An Integrated Method of Adaptive Enhancement for Unsupervised Segmentation of MRI Brain Images [1]

Jing-Hao XUE [a], Aleksandra PIZURICA [a], Wilfried PHILIPS [a,*],
Etienne KERRE [b], Rik VAN DE WALLE [c], Ignace LEMAHIEU [c]

[a]*Dept. of Telecommunications and Information Processing, Ghent University, Belgium*

[b]*Dept. of Applied Mathematics and Computer Science, Ghent University, Belgium*

[c]*Dept. of Electronics and Information Systems, Ghent University, Belgium*

Abstract

This paper presents an integrated method of the adaptive enhancement for an unsupervised global-to-local segmentation of brain tissues in three-dimensional (3-D) MRI (Magnetic Resonance Imaging) images. Three brain tissues are of interest: CSF (CerebroSpinal Fluid), GM (Gray Matter), WM (White Matter). Firstly, we de-noise the images using a newly proposed versatile wavelet-based filter, and segment the images with minimum error global thresholding. Subsequently, we combine a spatial-feature-based FCM (Fuzzy C-Means) clustering with 3-D clustering-result-weighted median and average filters, so as to further achieve a locally adaptive enhancement and segmentation. This integrated strategy yields a robust and accurate segmentation, particularly in noisy images. The performance of the proposed method is validated by four indices on MRI brain phantom images and on real MRI images.

Key words: Adaptive enhancement, Unsupervised segmentation, Brain tissues, MRI.

1 Introduction

Segmentation of brain tissues in MRI (Magnetic Resonance Imaging) images plays a crucial role in three-dimensional (3-D) volume visualization, quantitative morphometric analysis and structure-function mapping for both scientific and clinical investigations. For instance, in order to be able to combine EEG (ElectroEncephaloGram) data and MRI images for the localization of epileptic sources within the brain, an anatomic head model is required; this model describes the brain in terms of segments of CSF (CerebroSpinal Fluid), GM (Gray Matter), WM (White Matter), skull and scalp which have significantly different electric conductivities (Van Hoey et al., 2000). In this paper, we deal with the segmentation of CSF, GM and WM in MRI brain images.

Numerous MRI segmentation methods have been reported (Bezdek et al., 1993; Zijdenbos and Dawant, 1994; Clarke et al., 1995; Niessen et al., 1999; Pham et al., 2000; Ruan et al., 2000; Xue et al., 2001; Ruan et al., 2002). Niessen et al. (1999) roughly grouped these methods into three main categories: classifica-

tion methods, region-based methods and boundary-based methods. Just as pointed out in Niessen et al. (1999), the methods in the first two categories are limited by the difficulties due to intensity inhomogeneities, partial volume effects and susceptibility artifacts, while those in the last category suffer from spurious edges. Furthermore, all methods are degraded by noise perturbations in low contrast and low SNR (Signal-to-Noise Ratio) images, e.g., the images used in EEG/MRI analysis where the slices are thin and the measuring time is short.

In this context, we propose an integrated method to achieve an adaptive enhancement for the unsupervised global-to-local segmentation of CSF, GM and WM. In our method, a region-based global algorithm (minimum error thresholding) and an unsupervised local classification algorithm (Fuzzy C-Means clustering) are used for segmentation. In order to remove noise and artifacts, a versatile filter (Pizurica et al., 2003) based on wavelet domain techniques, and locally adaptive 3-D weighted median and average filters based on clustering results are also proposed and embedded into our method. Only single-channel (T1-weighted) MRI images are addressed. In this paper, we do not pay much attention to the image registration. Nevertheless, the proposed method can be extended to work on registered multiple pulse sequences, like T1-, T2- and Proton-Density-weighted MRI images.

The contribution of this paper is the integration of locally adaptive image enhancement and global-to-local segmentation in a 3-D framework, which achieves a more robust and accurate segmentation.

This paper is organized as follows. Section 2 outlines the proposed method. Section 3 presents the versatile wavelet-based de-noising algorithm. Section 4 and 5 describe the minimum error thresholding and the FCM (Fuzzy C-Means) clustering based on a feature space of pairs (intensity, 3-D locally averaged intensity), respectively. The clustering-based locally adaptive enhancement scheme is proposed in Section 6. Section 7 validates our proposed method with an MRI brain phantom and real images. Section 8 gives a summary and makes conclusions.

2 Outline of Proposed Integrated Method

First, we de-noise the images using the versatile wavelet-based filter. Second, we segment the images with minimum error global thresholding. Third, we classify the voxels (counterpart of pixels in a 3-D volume) into three brain tissues through FCM clustering, using the global thresholding result to initialize the FCM parameters. The feature space is constructed by intensity pairs (intensity, 3-D locally averaged intensity) associated with each of the voxels in the MRI images. Subsequently, we further enhance the images with locally adaptive weighted median and average filters; the elements of the 3-D filtering templates are selected according to the clustering result and weighted by the fuzzy membership degrees. Finally, we employ the FCM clustering once more to achieve a more robust and accurate segmentation.

3 Versatile Wavelet-based De-noising

In medical image enhancement, a trade off between noise reduction and the preservation of actual image features has to be made in a way that enhances the diagnostically relevant image content. To achieve a good performance in this respect, a de-noising algorithm has to adapt to image discontinuities. The wavelet representation naturally facilitates the construction of such spatially adaptive algorithms. It compresses the essential information in an image into relatively few, large coefficients; these large coefficients correspond to the main image details at different resolution scales. Thanks to this property, additive noise can be effectively suppressed even by simple thresholding (Donoho and Johnstone, 1995) of the wavelet coefficients.

In a wavelet decomposition of an image (Mallat, 1999), a wavelet coefficient $w_{k,j}^D$ represents its band-pass content at resolution scale 2^j $(1 \leq j \leq J)$, spatial position [2] k and orientation D. The lowpass

---

[2] For notational simplicity, we number each voxel with a single index k in this paper.

2

image content is represented by scaling coefficients $u_{k,J}$. Typically, three orientation subbands are used: $D \in \{LH, HL, HH\}$, leading to three detail images at each scale, characterized by horizontal, vertical and diagonal directions.

Here we apply a versatile and spatially adaptive wavelet-based de-noising algorithm (Pizurica et al., 2003) which is useful to medical images with diverse types of noise, and is of low-complexity, both in its implementation and execution time. The algorithm uses a non-decimated wavelet transform.

Assume the following noise model $w_k = y_k + n_k$, where y_k is the unknown noise-free wavelet coefficient [3] and n_k an arbitrary noise contribution. Let X_k denotes a random variable, which takes values x_k from the binary label set $\{0, 1\}$. The label value $x_k = 1$ denotes that the wavelet coefficient w_k represents a "signal of interest", and the opposite label value $x_k = 0$ denotes that w_k is dominated by noise. Starting from this model, it was motivated in Pizurica (2002) to estimate the noise-free coefficients as:

$$\hat{y}_k = \frac{r \xi_k \eta_k}{1 + r \xi_k \eta_k} w_k, \qquad (1)$$

with $\xi_k = \frac{p_{M_k|X_k}(m_k|1)}{p_{M_k|X_k}(m_k|0)}$, $\eta_k = \frac{p_{E_k|X_k}(e_k|1)}{p_{E_k|X_k}(e_k|0)}$, $r = \frac{P(X_k=1)}{P(X_k=0)}$, where m_k is the coefficient magnitude $m_k = |w_k|$ and e_k is a local spatial activity indicator (LSAI), which is a function of the surrounding wavelet coefficients. In particular, the LSAI is here the averaged energy of the neighboring coefficients of w_k, where the neighbors are the surrounding coefficients in a square window at the same scale and the "parent" coefficient from the first coarser scale. The probability $P(X_k = 1)$ is the fraction of the wavelet coefficients in a given subband, which represent the signal of interest. The estimator (1) thus shrinks thus each wavelet coefficient according to the following three sources of information: (i) the coefficient magnitude; (ii) the locally averaged magnitude within a given window and (iii) the global coefficients distribution.

---

[3] Whenever there can be no confusion, we omit the indices of the wavelet coefficients that denote the scale and the orientation.

The probabilities and probability density functions used above are estimated directly from the observed image using a preliminary coefficient classification which is detailed in Pizurica et al. (2003) and Pizurica (2002). The classification relies on the persistence of significant image features across the resolution scales. This preliminary coefficient classification is used to empirically estimate the statistical distributions of the coefficients that represent useful image features on the one hand and mainly noise on the other. The adaptation to the spatial context in the image is achieved by using a wavelet domain indicator of the local spatial activity.

3.1 Implementation of De-noising MRI Images

Noise in the MRI magnitude images is Rician, having a signal dependent mean, and the Rician distribution approaches a Gaussian distribution when the SNR is high (Nowak, 1999). In Nowak (1999), it was noted that, due to the signal-dependent mean of the Rician noise, both wavelet and scaling coefficients of noisy MRI images are *biased* estimates of their noise-free counterparts. It was also shown that one can efficiently overcome this problem by filtering the *square* of the MRI magnitude images in the wavelet domain. In the squared magnitude images, data are non-central chi-square distributed, and the wavelet coefficients are no longer biased estimates of their noise-free counterparts. The bias still remains in the scaling coefficients, but is not signal-dependent and it can be removed easily (Nowak, 1999): at the resolution scale 2^j, from each scaling coefficient $2^{j+1}\sigma_c$ should be subtracted, where σ_c^2 is the underlying complex Gaussian noise variance.

The complete de-noising algorithm we apply is thus:
Step 1: Compute the square of the MRI magnitude image;
Step 2: Compute the non-decimated wavelet transform with J decomposition levels (in practice, we used J=4);
Step 3: Estimate the wavelet coefficients using estimator (1);
Step 4: Subtract $2^{J+1}\sigma_c$ from the scaling coefficients;
Step 5: Apply the inverse wavelet transform;

3

Step 6: Compute the square root of the image.

This de-noising algorithm operates on 2-D images, and we apply it to 3-D MRI volumes in a slice-by-slice manner. With this algorithm, we can improve the SNR fastly and effectively; furthermore, the distribution of the remaining noise is approximated more precisely by a Gaussian distribution. Furthermore, higher SNR will also give rise to a better segmentation in further FCM clustering, considering more compact and well-separated intensities of brain tissues.

To illustrate the de-noising algorithm, Fig. 1(a) shows an MRI image simulated from a normal brain phantom (Kwan et al., 1999) with 9% noise level and 40% intensity inhomogeneities; Fig. 1(b) shows the corresponding wavelet-based de-noising result.

4 Minimum Error Thresholding

In our integrated global-to-local method, we use FCM clustering (Bezdek, 1981) to achieve spatially adaptive segmentation. FCM has been applied widely to MRI segmentation (Bezdek et al., 1993; Clarke et al., 1995; Clark et al., 1994), and regarded as one of the most promising methods (Clarke et al., 1995). As an unsupervised clustering method, the performance of FCM, particularly its validity and speed of convergence, depends on the initialization of its parameters, e.g., the centers (prototypes) of clusters and the membership degrees of each voxel to different clusters.

Practically, due to the great diversity of MRI images from different subjects and imaging settings, and due to an aim of reducing the human interactivity in favor of a less labor-intensive and fast segmentation, little prior knowledge of the parameters of FCM is available. However, these parameters can be estimated from an initial segmentation which constructs automatically the training set of classified voxels in the original images.

In order to generate a fast and reliable estimator of the FCM's parameters, intensity-based thresholding

is used in our work.

Intensity-based thresholding is the oldest and most widely used segmentation algorithm (Sahoo et al., 1988; Pal and Pal, 1993; Zhang, 1997), simple and fast. Since MRI data, which normally has a Rician distribution, tends to have a more Gaussian-like distribution after wavelet de-noising, we chose minimum error thresholding as proposed by Kittler and Illingworth (1986) based on Gaussian distributions of object and background intensities.

Although this method was originally proposed for binarization, it can be extended to segment three clusters, i.e., CSF, GM and WM. In T1-weighted MRI images, the black non-head background is always set to zero intensity and ignored in further processing; WM is indeed brighter than GM, and CSF is the darkest one among these three tissues.

Consider a threshold pair $\mathbf{t} = (t_1, t_2)$, $1 < t_1 < t_2 < f_{max}$, where f_{max} is the maximum intensity in a 3-D MRI volume. We use this pair to classify all the voxels having an intensity f into a cluster c: c is CSF (if $f < t_1$), GM (if $t_1 \leq f < t_2$) or WM (if $f \geq t_2$).

Consider a probability density $p(f)$ of the intensity f, the prior probability $P(c|\mathbf{t})$ of the cluster c, and a conditional density $p(f|c, \mathbf{t})$ of f given the cluster c under the threshold pair $\mathbf{t}$. One criterion of an optimal classification is to maximize a sum $\sum_c \sum_{f \in c} [p(f)P(c|\mathbf{t})p(f|c, \mathbf{t})]$ (Kittler and Illingworth, 1986). Suppose that we approximate $p(f)$ by the intensity histogram $h(f)$ of the MRI volume, and approximate $p(f|c, \mathbf{t})$ by three Gaussian distributions $N(\mu_{c,\mathbf{t}}, \sigma_{c,\mathbf{t}}^2)$ for c as being CSF, GM or WM respectively. In that case, the mentioned criterion can be viewed as a measure of fitting between the observed intensity distribution and the Gaussian approximation.

We seek the pair $\mathbf{t}$ corresponding to the maximum fitting. An alternative criterion can be obtained by minimizing $\sum_c \sum_{f \in c} [p(f)(-2) \ln(P(c|\mathbf{t})p(f|c, \mathbf{t}))]$ so as to simplify the computation. As mentioned above, we assume $p(f) = h(f)$ and $p(f|c, \mathbf{t}) \sim N(\mu_{c,\mathbf{t}}, \sigma_{c,\mathbf{t}}^2)$. Meanwhile, $P(c|\mathbf{t})$, $\mu_{c,\mathbf{t}}$ and $\sigma_{c,\mathbf{t}}$ can be estimated from the histogram $h(f)$

as $P(c|\mathbf{t}) = \sum_{f \in c} h(f)$, $\mu_{c,\mathbf{t}} = \frac{1}{P(c|\mathbf{t})} \sum_{f \in c} h(f)f$, $\sigma_{c,\mathbf{t}}^2 = \frac{1}{P(c|\mathbf{t})} \sum_{f \in c} h(f)f^2 - \mu_{c,\mathbf{t}}^2$. Consequently, the objective function can be simplified and rewritten as (Kittler and Illingworth, 1986)

$$J(\mathbf{t}) = \sum_c [P(c|\mathbf{t}) \ln \sigma_{c,\mathbf{t}} - P(c|\mathbf{t}) \ln P(c|\mathbf{t})] . \quad (2)$$

The global threshold pair can be found by minimizing $J(\mathbf{t})$ in equation (2) (Kittler and Illingworth, 1986; Ye and Danielsson, 1988), iteratively or exhaustively. The thresholding result of Fig. 1(a) is shown in Fig. 1(c); it will be used to initialize the parameters of the following FCM clustering.

5   Segmentation with FCM

Consider a dataset $X = \{\mathbf{x}_k\}_{k=1}^n \subset \mathbb{R}^q$ where q is the dimension of the desired feature space, $\mathbf{x}_k$ denotes the feature vector of k-th voxel; furthermore, consider a set of fuzzy clusters $\{F_i\}_{i=1}^C$ in $\mathbb{R}^q$ with its corresponding crisp version $\{H_i\}_{i=1}^C$.

Given a fuzzy cluster F_i, FCM assigns to each voxel $\mathbf{x}_k$ in the dataset X a degree of membership to the cluster F_i which is denoted as $u_i(\mathbf{x}_k)$ (hereinafter abbreviated as u_{ik}). $u_{ik} \in [0,1]$, and $\sum_{i=1}^C u_{ik} = 1$, $\forall k \in \{1, \cdots, n\}$.

The optimal assignment is accessed via minimizing approximately the sum of intra-cluster squared errors as

$$J_{FCM}(U, V : X) = \sum_{i=1}^C \sum_{k=1}^n (u_{ik})^m (\|\mathbf{x}_k - \mathbf{v}_i\|)^2, \quad (3)$$

where U is the matrix $[u_{ik}]_{C \times n}$, and V is the set $\{\mathbf{v}_i\}_{i=1}^C$. Furthermore, $\mathbf{v}_i \in \mathbb{R}^q$ is the prototype of i-th cluster (i.e., the prototype of F_i). $m \geq 1$ is a weighting exponent which determines the degree of fuzziness of FCM. $\|\mathbf{x}_k - \mathbf{v}_i\|$ is an inner product induced norm on $\mathbb{R}^q$ to measure the distance from $\mathbf{x}_k$ to $\mathbf{v}_i$; we use the Euclidean norm and set $m = 2$.

In order to bring local spatial neighborhood information from the images into the FCM clustering, we compute a feature vector (intensity, 3-D locally averaged intensity), which has been employed in 2-D

image thresholding techniques (Sahoo et al., 1988), for each voxel, i.e., $\mathbf{x}_k = (f_k, \bar{f}_k)' \in \mathbb{R}^2$ where f_k is the intensity of k-th voxel, and $\bar{f}_k$ the corresponding local average. The latter can be obtained by applying any averaging template; we choose a simple average over a 6-connected spatial neighborhood considering 3-D connectivities. Generally the voxels inside a tissue are located closer than the edge voxels to an equi-value curve representing $f_k = \bar{f}_k, \forall k \in \{1, \cdots, n\}$.

In our case three clusters $\{F_i\}_{i=1}^3$ are defined respectively for CSF, GM and WM, so $C = 3$. We implement FCM to segment brain tissues as follows.

Step 1: Initialize U using the minimum error thresholding result as follows: $\forall k$, if the k-th voxel (corresponding to $\mathbf{x}_k$) is segmented into H_i, set $u_{ik} = 1$; otherwise set $u_{ik} = 0$.

Step 2: Initialize all the elements in the set V of prototypes as follows: $\forall i$,

$$\mathbf{v}_i = \sum_{k=1}^n (u_{ik})^m \mathbf{x}_k \Big/ \sum_{k=1}^n (u_{ik})^m , \quad (4)$$

and set $\mathbf{v}_i^{(0)} = \mathbf{v}_i$. $m = 2$ is the weighting exponent in equation (3).

Step 3: Update U as follows: first, for every voxel $\mathbf{x}_k$, count the number β_k of prototypes which satisfy $\|\mathbf{x}_k - \mathbf{v}_i\| = 0$ for any i, Θ_k denotes the set of these prototypes; usually β_k equals 0 or 1, if no two clusters share the same prototype. Second, if $\beta_k \neq 0$, then assign $\mathbf{x}_k$ an average membership degree to the clusters with their prototypes in Θ_k, i.e., $\forall \mathbf{v}_i \in \Theta_k$, set $u_{ik} = 1/\beta_k$, and $\forall \mathbf{v}_i \notin \Theta_k$, set $u_{ik} = 0$; otherwise, from the minimization of equation (3), we can update the memberships as (Bezdek, 1981) $u_{ik} = \left[\sum_{j=1}^C (\|\mathbf{x}_k - \mathbf{v}_i\|/\|\mathbf{x}_k - \mathbf{v}_j\|)^{\frac{2}{m-1}} \right]^{-1}$, $\forall i, k$.

Step 4: Update V again with equation (4).

Step 5: Stop the FCM iteration if $\max_{i \in [1,C]} \|\mathbf{v}_i - \mathbf{v}_i^{(0)}\|_\infty < \epsilon$, where $\epsilon = 0.5$ for sub-level precision. Otherwise, set $\mathbf{v}_i^{(0)} = \mathbf{v}_i$, then return to Step 3. $\|\cdot\|_\infty$ is the L-infinity vector norm equal to the maximum of the absolute values of the entries in the corresponding vector. Similar results can be obtained using other equivalent norms.

Step 6: Segment X into $\{H_i\}_{i=1}^C$ by maximum membership de-fuzzification as follows: $\mathbf{x}_k \in H_i$, if $\max_{j\in[1,C]}(u_{jk}) = u_{ik}$. If the maximum is reached by two or more clusters, one of these clusters is selected randomly.

The FCM clustering result of Fig. 1(a) is shown in Fig. 1(d); we use it for the segmentation-based locally adaptive weighted median and weighted average filterings in next section.

6   Adaptive Enhancement for Segmentation

Conventional linear/nonlinear filters always employ fixed-shape and fixed-size templates in a sliding window (here denoted as $\mathcal{W}_c$) to perform convolutions (Astola and Kuosmanen, 1997). The voxel to be filtered is generally the center voxel (denoted as $\mathbf{x}_c$) of $\mathcal{W}_c$. The entries in the templates can be selected in a nonlinear manner using statistics calculated from $\mathcal{W}_c$. Normally stationarity is assumed in the sliding window $\mathcal{W}_c$. However, this assumption is not always true for MRI images, especially for the edge voxels which are moreover affected by partial volume effects.

However, the FCM clustering result gives us some useful information about partial volume effects via the membership matrix U, and about local degree of spatial correlation and stationarity. Hence, we suppose that a more stationary neighborhood of the center voxel $\mathbf{x}_c$ can be constructed using the voxels which were segmented into the same cluster as $\mathbf{x}_c$ and which also are within the current sliding window (in our case, the sliding window $\mathcal{W}_c$ is a sliding cube; the influence of the size of $\mathcal{W}_c$ on the final segmentation results will be investigated quantitatively in Section 7). The resulting neighborhood, varying in the size and shape, considers both the correlation in the feature space via the segmentation, and the correlation in the spatial domain via the sliding window. We will call it the "reduced neighborhood".

In this context, we proposed a locally adaptive enhancement scheme based on the clustering result. The enhancement employs weighted median filters (Yin et al., 1996) or weighted average filters for different types of voxels in a 3-D adaptive framework.

Generally a voxel within a cluster H_i and far away from edges has a higher membership degree to H_i than edge voxels and noise classified into H_i. A weighted average filter is applied to such a voxel in its reduced neighborhood. For the voxels whose maximum membership degrees are not high enough (e.g., less than 80%), we use weighted median filters in their reduced neighborhoods to remove noise and to avoid destroying details and structures as what average filtering suffers from.

Furthermore, we use maximum membership degrees as weights of the elements within a filtering template. To weight the average and median with the maximum membership degrees will favor the replacement of edge voxels with interior voxels of a segment. This will enhance structures like edges and will increase the homogeneity within each region, and thus improves the performance of the following FCM clustering step.

6.1   Implementation of Adaptive Enhancement

First we introduce some notations. $\mathbf{x}$ represents a voxel in the images; $f(\mathbf{x})$ is its intensity. $u_j(\mathbf{x})$ denotes the membership degree of $\mathbf{x}$ to the j-th cluster; it can be obtained from the matrix U in the FCM. $u(\mathbf{x}) = \max_{j\in[1,C]}(u_j(\mathbf{x}))$ is the maximum membership degree of $\mathbf{x}$ to any of the C clusters. $\mathbf{x}_c$ is the voxel to be currently enhanced and also the center voxel of a sliding window $\mathcal{W}_c$. If $\mathbf{x}_c$ was classified into i_c-th cluster H_{i_c} by the maximum membership de-fuzzification in FCM, then $u(\mathbf{x}_c) = u_{i_c}(\mathbf{x}_c)$. Then a novel FCM-weighted adaptive median and average filtering on this voxel $\mathbf{x}_c$ is proposed and implemented as follows.

Step 1: Construct a reduced neighborhood Ω of $\mathbf{x}_c$ as $\Omega = \{\mathbf{x} \mid u(\mathbf{x}) = u_{i_c}(\mathbf{x}), \mathbf{x} \in \mathcal{W}_c\}$, and calculate the average maximum membership degree in Ω as follows: $\bar{u}_\Omega = \sum_{\mathbf{x}\in\Omega} u(\mathbf{x})/\lambda$, where λ is the number of voxels in Ω. Obviously, $\Omega \subseteq \mathcal{W}_c$. Then we check whether λ is less than three.

If $\lambda < 3$, $\mathbf{x}_c$ is possibly belonging to a small piece of isolated noise, and therefore we update $f(\mathbf{x}_c)$ using a simple weighted average over its 6-connected 3-D neighborhood, and then go to Step 4. The weights of the 6 neighbors are their maximum membership degrees.

Step 2: If $u(\mathbf{x}_c)$ exceeds a given threshold $u_{th}(i_c)$, we suppose that $\mathbf{x}_c$ is an interior voxel of the i_c-th cluster, and that the clustering of this voxel is reliable.

The thresholds $u_{th}(i_c)$, $i_c \in [1, C]$ are parameters that measures the reliability and validity of the FCM clustering result. $u_{th}(i_c)$ is selected automatically in the range $[0, 1]$ through binarizing the histogram of maximum membership degree $u(\mathbf{x})$ of all the voxels belonging to the i_c-th cluster within a 3-D volume. We choose Otsu method (Otsu, 1979) for the binarization. The Otsu method here selects a threshold to maximize the between-class variance of the binarized interior and non-interior voxels within the i_c-th cluster.

We update $f(\mathbf{x}_c)$ using a weighted averaging over voxels in Ω as

$$f(\mathbf{x}_c) = \sum_{\mathbf{x} \in \Omega} u(\mathbf{x}) f(\mathbf{x}) \; / \; \sum_{\mathbf{x} \in \Omega} u(\mathbf{x}) \; . \qquad (5)$$

The average is affected more by voxels with higher membership degrees, which are normally voxels closer to the prototype of the cluster to which $\mathbf{x}_c$ belongs. In such a way, the homogeneity within each cluster is improved.

Step 3a: If $u(\mathbf{x}_c) \leq u_{th}(i_c)$, $\mathbf{x}_c$ is likely to be an edge voxel or a noisy voxel. Furthermore, if the average maximum membership degree $\bar{u}_\Omega$ in Ω satisfies $\bar{u}_\Omega \leq u_{th}(i_c)$, we suppose that the clustering of the voxels in Ω is not reliable, and that Ω cannot be used to decide a filtering template. In this case, we ignore any filtering operation, and thus avoid augmenting the misclassifications of those voxels with similar degree to two clusters.

Step 3b: If $u(\mathbf{x}_c) \leq u_{th}(i_c)$ and $\bar{u}_\Omega > u_{th}(i_c)$, we update $f(\mathbf{x}_c)$ using a weighted median filtering in Ω. First we sort the intensities of all the voxels within Ω in ascending order, denoting the sorted intensi-

ties as $\{f_1, f_2, \cdots, f_\lambda\}$, i.e., $f_1 \leq f_2 \leq \cdots \leq f_\lambda$. The corresponding maximum membership degrees are $\{u_1, u_2, \cdots, u_\lambda\}$. The weighted intensity median f_{med}, $f_{med} \in \{f_i\}_{i=1}^\lambda$ is calculated as

$$f_{med} = \text{median}(\{u_1 \lozenge f_1, u_2 \lozenge f_2, \cdots, u_\lambda \lozenge f_\lambda\}) \; , \qquad (6)$$

where $\lozenge$ denotes a duplication operator.

Subsequently, we decide the index med of the median f_{med} as

$$med = \min(\{i \mid \sum_{j=1}^i u_j > \frac{1}{2} \sum_{j=1}^\lambda u_j, i \in [1, \lambda]\}) \; , \qquad (7)$$

i.e., starting from f_1, we accumulate the weights until the amount is larger than half the sum of all the weights in Ω. If all the weights are equal to one, we can obtain the standard median filter. Since the weight of each voxel is the corresponding membership degree, the computation of $f(\mathbf{x}_c) = f_{med}$ favors the intensities with highest degree in the neighborhood, and the misclassification will be reduced in the subsequent segmentation step.

Step 4: Update $f(\mathbf{x}_c)$ for the enhancement of next voxel, then move $\mathbf{x}_c$ to the next position, update the sliding window $\mathcal{W}_c$; then return back to Step 1 unless every voxel in the volume has already been processed once as a center voxel.

6.2 Segmentation after Adaptive Enhancement

The FCM will be applied once more to segment the images which has been filtered by the aforementioned adaptive enhancement. The enhancement and segmentation results of the simulated image Fig. 1(a) can be found respectively in Fig. 1(e) and Fig. 1(f). Meanwhile, in Fig. 2 we also show a real MRI image in axial view with its respective wavelet-based de-noising, thresholding, FCM clustering, adaptive enhancement and final segmentation results. After comparing both the Fig. 1(d) with Fig. 1(f) and the Fig. 2(d) with Fig. 2(f), we find that the adaptive enhancement provides a refinement of edge preservation and noise removal, especially for the "gray" noise in lower parts of WM in Fig. 1(d) which was

caused by the high intensity inhomogeneities, and in the upper region of white matter in Fig. 2(d).

(a)   (b)   (c)
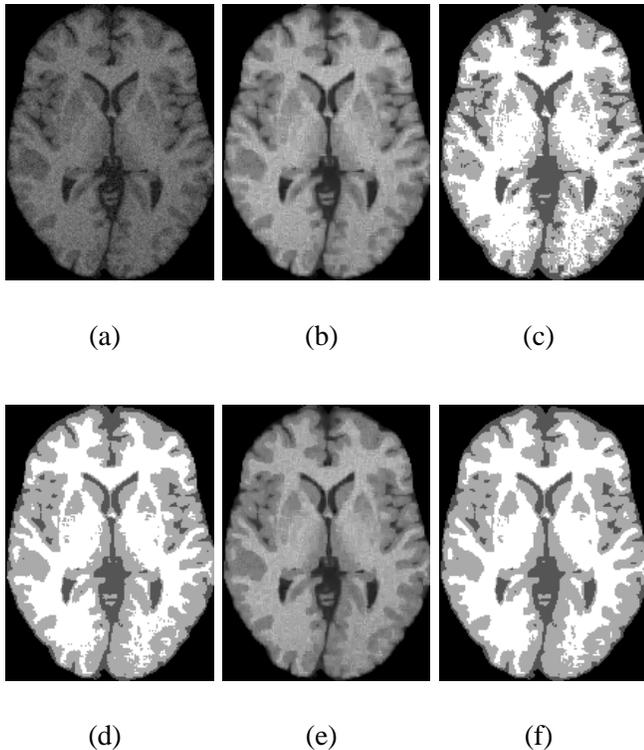

(d)   (e)   (f)

Figure 1. (a) Original image simulated from MRI brain phantom with 9% noise level and 40% intensity inhomogeneities, and its processed versions with (b) wavelet-based de-noising; (c) minimum error thresholding; (d) FCM clustering; (e) adaptive enhancement; (f) final segmentation result.

To be noted is that the performance of such an adaptive median filtering depends on the accuracy of a preceding segmentation. Therefore using an iteration of filterings and segmentations can improve the accuracy of the final segmentation. But the price of this iteration is the computational complexity and the risk of overenhancing some edges between different segments. From our experiments, we find that more than two iterations give little improvement on the segmentation results.

7 Quantitative Validation

To quantitatively validate our method, test images with known "ground truth" are required. For this purpose, we used a realistic digital brain phan-

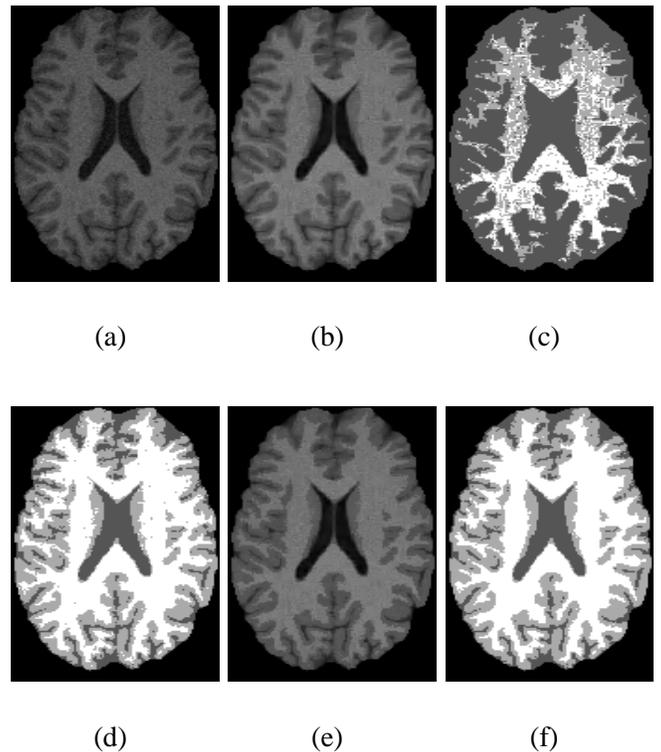(a)   (b)   (c)

(d)   (e)   (f)

Figure 2. (a) Original image sampled from real MRI volumes, and its processed versions with (b) wavelet-based de-noising; (c) minimum error thresholding; (d) FCM clustering; (e) adaptive enhancement; (f) final segmentation result.

tom (Kwan et al., 1999) considering the partial volume effects. A discrete anatomical model of three brain tissues is derived from the phantom by assigning the voxel a label of the tissue which contributes the most to that voxel. This model serves as the "ground truth" in our quantitative validation.

Based on the above phantom, four realistic MRI volumes are simulated with T1-weighted sequences, slice thickness of $1mm$, volume size of $217 \times 181 \times 181$, intensity inhomogeneities of 20%, and noise levels of 3%, 5%, 7% and 9% respectively (Kwan et al., 1999) for the validation. The skull, scalp, unnecessary background and slices with few brain voxels are first removed with the guidance of the "ground truth", thus the brain of interest consisting of CSF, GM and WM is extracted and then segmented by our proposed method. The enhancement and segmentation results of two sample images from the volume with 9% noise level are shown in Fig. 3 with their "ground truth." 3-D volumes simulated from this phantom are also used for quantitative val-

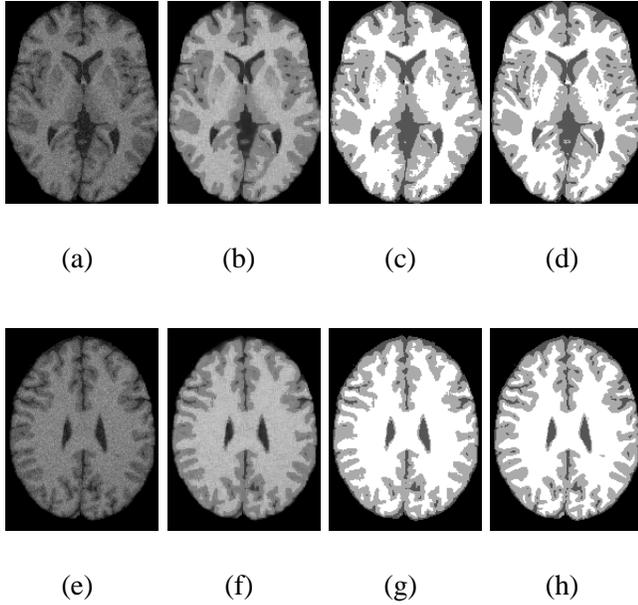idations in Kollokian (1996) and Ruan et al. (2000, 2002).

Figure 3. (a) and (e) are two sample images simulated from MRI brain phantom with 9% noise level and 20% intensity inhomogeneities, with (b) the adaptive enhancement of (a), (c) the segmentation result of (a) obtained by our proposed method, and (d) the "ground truth" of (a). Corresponding to (b), (c) and (d) respectively, (f), (g) and (h) are the processed versions of (e).

We employ four different indices (false positive ratio γ_{fp}, false negative ratio γ_{fn}, similarity index ρ (Zijdenbos et al., 1994), and kappa statistic κ (Zijdenbos et al., 1994)) for each of three brain tissues as quantitative measures to validate the accuracy and reliability of our method.

For a given brain tissue $i, i = 1, 2, 3$ for CSF, GM and WM respectively, suppose that A_i and B_i represent the sets of voxels labeled into i by the "ground truth" and by our method respectively. $|A_i|$ denotes the number of voxels in A_i.

The widely-used false positive ratio γ_{fp}, representing the error due to the misclassification in a brain tissue i, is defined as $\gamma_{fp} = (|B_i| - |A_i \cap B_i|)/|A_i|$. Likewise, the false negative ratio γ_{fn}, representing the error due to the loss of desired voxels of i, is defined as $\gamma_{fn} = (|A_i| - |A_i \cap B_i|)/|A_i|$.

The similarity index ρ is an intuitive and plain in-

dex to consider the matching volume/area between A_i and B_i, defined as $\rho = 2|A_i \cap B_i|/(|A_i| + |B_i|)$. ρ is sensitive to discrepancies in shape, location, and size; $\rho > 0.7$ indicates an excellent similarity (Zijdenbos et al., 1994).

In fact, the similarity index is a special case of kappa statistic (Zijdenbos et al., 1994). The kappa statistic is a chance-corrected measure of agreement between two results, defined as $\kappa = (P_o - P_e)/(1 - P_e)$, where P_o is the observed proportion of agreement between two results, defined as $P_o = (|A_i \cap B_i| + |\bar{A}_i \cap \bar{B}_i|)/\sum_j |A_j|$ for tissue i, where $\bar{A}_i$, $\bar{B}_i$ denotes the complement of A_i, B_i respectively; P_e is the expected proportion of agreement due to chance alone, $P_e = (|A_i| \cdot |B_i| + |\bar{A}_i| \cdot |\bar{B}_i|)/(\sum_j |A_j|)^2$. An agreement greater than chance alone results in $\kappa \in (0, 1]$. $\kappa > 0.75$ was suggested a strong agreement above chance in Fleiss (1981); $\kappa \in [0.81, 1]$ indicates an "almost perfect" agreement according to Landis and Koch (1977).

To compare the validation results reported in Ruan et al. (2000) and Kollokian (1996), a simple kappa statistic κ_a is also calculated as κ, where $P_o = \sum_i |A_i \cap B_i|/\sum_i |A_i|$, and $P_e = \sum_i(|A_i| \cdot |B_i|)/(\sum_i |A_i|)^2$. κ_a considers all the classifications of three tissues as a whole.

In order to simultaneously investigate the sensitivity of our proposed method to noise, we plot the validation results of the aforementioned four realistic MRI volumes for four noise levels in Fig. 4 (the edge length of the sliding cube $\mathcal{W}_c$ is 5 voxels here).

From Fig. 4, we can find that both the false positive ratio γ_{fp} and false negative ratio γ_{fn} of CSF approximately lie in $[0.05, 0.10]$; they are the largest among three tissues considering the same indices under almost all the conditions with different noise. It means that more than 5% of voxels which are labeled as CSF are wrongly taken from other tissues, and more than 5% of desired voxels of CSF has been misclassified into GM and WM. On the other hand, we can say that more than 90% of voxels of CSF are labeled correctly. And both γ_{fp} and γ_{fn} of WM are the smallest and are less than 8%, which means that more than 92% of voxels of WM are also labeled
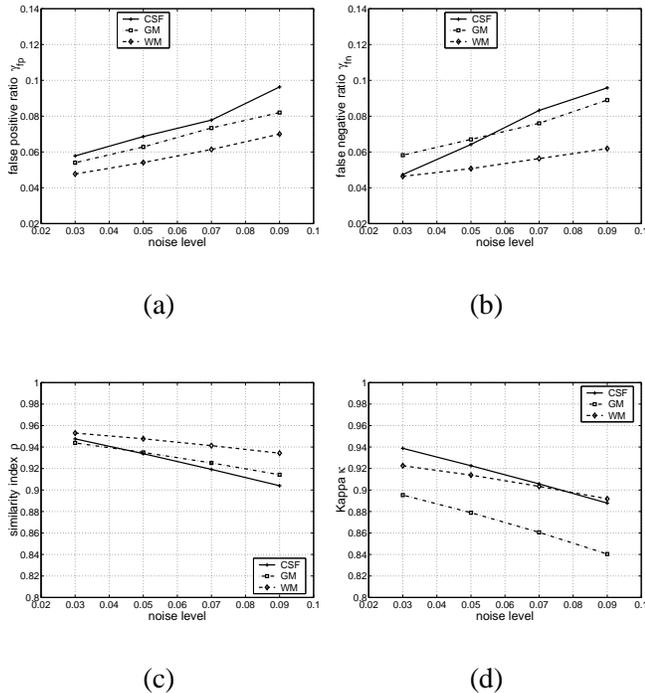
Figure 4. Validation results for different noise levels with measures of (a) false positive ratio γ_{fp}; (b) false negative ratio γ_{fn}; (c) similarity index ρ; (d) kappa statistic κ.

correctly.

The similarity indices ρ of all the tissues are larger than 0.90. Hence, the overlap degree between our segmentation results and the "ground truth" is higher than 90%, even for a bad condition with 9% noise level. Meanwhile, the obtained values of kappa statistic κ in Fig. 4 also indicate an "almost perfect" agreement between our segmentation results and the "ground truth".

In Kollokian (1996), seven pattern classifiers were applied to segment brain tissues by using 3-D simulated T1-, T2-, and Proton-Density-weighted volumes. In the worst case presented (20% intensity inhomogeneities and 9% noise level), the best result was obtained by the Back-propagation Artificial Neural Network, where κ_a is about 0.81. The κ_a obtained by the FCM classifier in Kollokian (1996) is about 0.58. Another method proposed in Ruan et al. (2000), using Markov random field and partial volume modeling for the T1-weighted volume, yielded a better result, where $\kappa_a = 0.85$ in the above case. Our method proposed in this paper estimates κ_a from the same T1-weighted volume, and offers approxi-

mately 0.87.

From Fig. 4, we can also find that: for any of the three brain tissues, the maximum difference between two noise levels (3% and 9%) is about 5% in all the indices; with the noise level increasing, the false positive and negative ratios are monotonically increasing, while ρ and κ are monotonically decreasing.

Meanwhile, we also investigate the influence of the size of the sliding window $\mathcal{W}_c$ on the final segmentation results. Four different edge lengths (3, 5, 7 and 9 voxels respectively) of $\mathcal{W}_c$ are applied to the four aforementioned realistic MRI volumes for validation. The validation results in terms of the index κ_a are shown in Fig. 5. From this figure, we find that, although the maximum difference in κ_a is less than 0.02 with regard to any individual noise levels, there is a steeper slope between the edge lengths of 3 and 5 voxels under all the four conditions with different noise. Hence, we choose 5 voxels as the default edge length of $\mathcal{W}_c$.
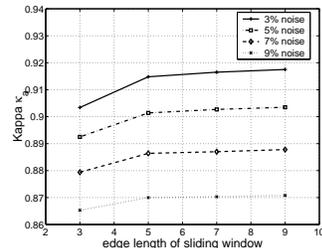

Figure 5. Influence of the size of the sliding window $\mathcal{W}_c$ on the final segmentations. Four edge lengths of $\mathcal{W}_c$ (3, 5, 7 and 9 voxels of length respectively) are applied to four realistic MRI volumes with different noise, and κ_a is used for quantitative validation.

Due to the lack of "ground truth", real MRI brain volumes for EEG/MRI analysis are used here only for qualitative validation. These volumes were acquired on a Siemens Magnetom Symphony 1.5T MR scanner at the Ghent University Hospital, by using the Turbo-FLASH sequences. The skull and scalp are first removed by the ATOMIA software (Moretti et al., 2000) before applying our method. As you have seen, an image sampled from one of the above volumes is shown in the Fig. 2(a). From Fig. 2, we can see that, although the minimum error thresholding did not offer a satisfactory initialization for FCM clustering, our proposed method still removes

10

noise and preserves edges well, and provides an acceptable segmentation result as most of tissues are segmented correctly.

8 Conclusion

In this paper, we have presented an integrated method of the adaptive enhancement for the unsupervised global-to-local segmentation of three brain tissues (CSF, GM and WM) in single-channel 3-D MRI images. To enable the effective and robust implementation of such an enhancement and segmentation, we have first integrated a versatile wavelet de-noising algorithm with the minimum error thresholding based on a global intensity threshold, then combined an FCM clustering using 3-D spatial context with locally adaptive weighted median and weighted average filters which are selected and weighted by the clustering result. This method yields a good segmentation performance even in noisy images, and it has been validated on simulated and real MRI brain images. The validation results demonstrate an encouraging future of practical applications of the proposed method.

References

Astola, J., Kuosmanen, P., 1997. Fundamentals of Nonlinear Digital Filtering. CRC Press, Boca Raton.

Bezdek, J., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York.

Bezdek, J., Hall, L., Clarke, L., 1993. Review of MR image segmentation techniques using pattern recognition. Medical Physics 20 (4), 1033–1048.

Clark, M., Hall, L., Goldgof, D., Clarke, L., Velthuizen, R., Silbiger, M., 1994. MRI segmentation using fuzzy clustering techniques. IEEE Engineering in Medicine and Biology Magazine 13 (5), 730–742.

Clarke, L., Velthuizen, R., Camacho, M., Heine, J., Vaidyanathan, M., Hall, L., Thatcher, R., Silbiger, M., 1995. MRI segmentation: methods and application. Magnetic Resonance Imaging 13 (3), 343–368.

Donoho, D., Johnstone, I., 1995. Adapting to unknown smoothness via wavelet shrinkage. Journal of the American Statistical Association 90 (432), 1200–1224.

Fleiss, J., 1981. Statistical Methods for Rates and Proportions, 2nd Edition. John Wiley & Sons, New York.

Kittler, J., Illingworth, J., 1986. Minimum error thresholding. Pattern Recognition 19 (1), 41–47.

Kollokian, V., November 1996. Performance analysis of automatic techniques for tissue classification in MRI of the humain brain. Master's thesis, Concordia University, Montreal, Canada.

Kwan, R.-S., Evans, A., Pike, G., 1999. MRI simulation-based evaluation of image-processing and classification methods. IEEE Trans. Medical Imaging 18 (11), 1085–1097, http://www.bic.mni.mcgill.ca/brainweb/.

Landis, J., Koch, G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174.

Mallat, S., 1999. A Wavelet Tour of Signal Processing, 2nd Edition. Academic Press, San Diego.

Moretti, B., Fadili, M., Ruan, S., Bloyet, D., Mazoyer, B., 2000. Phantom-based performance evaluation - application to brain segmentation from magnetic resonance images. Medical Image Analysis 4 (4), 303–316.

Niessen, W., Vincken, K., Weickert, J., Haar Romeny, B., Viergever, M., 1999. Multiscale segmentation of three-dimensional MR brain images. International Journal of Computer Vision 31 (2/3), 185–202.

Nowak, R., 1999. Wavelet-based Rician noise removal for magnetic resonance imaging. IEEE Trans. Image Processing 8 (10), 1408–1419.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Systems, Man, and Cybernetics 9 (1), 62–66.

Pal, N., Pal, S., 1993. A review on image segmentation techniques. Pattern Recognition 26 (9), 1277–1294.

Pham, D., Xu, C., Prince, J., 2000. Current methods in medical image segmentation. Annual Review of Biomedical Engineering 2, 315–337.

Pizurica, A., June 2002. Image denoising using wavelets and spatial context modeling. Ph.D. thesis, Ghent University, Gent, Belgium.

Pizurica, A., Philips, W., Lemahieu, I., Acheroy, M., 2003. A versatile wavelet domain noise filtration technique for medical imaging. IEEE Trans. Medical Imaging, special issue on Wavelets in Medical Imaging In press.

Ruan, S., Jaggi, C., Xue, J.-H., Fadili, J., Bloyet, D., 2000. Brain tissue classification of Magnetic Resonance Images using partial volume modeling. IEEE Trans. Medical Imaging 19 (12), 1179–1187.

Ruan, S., Moretti, B., Fadili, J., Bloyet, D., 2002. Fuzzy Markovian segmentation in application of Magnetic Resonance Images. Computer Vision and Image Understanding 85 (1), 54–69.

Sahoo, P., Soltani, S., Wong, A., Chen, Y., 1988. A survey of thresholding techniques. Computer Vision, Graphics and Image Processing 41 (2), 233–260.

Van Hoey, G., De Clercq, J., Vanrumste, B., Van de Walle, R., Lemahieu, I., D'Havé, M., Boon, P., 2000. EEG dipole source localization using artificial neural networks. Physics in Medicine and Biology 45 (4), 997–1011.

Xue, J.-H., Ruan, S., Moretti, B., Revenu, M., Bloyet, D., 2001. Knowledge-based segmentation and labeling of brain structures from MRI images. Pattern Recognition Letters 22 (3-4), 395–405.

Ye, Q.-Z., Danielsson, P.-E., 1988. On minimum error thresholding and its implementations. Pattern Recognition Letters 7 (4), 201–206.

Yin, L., Yang, R., Gabbouj, M., Neuvo, Y., 1996. Weighted median filters: a tutorial. IEEE Trans. Circuits and Systems II 43 (3), 157–192.

Zhang, Y., 1997. Evaluation and comparison of different segmentation algorithms. Pattern Recognition Letters 18 (10), 963–974.

Zijdenbos, A., Dawant, B., 1994. Brain segmentation and white matter lesion detection in MR images. Critical Reviews in Biomedical Engineering 22 (5-6), 401–465.

Zijdenbos, A., Dawant, B., Margolin, R., Palmer, A., 1994. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Trans. Medical Imaging 13 (4), 716–724.