vectors. Assuming \(n\) image channels, we can explicitly express \(a_k = [a_{k1}, a_{k2}, \ldots, a_{kn}]^T\), and \(t_k = [t_{k1}, t_{k2}, \ldots, t_{kn}]^T\). The gain field can be denoted as \(g_k = [g_{k1}, g_{k2}, \ldots, g_{kn}]^T\), with \(G_k = \text{diag}(g_{k1}, g_{k2}, \ldots, g_{kn})\) and Eq. (4) can be rewritten as

\[
[a_{k1}, \ldots, a_{kn}]^T = [g_{k1} t_{k1}, \ldots, g_{kn} t_{kn}]^T + \text{noise}(k).
\]

By ignoring the noise term and applying the log-transform on both sides of Eq. (5), denoting the log-transformed observed MR image data as \(y_k\), the MR image data can be approximated as \([1,23]\).

\[
y_k = b_k + x_k
\]

where \(b_k = [b_{k1}, b_{k2}, \ldots, b_{kn}]^T\) is the vectorial voxel representation of the bias field.
Generally, the bias field can be approximated as \( b_k = \sum q_i \phi_i(p_k) \), where \( \phi_i(p_k) \) is a smooth basis function and \( p_k \) represents the coordinates. Therefore, \( b_k \) can be written as

\[
\begin{align*}
\mathbf{b}_k &= \left[ \sum q_1 \phi_1(p_k), \sum q_2 \phi_2(p_k), \ldots, \sum q_n \phi_n(p_k) \right]^T \\
&= \mathbf{Q} \Phi(p_k)
\end{align*}
\]

(7)

with

\[
\Phi(p_k) = [\phi_1(p_k), \phi_2(p_k), \ldots, \phi_n(p_k)]^T
\]

and

\[
\mathbf{Q} = \begin{bmatrix}
q_{11} & q_{12} & \cdots & q_{1m} \\
q_{21} & q_{22} & \cdots & q_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
q_{n1} & q_{n2} & \cdots & q_{nm}
\end{bmatrix}
\]

where \( m \) is the number of basis functions in \( \Phi \). The whole bias field can be expressed as

\[
\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{N-1}, \mathbf{b}_N].
\]

(8)

In practice, the approximation of the bias field is over the 3D space and the spatial relation have to be included into the expansion of the smooth basis functions.

By the piecewise-homogeneous assumption, the influence of the bias field is included in the objective function (Eq. (3)). Introducing a normalizing factor \( e \), the modified objective function can be written as

\[
J_m = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p |y_i - \mathbf{Q} \Phi(p_k) + e - v_i|^2
\]

\[
+ \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \left( \sum_{x_i \in N_k} \frac{\alpha}{N_R} |y_r - \mathbf{Q} \Phi(p_r) + e - v_i|^2 \right)
\]

(9)

subject to

\[
\sum_{i=1}^c u_{ik} = 1, \quad u_{ik} \in [0, 1], \quad 0 < \sum_{k=1}^N u_{ik} < N
\]

Since the amplitude, \( b_k \), is arbitrary, normalization of \( b_k \) is necessary to ascertain that the sum of \( b_k \) is zero to satisfy the convergence condition of alternating optimization (AO). The normalization factor, \( e \), will maintain the sum of \( b_k \) to zero and plays an important role in the convergence of the iterations in the AO that can be applied to Eq. (9) to arrive at the (local) optima. The image segmentation is achieved by solving

\[
\min_{[u_{ik}, |v_i|, \{b_k\}]} J_m, \quad \text{subject to} \sum_{i=1}^c u_{ik} = 1.
\]

(10)

### 2.2. Extension of AFCM algorithm with G–K measure

Since the Euclidean distance measure is used in the objective function to compute the distance between data points and prototypes of clusters, FCM method works well for the spherical shaped clusters with equal volumes. To take the cluster shape into consideration, we incorporate the covariance matrix of each cluster into the calculation of distance measures using G–K algorithm as described in the following section.

In the G–K algorithm, a more sophisticated distance measure, the Mahalanobis distance, based on positive definite, symmetric scatter matrix (or covariance matrix) is used to account for the scatter shape of each cluster and the ellipsoidal occupancy of clusters in the feature space [15]. The fuzzy covariance matrix, \( S_m \), is given by \([15,17]\).

\[
S_i = \sum_{k=1}^N u_{ik}^p (y_i - \mathbf{Q} \Phi(p_k) + e - v_i)(y_i - \mathbf{Q} \Phi(p_k) + e - v_i)^T
\]

(11)

We describe the algorithm for extending the G–K algorithm to multi-channel MR image segmentation by denoting

\[
D_{ik} = (y_i - \mathbf{Q} \Phi(p_k) + e - v_i)^T A_i (y_i - \mathbf{Q} \Phi(p_k) + e - v_i)
\]

(12)

and

\[
\Upsilon_{ik} = \sum_{y_r \in N_k} \left( \frac{\alpha}{N_R} |y_r - \mathbf{Q} \Phi(p_r) + e - v_i|^2 \right)
\]

\[
\times A_i (y_r - \mathbf{Q} \Phi(p_r) + e - v_i)
\]

(13)

where \( A_i = \sqrt{\text{det}(S_i)}S_i^{-1} [15,17] \) denotes the norm matrix.

The Lagrangean multiplier is adopted to include the constraints into the optimization, and the augmented objective function becomes

\[
F_m = \sum_{i=1}^c \sum_{k=1}^N (u_{ik}^p D_{ik} + u_{ik}^p \Upsilon_{ik}) + \lambda \left( 1 - \sum_{i=1}^c u_{ik} \right)
\]

(14)

Taking the derivative of \( F_m \) with respect to \( u_{ik} \) for \( p > 1 \) and equating to zero, and with the constraint \( \sum_{i=1}^c u_{ik} = 1 \), we get

\[
u_{ik} = \frac{1}{\sum_{j=1}^c \left( \frac{D_{ik} + \Upsilon_{ik}}{D_{jk} + \Upsilon_{jk}} \right)^{1/(p-1)}}
\]

(15)

For a positive definite matrix \( L \), for any vector \( x \), we know that

\[
\frac{\partial}{\partial x} (x^T L x) = 2Lx
\]

(16)

Taking the derivative of \( F_m \) with respect to \( v_i \) and equating to zero, and using the result in Eq. (16) with the property \( A_i^{-1}A_i = I \), we have

\[
v_i = \frac{\sum_{k=1}^N u_{ik}^p (y_i - \mathbf{Q} \Phi(p_k) + e) + \frac{\alpha}{N_R} \sum_{x_i \in N_k} (y_r - \mathbf{Q} \Phi(p_r) + e)}{(1 + \alpha) \sum_{k=1}^N u_{ik}^p}
\]

(17)
Similarly, the bias field can be estimated by equating the derivative of $F_m$ with respect to $b_k$ to zero and using the matrix derivative result

$$\frac{\partial}{\partial X}(Xa + b)^TL(Xa + b) = (L + L^T)(Xa + b)a^T$$  \hspace{1cm} (18)$$
we have

$$c\sum_{i=1}^{c} N \frac{P}{Nk}(A_i + A_i^T)Q([\Phi(p_k)][\Phi(p_k)]^T)
+ \sum_{y_r \in N_k} \frac{\alpha}{N_R}([\Phi(p_r)][\Phi(p_r)]^T)

= \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^P(A_i + A_i^T)(y_k + e - v_i)[\Phi(p_k)]^T
+ \sum_{y_r \in N_k} \frac{\alpha}{N_R}([y_r + e - v_i][\Phi(p_r)]^T)$$  \hspace{1cm} (19)$$

Let

$$O_i = (A_i + A_i^T)$$  \hspace{1cm} (20)$$
$$M_i = \sum_{k=1}^{N} u_{ik}^P([\Phi(p_k)][\Phi(p_k)]^T) + \sum_{y_r \in N_k} \frac{\alpha}{N_R}([\Phi(p_r)][\Phi(p_r)]^T)$$  \hspace{1cm} (21)$$

and

$$R = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^P(A_i + A_i^T)(y_k + e - v_i)[\Phi(p_k)]^T
+ \sum_{y_r \in N_k} \frac{\alpha}{N_R}([y_r + e - v_i][\Phi(p_r)]^T)$$  \hspace{1cm} (22)$$

Eq. (19) can be written as

$$O_1 Q M_1 + O_2 Q M_2 + \cdots + O_c Q M_c = R$$  \hspace{1cm} (23)$$
which equals to

$$(O_1 \otimes M^T + O_2 \otimes M^T + \cdots + O_c \otimes M^T)Q := R :$$  \hspace{1cm} (24)$$
where $\otimes$ represents the Kronecker product in Eq. (24). The definition of ";:" in Eq. (24) is as follows: for arbitrary matrix $W$, the note “$W;$" is defined as the vector formed by concatenating all the columns of matrix $W$, e.g. if $h = W_{[m \times n]}$, then $h_{i + m(j-1)} = w_{i, j}$. Therefore, we have

$$Q := \left( \sum_{i=1}^{c} O_i \otimes M^T \right)^{-1} R$$  \hspace{1cm} (25)$$

The computation of $Q$ is fast since the dimensions of $O$ and $M$ are small. The above solutions lead to the smooth approximation of the bias field as

$$b_k(\text{smooth}) = \left[ \sum_i q_1 \phi_i(p_k), \sum_i q_2 \phi_i(p_k), \ldots, \sum_i q_n \phi_i(p_k) \right]^T = Q \Phi(p_k)$$  \hspace{1cm} (26)$$

As before, the smoothed $b_k$ should be normalized to eliminate the arbitrary value of the amplitude of $b_k$ and ascertain that the sum of $b_k$ is zero to satisfy the convergence condition of AO.

2.3. Image acquisition

For evaluation of the above algorithms on actual brain images, dual FSE MR images of the whole brain (from vertex to foramen magnum) were acquired on 13 healthy normal volunteers (3 women and 10 men; age range 23–55 years). Since the dual echo images are acquired in an interleaved manner, the images are in perfect registration with each other and do not require post-acquisition image alignment. Images were acquired either on a General Electric (1.5 T) or a Philips (3 T) scanner, with the following parameters: field-of-view of 240 mm $\times$ 240 mm, image matrix of 256 $\times$ 256, and echo train length of 8. A quadrature birdcage resonator was used both for RF transmission and signal reception at 1.5 T using the following parameters: TE1/TE2/TR = 12 ms/86 ms/6800 ms, where TE and TR represent the echo and repetition times, respectively. A total of 42 contiguous and interleaved slices, each of 3 mm thick, were acquired. On the Philips 3 T Intera scanner a six channel SENSE coil was used for signal reception while the whole body coil was used for RF transmission. A SENSE factor of 2 was used for these scans. MR images were acquired with the following scan parameters: TE1/TE2/TR = 9.5 ms/90 ms/6800 ms. The total number of slices at 3 T was 44, each of 3 mm thick.

Prior to segmentation, the extrameningeal tissues from the images were removed using a semi-automatic procedure that is described elsewhere [12,29] and these stripped brain images were used as the input to the algorithms. The output of the algorithms included inhomogeneity corrected images, cluster centers, bias field, and memberships of the image volume. All the proposed algorithms were implemented under Interactive Data Language (IDL) environment in Windows.

2.4. Evaluation

Initially, the performance of the GFCM algorithm was evaluated quantitatively using the BrainWeb images [11,18]. The BrainWeb images consist of 3 mm thick normal proton density (PD) and T2 weighted images with 3% noise and 40% inhomogeneity added. We assumed the number of clusters to be four and our results suggest this number to be appropriate (see Section 3 for the rationale for using four clusters). For the BrainWeb images these four clusters represent WM, GM, and CSF, and dura matter that is consistent with Suckling et al. [30]. This is different from most of the studies performed using the single-channel images (such as T1 weighted) where usually only three classes—WM, GM, and CSF—are included [1,20,23,34].

The convergence of clustering iterations is controlled by the $L_2$ norm of the cluster center’s difference between two consecutive iterations (denoted by $\varepsilon$). Two values for $\varepsilon$ were used: $\varepsilon = 0.05$ and 0.01. At least visually, the clustering results for the two $\varepsilon$ values were comparable, as assessed by an expert, but the computation cost for the smaller value of $\varepsilon$ was almost twice. For instance, the computational time for AFCM was around
1 min for $\varepsilon = 0.05$, but around 2 min for $\varepsilon = 0.01$. Therefore, the value of $\varepsilon = 0.05$ was chosen to generate the initial inputs to the G–K algorithm using the AFCM. Other parameters that control the contextual constraints are indicated at the relevant places. All the segmentation (clustering) was performed in 2D feature space.

For quantitative comparison of segmentation based on AFCM and G–K algorithms, we compared the tissue volumes (Seg) based on the segmentation of the BrainWeb images, with the reference volumes (Ref) generated by the ground truth (using the crisp data) and computing the four similarity measures defined in Eqs. (27–30) [2]. In these equations, POE, PUE, PCE refer to the over-, under-, and the correctly estimated percentage of tissue volumes, respectively, and SI is the similarity index.

\[
\text{SI} = \frac{2 \times (\text{Ref} \cap \text{Seg})}{\text{Ref} + \text{Seg}} \times 100
\]

\[
\text{POE} = \frac{\text{Ref} \cap \text{Seg}}{\text{Ref}} \times 100
\]

\[
\text{PUE} = \frac{\text{Ref} \cap \text{Seg}}{\text{Seg}} \times 100
\]

\[
\text{PCE} = \frac{\text{Ref} \cap \text{Seg}}{\text{Ref}} \times 100
\]

The performance of the algorithm was also evaluated using the above procedure on the FSE images of normal human brain. In this case, the segmentation by an expert neurologist (more than 20 years of experience in MRI of MS) was considered as the ground truth. The segmentation by the expert was based on automatic segmentation using the method proposed by Sajja et al. [29], followed by manual validation. For the FSE images these four classes primarily represent WM, GM, CSF, and GM + CSF. Without the inclusion of GM + CSF as a separate cluster, the large spread between GM and CSF in the feature space led to less favorable classification results.

3. Results and discussion

The BrainWeb images were segmented into four tissue classes: WM, GM, CSF, and dura matter. The intensity inhomogeneity correction was performed iteratively along with the classification of tissues as described in Section 2. As mentioned earlier, the BrainWeb images were classified using the neighborhood contextual constraint based on intensity [1]. In these studies $\alpha$ was set to 0.01 to balance between the ability in overcoming noise and the effect of blurring the segmentation; the power of the membership, $p$, was set to 2 for all studies. For the contextual constraints in 3D the direct neighborhood was set to 26. We observed that the inclusion of neighborhood contextual constraints in the AO iterations did slow down the process significantly. Therefore, the contextual constraints were applied only when the clustering without contextual constraints was completed. The iteration with contextual constraints required only limited number of loops (one loop in our studies).

As an example, images segmented using the GFCM along with the corresponding BrainWeb image at one slice location are shown in Fig. 1. As can be seen from these images, GFCM, at least visually, has provided quite satisfactory results. Table 1 summarizes the quantitative evaluation of AFCM and GFCM based on the similarity measures for all the segmented tissues. The performance of GFCM can be appreciated by the high values of the similarity index. A comparison of the results in Table 1 clearly shows that G–K algorithm improved all the similarity measures.

| Similarity measures based on AFCM and GFCM without and with neighborhood contextual constraints ($\alpha = 0.01$, one loop in iterations) on the Brain Web images |
|-----------------|-----------------|-----------------|-----------------|
|                 | AFCM            | GFCM            | GFCM + $\alpha$ |
| GM              | WM              | CSF             | GM              | WM              | CSF             | GM              | WM              | CSF             |
| SI              | 89.7275         | 75.5651         | 73.3794         | 89.6643         | 88.7303         | 97.7046         | 91.2838         | 90.6280         | 98.8351         |
| PUE             | 13.2976         | 33.1987         | 0.2452          | 13.9910         | 8.3321          | 0.2174          | 13.1386         | 4.7804          | 0.3616          |
| PCE             | 86.7024         | 66.8013         | 99.7548         | 86.0090         | 91.6679         | 99.7826         | 86.8614         | 95.2196         | 99.6384         |

The similarity measures are expressed as %.
The FSE images were classified into four classes: GM, WM, CSF and GM + CSF. The use of four clusters in the FSE images (GM, WM, CSF, GM + CSF) can also be rationalized by the fact that the proximity of CSF and GM, particularly in the cortex, often results in significant overlap between these two tissues. As an example Fig. 2 shows the segmentation (membership) of one section of brain in the cerebellar region. It is known that on the FSE images the tissue intensities in the cerebellar regions are different from those of the superior parts of the brain [30]. These intensity differences cannot be corrected by merely applying the inhomogeneity correction. Therefore, the cerebellar and the posterior fossa regions are very difficult to segment using automatic techniques [29,30,34] and the segmentation of the cerebellum area is usually performed by adopting regional or localized methods [29,30,34]. Thus, it is gratifying that segmentation of the cerebellar region has been considerably improved with GFCM.

Inclusion of the G–K algorithm to account for the non-spherical occupancy of the feature space is a major component of GFCM. Therefore, we quantitatively evaluated the importance of G–K algorithm by segmenting the Brain Web and FSE images with and without incorporation of the G–K algo-

![Fig. 2. Axial FSE images in the cerebellum region (a) and (b), AFCM membership (c)–(f) and GFCM membership (g)–(j). Classification is based on WM, GM, GM + CSF, and CSF, respectively.](image)

![Fig. 3. Effect of the inclusion of contextual constraints on the GFCM classification. The PD and T2 Brain Web images are shown in (a) and (b), respectively. Images (c) and (d) demonstrate classification before and after the application of the neighborhood intensity based contextual constraint, respectively.](image)

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similarity Index (SI, expressed as %) calculated on 13 FSE images</td>
</tr>
<tr>
<td>AFCM</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>GM</td>
</tr>
<tr>
<td>77.4859</td>
</tr>
<tr>
<td>78.1612</td>
</tr>
<tr>
<td>78.5093</td>
</tr>
<tr>
<td>82.1223</td>
</tr>
<tr>
<td>81.3736</td>
</tr>
<tr>
<td>76.7665</td>
</tr>
<tr>
<td>82.2208</td>
</tr>
<tr>
<td>82.5891</td>
</tr>
<tr>
<td>77.3762</td>
</tr>
<tr>
<td>84.2675</td>
</tr>
<tr>
<td>86.1678</td>
</tr>
<tr>
<td>84.5455</td>
</tr>
<tr>
<td>SI</td>
</tr>
<tr>
<td>Mean</td>
</tr>
</tbody>
</table>

The SI values were calculated for both AFCM and GFCM without neighborhood contextual constraints.
Table 3
Percentage of overestimated (POE, expressed as %) tissue volumes calculated on 13 FSE images

<table>
<thead>
<tr>
<th></th>
<th>AFCM</th>
<th>GFCM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GM</td>
<td>WM</td>
</tr>
<tr>
<td>POE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.9762</td>
<td>17.8663</td>
<td>23.2893</td>
</tr>
<tr>
<td>10.4612</td>
<td>19.3913</td>
<td>41.0715</td>
</tr>
<tr>
<td>15.2508</td>
<td>17.0538</td>
<td>49.1291</td>
</tr>
<tr>
<td>19.5846</td>
<td>10.8091</td>
<td>22.7711</td>
</tr>
<tr>
<td>20.3589</td>
<td>19.5538</td>
<td>18.9921</td>
</tr>
<tr>
<td>11.8643</td>
<td>11.0086</td>
<td>80.251</td>
</tr>
<tr>
<td>12.0159</td>
<td>15.6141</td>
<td>23.1833</td>
</tr>
<tr>
<td>Mean</td>
<td>14.26744</td>
<td>16.31853</td>
</tr>
</tbody>
</table>

The POE values were calculated for both AFCM and GFCM without neighborhood contextual constraints.

Table 4
Percentage of underestimated (PUE, expressed as %) tissue volumes calculated on 13 FSE images

<table>
<thead>
<tr>
<th></th>
<th>AFCM</th>
<th>GFCM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GM</td>
<td>WM</td>
</tr>
<tr>
<td>PUE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.5273</td>
<td>10.8044</td>
<td>34.8732</td>
</tr>
<tr>
<td>22.5141</td>
<td>5.46488</td>
<td>42.0648</td>
</tr>
<tr>
<td>21.4907</td>
<td>11.6368</td>
<td>39.1233</td>
</tr>
<tr>
<td>18.6264</td>
<td>9.75982</td>
<td>28.9487</td>
</tr>
<tr>
<td>17.7792</td>
<td>10.2055</td>
<td>25.4576</td>
</tr>
<tr>
<td>22.6238</td>
<td>8.26889</td>
<td>35.9584</td>
</tr>
<tr>
<td>15.7325</td>
<td>4.54785</td>
<td>37.8708</td>
</tr>
<tr>
<td>13.8322</td>
<td>5.6788</td>
<td>37.9133</td>
</tr>
<tr>
<td>15.4545</td>
<td>5.55601</td>
<td>30.8017</td>
</tr>
<tr>
<td>Mean</td>
<td>19.14935</td>
<td>8.24155</td>
</tr>
<tr>
<td>STD</td>
<td>3.096724</td>
<td>2.666082</td>
</tr>
</tbody>
</table>

The PUE values were calculated for both AFCM and GFCM without neighborhood contextual constraints.

Table 5
Percentage of correctly estimated (PCE, expressed as %) tissue volumes calculated on 13 FSE images

<table>
<thead>
<tr>
<th></th>
<th>AFCM</th>
<th>GFCM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GM</td>
<td>WM</td>
</tr>
<tr>
<td>PCE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79.4727</td>
<td>89.1956</td>
<td>65.1268</td>
</tr>
<tr>
<td>77.4859</td>
<td>94.5355</td>
<td>57.9352</td>
</tr>
<tr>
<td>78.1612</td>
<td>93.7035</td>
<td>67.5857</td>
</tr>
<tr>
<td>78.5093</td>
<td>88.3633</td>
<td>60.8767</td>
</tr>
<tr>
<td>82.1223</td>
<td>90.9386</td>
<td>73.9443</td>
</tr>
<tr>
<td>81.3736</td>
<td>90.2402</td>
<td>71.0513</td>
</tr>
<tr>
<td>76.7665</td>
<td>92.876</td>
<td>73.6889</td>
</tr>
<tr>
<td>82.2208</td>
<td>89.7945</td>
<td>74.5424</td>
</tr>
<tr>
<td>82.5891</td>
<td>87.2643</td>
<td>60.275</td>
</tr>
<tr>
<td>77.3762</td>
<td>91.7311</td>
<td>64.0416</td>
</tr>
<tr>
<td>84.2675</td>
<td>95.4521</td>
<td>62.1292</td>
</tr>
<tr>
<td>86.1678</td>
<td>94.3212</td>
<td>62.0867</td>
</tr>
<tr>
<td>Mean</td>
<td>84.6939</td>
<td>92.3825</td>
</tr>
<tr>
<td>STD</td>
<td>89.1956</td>
<td>65.1268</td>
</tr>
</tbody>
</table>
The PCE values were calculated for both AFCM and GFCM without neighborhood contextual constraints. The algorithm and comparing the results with the ground truth using the similarity measures described above. The improvement in the tissue segmentation by including the G–K algorithm was also visually evaluated on the FSE images. Fig. 3 shows the segmentation of Brain Web images with and without the contextual constraints for G–K algorithm, and Table 1 summarizes the quantitative comparisons. After the G–K algorithm without the contextual constraints was completed, only a single itera-

![Graphs](image_url)

Fig. 4. Values of SI (a–c) and PCE (d–f) for GM (a and d), WM (b and e) and CSF (c and f) for the 13 FSE images with $\alpha = 0.01$, $\alpha = 0.001$, $\alpha = 0.005$, and $\alpha = 0.0005$. 

Table 5 (Continued)

<table>
<thead>
<tr>
<th></th>
<th>AFCM</th>
<th></th>
<th></th>
<th></th>
<th>GFM</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>80.85065</td>
<td>91.75845</td>
<td>66.34478</td>
<td>90.62436</td>
<td>93.01898</td>
<td>75.58457</td>
<td>84.0477</td>
<td></td>
</tr>
<tr>
<td>STD</td>
<td>3.096724</td>
<td>2.666064</td>
<td>5.707285</td>
<td>3.787879</td>
<td>3.519487</td>
<td>8.88022</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
tion was performed following the application of the contextual constraints. It is important to point out that the inclusion of neighborhood intensity based contextual constraint places increased computational burden. Further optimization of $\alpha$ and number of iterations is necessary for improving the performance of contextual constraint.

The same iteration stop criteria and contextual constraints used for the Brain Web images were employed for the FSE images. It should be pointed out that the FSE images were hard segmented to include only the three pure classes, GM, WM, and CSF (through winner takes all on the membership). Inclusion of three pure classes is necessary since the mixed class cannot be delineated by the human expert. Tables 2–5 summarize the SI, POE, PUE and PCE for the FSE images on 13 normal volunteers. Fig. 4 shows the similarity measures (SI and PCE) for all the 13 FSE images with contextual constraints for $\alpha = 0.01$.

Fig. 5. Typical FSE volume images (PD (a) and T2 (b) and segmented images using AFCM (c), GFCM (d), and GFCM with contextual constraint ($\alpha = 0.01$) (e). In the segmented images, dark, grey, and bright represent GM, WM, and CSF, respectively.
0.001, 0.005, and 0.0005. It can be seen from these plots that the results are relatively insensitive to the value of $\alpha$ over a wide range. Because of the space limitations, we included only SI and PCE, the two more important similarity measures, in these plots.

Fig. 5 shows the PD and T2 weighted FSE images at different locations. The corresponding segmented images with AFCM, GFCM, and GFCM with contextual constraint ($\alpha = 0.01$) are shown in Fig. 5c–e, respectively. Consistent with the quantitative analysis based on the similarity measures, these images demonstrate that visually GFCM performed better than AFCM, and GFCM with and without contextual constraint provided comparable results.

It is difficult to quantitatively compare our results with other published results since very few publications evaluated the performance of the segmentation using the metrics employed in the current studies. Of the four similarity measures employed in these studies, SI is considered to be the most important metric.
Fig. 5. (Continued)

since it is a measure of the agreement between the proposed segmentation and the ground truth. Higher the SI value, better is the agreement. The high SI values (∼90%) that we achieved for segmented BrainWeb images, where the ground truth is known, suggests excellent performance of the GFCM method. The SI value for the FSE images, however, is slightly lower, particularly for the CSF. This slightly lower value perhaps is due to the fact that the ground truth is not known in this case. Our results also demonstrate the improved performance of GFCM relative to AFCM. We did not observe significant differences in the segmentation quality by including the contextual constraint. However, this conclusion is based on limited studies without complete optimization of the parameters (α and number of iterations).

In these studies we have included only normal brains. Additional dimensions (such as FLAIR images) in the feature space may be required to increase the separability of more classes such as lesions [3,13,16].
Fig. 5. (Continued)
4. Conclusions

A comprehensive GFCM technique for multi-spectral segmentation of MR images is described. This technique incorporates an efficient bias field correction along with contextual constraints. In addition, this method takes into account the non-spherical occupancy of the feature space by replacing the Euclidean distance with cluster scatter measures to define the Mahalanobis distance. Qualitative and quantitative evaluation on Brain Web and real brain images indicates that GFCM outperforms the AFCM method.

Acknowledgments

This work is supported by the National Institutes of Health grants EB002095 and 1 S10 RR19186. We acknowledge Vipulkumar Patel for acquiring the FSE images.

References


Renjie He, got his Ph.D. in Electrical Engineering with specialization in Biomedical Engineering from Tsinghua University, Beijing, China, in 1998, he is assistant professor with Department of Diagnostic and Interventional Imaging, University of Texas Medical School at Houston.

Sushmita Datta, got her Ph.D. in mathematics form Indian Institute of Technology, Kanpur, India, in 2003, she is research associate of Department of Diagnostic and Interventional Imaging, University of Texas Medical School at Houston.

Balasrinivasa Rao Sajja, got his Ph.D. in mathematics form Indian Institute of Technology, Kanpur, India, in 2002, he is assistant professor with Department of radiology, University of Nebraska Medical Center.

Ponnada Narayana, got his Ph.D. in Physics from Indian Institute of Technology, Kanpur, India, in 1969, He is professor with Department of Diagnostic and Interventional Imaging, University of Texas Medical School at Houston.