Brain magnetic resonance image segmentation using novel improvement for expectation maximizing

Mohammad A. Balafar, PhD, MS, Abdul-Rahman Ramli, PhD, MS, Syamsiah Mashohor, PhD, MS.

ABSTRACT

Objectives: To improve the quality of expectation maximizing (EM) for brain image segmentation, and to evaluate the accuracy of segmentation results.

Methods: This brain segmentation study was conducted in Universiti Putra Malaysia in Serdang, Malaysia between February and November 2010 on simulated and real images using novel improvement for EM. The EM-1 (proposed algorithm) was compared with neighborhood based extensions for fuzzy C-mean (FCM). The EM-1 was also applied to all 20 normal real MRI volumes and compared with reported results from the Internet Brain Segmentation Repository.

Results: In simulated images, the EM-1 outperforms neighborhood based extensions for FCM. The average similarity index value of the proposed algorithm for all 20 normal images is 0.802. The EM-1 produces the average Jaccard indices \( \rho \) higher than other algorithms and near to manual results. The average similarity indices \( \rho \) for EM-1 and FCM extensions (FCM with spatial information [FCM-S], Fast Generalized FCM [FGFCM]) for all 20 normal images are: EM-1=0.802, FCM-S=0.7517, enhanced FCM=0.7581, and FGFCM=0.7597.

Conclusion: Experimental results show that the proposed algorithm performs better than other studied algorithms on various noise levels in terms of similarity index, \( \rho \).

Neurosciences 2011; Vol. 16 (3): 242-247

From the Department of Computer & Communication Systems, Faculty of Engineering, University Putra Malaysia, Serdang, Selangor, Malaysia.

Received 25th October 2010. Accepted 13th February 2011.

Current address correspondence and reprint request to: Dr. Mohammad A. Balafar, Department of Computer Engineering, Faculty of Electrical & Computer Engineering, Tabriz University, Tabriz, Azerbaijane Shargi, Iran. Tel. +98 (411) 3848641. E-mail: balafarila@yahoo.com
area for a long time. There are many segmentation algorithms, but there is no generic algorithm for total successful segmentation of medical images. Many image techniques have been used for image segmentation, such as thresholding, region growing, statistical models, active control models, and clustering. The distribution of intensities in medical images is usually very complex, and therefore, the determination of a threshold for the images is difficult and therefore, thresholding methods have not been very successful with these images. Mostly, the thresholding method is combined with other methods. The region growing method extends thresholding by combining it with connectivity. This method requires seeds for each region, and has the same problem of thresholding for determining the suite threshold for homogeneity. Clustering methods are common for MRI brain segmentation. Expectation-maximization (EM) and fuzzy c-mean (FCM) are the most popular clustering algorithms. The Gaussian mixture model (GMM) is a popular segmentation method. The EM is used to estimate the parameters of this model. The FCM and EM only consider the intensity of images and in noisy images, intensity is not trustworthy. Usually, spatially adjacent pixels belong to the same cluster. Many algorithms are introduced to make FCM, and EM robust against noise, but they need to be improved. Usually, spatially adjacent pixels belong to the same cluster. Many researchers have attempted to incorporate spatial information into FCM and EM to overcome the noise problem. Zhang et al proposed a novel Gaussian hidden Markov Random Field (HMRF) model to integrate spatial information into the Gaussian model. They used a Markov Random Field-Maximum A Posteriori (MRF-MAP) approach to estimate the model solution. Recently, Tang et al proposed a neighborhood-weighted GMM to overcome misclassification on the boundaries and on inhomogeneous regions of MRI brain images with noise. In this paper, a new improvement for EM is proposed, which incorporates neighborhood information into the GMM.

**Standard Gaussian model.** The Gaussian mixture model assumes \(M\) mixed component densities (Gaussian distribution) for each pixel (voxel) with \(M\) mixing coefficients. Each component is assigned to one target class and the goal is to obtain the class probabilities of each pixel (voxel). The probability distribution of the \(j\)th component is denoted by \(p_j(x_i | \theta)\), where \(x_i\) is \(ith\) pixel in input image and \(\theta\) is the parameter (mean \(\mu_j\) and covariance matrix \(\Sigma_j\)) of component \(j\). The probability distribution for each pixel (voxel) can be described as a mixture of probability distributions as follows:

\[
p(x_i | \theta) = \sum_{j=1}^{M} \alpha_j p_j(x_i | \theta_j)
\]  

(1)

where \(\alpha_j\) denotes the mixture coefficient with the constraint, \(\sum_{j=1}^{M} \alpha_j = 1\). The probability distribution of component \(j\) is modelled by a Gaussian distribution with mean \(\mu_j\) and covariance matrix \(\Sigma_j\):

\[
p_j(x_i | \theta_j) = p_j(x_i | \mu_j, \Sigma_j) = \frac{1}{\sqrt{2\pi \det(2\pi \Sigma_j)}} e^{-(x_i - \mu_j)^T \Sigma_j^{-1} (x_i - \mu_j)/2}
\]  

(2)

Usually, maximum likelihood (ML) estimation is used to find the parameters. The log-likelihood expression for the parameter \(\theta\) and the image \(X\) is defined as follows:

\[
\log(L(\theta | X)) = \log \prod_{i=1}^{N} p(x_i | \theta) = \sum_{i=1}^{N} \log \left( \sum_{j=1}^{M} \alpha_j p_j(x_i | \theta_j) \right)
\]  

(3)

Finding the ML solution from this equation is difficult. Usually, the expectation-maximization (EM) is used to obtain the parameters. The EM steps are listed in the following:

**Step 1:** Mean and covariance matrix are initialized using \(k\)-means and prior probability is initialized uniformly.

**Step 2:** Bayes’ rule is used to obtain the probability of data \(x_i\) belong to class \(\theta_j\) (E-step):

\[
p(j | x_i, \theta^j) = \frac{\alpha_j p_i(x_i | \theta^j)}{\sum_{k=1}^{M} \alpha_k p_i(x_i | \theta^k)}
\]  

(4)

**Step 3:** Probability obtained in E-step is used to obtain the mixing coefficient, mean, and covariance matrix (M-step):

\[
\alpha_j^{t+1} = \frac{1}{N} \sum_{i=1}^{N} p(j | x_i, \theta^j)
\]  

(5)

\[
\mu_j^{t+1} = \frac{\sum_{i=1}^{N} x_i p(j | x_i, \theta^j)}{\sum_{i=1}^{N} p(j | x_i, \theta^j)}
\]  

(6)

\[
\Sigma_j^{t+1} = \frac{\sum_{i=1}^{N} p(j | x_i, \theta^j)(x_i - \mu_j^{t+1})(x_i - \mu_j^{t+1})^T}{\sum_{i=1}^{N} p(j | x_i, \theta^j)}
\]  

(7)

The EM steps are repeated until convergence.

**Methods.** This brain segmentation study was conducted in the Universiti Putra Malaysia in Serdang, Malaysia between February and November 2010 on simulated and real images using novel improvement for EM. An improvement for the GMM is introduced by incorporating neighborhood information into...
Brain segmentation ... Balafar et al

likelihood function and EM steps. In the likelihood function (equation (3)), the average of neighborhood pixels distribution is added to the distribution value of pixel $x_i$ as neighborhood information:

$$\log(L(\theta | X)) = \log \prod_{i=1}^{n} p(x_i | \theta) = \sum_{i=1}^{n} \log \sum_{\theta_j} p(x_i | \theta_j) + \sum_{i=1}^{n} \frac{1}{L} \sum_{k=1}^{L} \log p(x_i | \theta_k)$$ \hspace{1cm} (8)

Where $x_i$ represents a neighbor of pixel $x$ and $K = 1, ..., L$ denotes the set of neighbors, which are determined by a window centered on $x_i$. The $L$ is the number of distribution values for neighbor pixels. The parameter $\beta$ determines the weight of neighborhood information. Incorporating neighborhood information improves the performance of clustering methods in high levels of noise, but the blurring effect degrades the performance of them in low noise levels. In order to overcome the degrading effect of algorithms in low levels of noise, the variance of noise is used to specify the weight of neighborhood information ($\beta$). The $\beta$ value is set to $\sigma$, where $\sigma$ is the variance of noise except for 3% noise level for which the $\beta$ value is set to 0. Because in 3% noise level, the degradation effect of neighborhood incorporation is more than its improvement effect, and standard EM performs better than the improved one. An improvement for EM named EM-1 is introduced to solve likelihood function. The EM is modified as follows:

a. In equation 4, average of probability of neighborhood pixels is multiplied to the probability value of pixel $x_i$ as neighborhood information:

$$p(j | x_i, \theta') = \frac{\alpha_j^{L} E_{\theta} p_j(x_i | \theta')}{\sum_{j=1}^{L} \alpha_j^{L} E_{\theta} p_j(x_i | \theta')} \hspace{1cm} (9)$$

Where $E_{\theta}$ is the average of the probability of neighbors of $x_i$ that belong to class $\theta_j$:

$$E_{\theta} = \sum_{j=1}^{L} \frac{\sum_{r=1}^{K} p(j | x_r, \theta')}{\sum_{r=1}^{K} p(j | x_r, \theta')} \hspace{1cm} (10)$$

b. In equation 6, the average of neighborhood pixels values with a weight is added to each pixel as neighborhood information:

$$\mu_{\theta'} = \frac{1}{\sum_{j=1}^{L} \sum_{r=1}^{K} p(j | x_r, \theta')} \sum_{j=1}^{L} \sum_{r=1}^{K} \frac{1}{\beta} (\mu_j - \mu_{\theta'}) + \frac{1}{\beta} \sum_{j=1}^{L} \sum_{r=1}^{K} p(j | x_r, \theta') \mu_{\theta'} \hspace{1cm} (11)$$

c. In equation 7, the average of distance of neighbor pixel from component center is added to the distance of pixel from component center as neighborhood information:

$$\sum_{j=1}^{L} \sum_{r=1}^{K} p(j | x_r, \theta')((x_i - \mu_{\theta'})^2 + \beta \sum_{j=1}^{L} \sum_{r=1}^{K} p(j | x_r, \theta') \mu_{\theta'})$$

In an MRI image, noise behaves in a Rician distribution.\(^{21}\) Noise distribution approaches Gaussian with increasing Signal to Noise Ratio (SNR) and approaches Rayleigh with decreasing SNR. The Rician distribution in the background is Rayleigh distributed, because the signal is usually considered as zero and the probability distribution function (PDF) becomes as follows:

$$p(O_i) = \frac{O_i}{\sigma^2} e^{-\frac{(O_i^2)}{2\sigma^2}} \hspace{1cm} (13)$$

Where $O_i = 1, ..., n$ is the set of observations in the background. The variance of noise can be estimated by equation 14:\(^{21}\)

$$\sigma_{\text{Noise}}^2 = \frac{1}{2n} \sum_{i=1}^{n} O_i^2 \hspace{1cm} (14)$$

Sometimes in real images, the background has no noise. In this case, the parameter $\beta$ is considered 3 because the real images usually contain the noise.

**Results.** The proposed extension of EM (EM-1) is simulated and applied on the simulated images from BrainWeb,\(^{22}\) and real images from the Internet Brain Segmentation Repository (IBSR).\(^{23}\) The results of the proposed algorithm are compared with results for the existing extensions of EM (DPM, rjMCMC, KVL, MPM-MAP), the existing neighborhood based extension of FCM (FCM_S,\(^{25}\) FGFCM\(^{25}\)) and reported results in IBSR. The results of algorithms are compared quantitatively to analyze their performance. The neighborhood size, $N$ for the proposed algorithm is set to $3 \times 3$. The similarity index $\rho$ is used to evaluate the algorithms quantitatively. The similarity index $\rho$ is the degree of a class of pixels matching between ground truth and segmentation result. It is defined as:

$$\rho = \frac{2 | X_i + Y_i |}{| X_i | + | Y_i |} \hspace{1cm} (15)$$

where $X_i$ represents class $i$ in ground truth and $Y_i$ represents the same class in the segmentation result.

**Simulated images.** The simulated MRI images are obtained from BrainWeb. A simulated data volume with T1-weighted sequence, slice thickness of 1 mm, and a volume size of $217 \times 181 \times 181$ is used. Non-
brain tissues are removed prior to segmentation. The number of tissue classes in the segmentation is set to 3: grey matter (GM), white matter (WM), and CSF. It should be noted that background pixels are ignored in this experiment. In order to compare different clustering algorithms, they are applied on whole volume at different noise levels, and average similarity is used to compare them. The average similarity indices $\rho$ for clustering algorithms are shown in Figure 1. The EM-1 has similarity indices higher than other algorithms. The EM-1 (proposed algorithm) is compared with neighborhood based extensions for FCM. Figure 2 shows the average similarity indices $\rho$ for EM1 and FCM extensions (FCM_S and FGFCM) when they are applied on 3D volume at different noise levels. Figure 2 shows that EM-1 has the highest similarity indices. Afterwards, the effect of different neighborhood sizes on the performance of proposed algorithms is analyzed. The proposed algorithm with different neighborhood sizes is applied on 3D volume, and again the average similarity is used to analyze the effect of different neighborhood sizes on the proposed algorithm. The average similarity index $\rho$ of the proposed algorithm when different neighborhood sizes (3, 5, and 7) are applied on the simulated image with 9% noise are: 0.9085, 0.904, and 0.894.

**Real images.** The proposed algorithm is also applied on real MRI images. The real MRI images are obtained from the IBSR by the Centre for Morphometric Analysis, Massachusetts General Hospital. Twenty normal data volume with T1-weighted sequence were used.

In the IBSR, manual segmentation results are provided along with brain MRI data to encourage introducing new segmentation algorithms and evaluate their performance. Trained investigators used semi-automated histograms on the spatially normalized images to obtain manual segmentation. Prior to clustering, inhomogeneity correction along with intensity adjusting is applied to real image volumes. First, the EM-1 (proposed algorithm) is applied to slices of a real MRI volume with size 256x256x65 and the similarity index $\rho$ for each slice is presented. Figure 3 shows the similarity indices of the proposed algorithm for every slice of the MRI volume. The EM-1 (proposed algorithm) is also applied to all 20 normal real MRI volumes and compared with reported results from the IBSR. Figure 4 shows the similarity index of the
The proposed algorithm (EM-1) when it is applied on each of 20 normal image volumes. The average similarity index value of the proposed algorithm for all 20 normal images is 0.802. The reported results are based on the Jaccard index. Therefore, the average Jaccard index values of algorithms for 20 normal real MRI volumes are used to evaluate them. Figure 5 shows the average Jaccard index values of different algorithms for all 20 normal images.

Discussion. To evaluate EM-1, the method is compared with existing neighborhood based extensions for EM on the whole volume of simulated images at different noise levels. The EM-1 has similarity indices higher than other algorithms, and when the noise level is increased, its similarity indices decrease more slowly than others. The EM-1 (proposed algorithm) is also compared with neighborhood-based extensions for FCM (FCM_S and FGFCM) when they are applied on 3D simulated volume at different noise levels. The EM-1 has the highest similarity indices. Following this, the effect of different neighborhood sizes on the performance of proposed algorithms was analyzed. When the neighborhood size is increased, the similarity indices of the proposed algorithm decreases. The proposed algorithm was also applied to real MRI images obtained from the IBSR. The EM-1 (proposed algorithm) is applied to all 20, 3D normal real MRI volumes and compared with reported results from IBSR. The average similarity index value of the proposed algorithm (EM-1) for all 20 normal images is 0.802. The EM-1 produces the average Jaccard indices ρ higher than other algorithms. The EM-1 is also compared with neighborhood based extensions for FCM. The average similarity indices ρ for EM-1 and FCM extensions (FCM_S, FCM-EN, FGFCM) for all 20 normal images are: EM-1=0.802, FCM-S=0.7517, FCM-EN=0.7581, FGFCM=0.7597. The EM-1 produced the highest similarity indices.

In conclusion, an improvement for EM has been introduced. In order to overcome the problem of standard EM in the presence of noise, the introduced algorithms are formulated by modifying the equations of the standard EM algorithm, which allow the neighborhood pixels to be incorporated in the labeling of a pixel. The introduced algorithm is tested on simulated MRI images, with different noise levels, and real images. The performance of the existing neighborhood based EM, and FCM algorithms and the proposed algorithm are compared qualitatively. The similarity index, ρ of the segmentation results is used to evaluate different algorithms. Experimental results show that the proposed algorithm performs better than other studied algorithms on various noise levels in terms of the similarity index, ρ.

In the future, we will consider undertaking research on other kinds of clustering methods to improve their functionality. Also, we will analyze the effects of different clustering methods in segmentation of medical images for the diagnosis of abnormal or various important matters in medical images.

References

Brain segmentation ... Balafar et al


23. MGH CMA. IBSR: The Internet Brain Segmentation Repository [Internet]. National Institute of Science: Harvard Medical School; 1997 [updated 2009 January 15; Cited 2010 October 1]; Available from URL: http://www.cma.mgh.harvard.edu/ibsr/.


**ETHICAL CONSENT**

All manuscripts reporting the results of experimental investigations involving human subjects should include a statement confirming that informed consent was obtained from each subject or subject’s guardian, after receiving approval of the experimental protocol by a local human ethics committee, or institutional review board. When reporting experiments on animals, authors should indicate whether the institutional and national guide for the care and use of laboratory animals was followed. Research papers not involving human or animal studies should also include a statement that approval/no objection for the study protocol was obtained from the institutional review board, or research ethics committee.