

# A protocol-independent brain MRI segmentation method

László G. Nyúl<sup>a</sup> and Jayaram K. Udupa<sup>b</sup>

<sup>a</sup>Department of Applied Informatics, University of Szeged, Szeged, Hungary

<sup>b</sup>Medical Image Processing Group, Department of Radiology,  
University of Pennsylvania, Philadelphia, PA

## ABSTRACT

We present a segmentation method that combines the robust, accurate, and efficient techniques of fuzzy connectedness with standardized MRI intensities and fast algorithms. The result is a general segmentation framework that more efficiently utilizes the user input (for recognition) and the power of computer (for delineation). This same method has been applied to segment brain tissues from a variety of MRI protocols. Images were corrected for inhomogeneity and standardized to yield tissue-specific intensity values. All parameters for the fuzzy affinity relations were fixed for a specific input protocol. Scale-based fuzzy affinity was used to better capture fine structures. Brain tissues were segmented as 3D fuzzy-connected objects by using relative fuzzy connectedness. The user can specify seed points in about a minute and tracking the 3D fuzzy-connected objects takes about 20 seconds per object. All other computations were performed before any user interaction took place. Segmentation of brain tissues as 3D fuzzy-connected objects from MRI data is feasible at interactive speeds. Utilizing the robust fuzzy connectedness principles and fast algorithms, it is possible to interactively select fuzzy affinity, seed point, and threshold parameters and perform efficient, precise, and accurate segmentations.

**Keywords:** Image segmentation, fuzzy connectedness, interactive segmentation, image scale, magnetic resonance imaging (MRI)

## 1. INTRODUCTION

The purpose of image segmentation is to extract object information from given images and to output this as a structure system. Segmentation is needed directly or indirectly for most of the operations performed on images. It is also the most difficult of all image operations.

Segmentation may be thought of as consisting of two related tasks — recognition (or detection) and delineation. Recognition is the high-level task of determining roughly the whereabouts of the object in the given image. Delineation is the low-level task of determining the precise spatial extent of the object and its point-by-point graded composition. In most recognition tasks, trained human operators outperform any computer algorithms. On the contrary, computer algorithms exist for delineation that are more precise, accurate, and efficient than human delineation of object regions or boundaries. Recognition and delineation are not completely disparate steps in segmentation.

Images are by nature fuzzy. Approaches to object information extraction from images should therefore attempt to retain uncertainties as realistically as possible. Local fuzziness has been accounted for in fuzzy delineation approaches in the past. A recently developed theoretical and algorithmic framework called fuzzy connectedness attempts to capture the notion of “hanging togetherness” of image elements in a global fuzzy sense for object definition.

Segmentation techniques using fuzzy connectedness principles have been used successfully in several large applications in recent years including segmentation of brain and separation of brain tissues from MR images, MS lesion and brain tumor quantification. However, there were several parts in these techniques where improvements can be achieved. All segmentation techniques depend on parameters. In previous applications of the fuzzy connectedness method, some of the parameters were set in advance, while others could only be extracted from the data being segmented, requiring

---

Further author information: (Send correspondence to J.K.U.)

L.G.N.: E-mail: nyul@inf.u-szeged.hu, Telephone: +36 62 420-184, Address: Department of Applied Informatics, University of Szeged, P.O.Box 652, H-6701, Hungary

J.K.U.: E-mail: jay@mipg.upenn.edu, Telephone: 1 215 662-6780, Address: MIPG, University of Pennsylvania, Department of Radiology, 423 Guardian Drive – Fourth Floor, Blockley Hall, Philadelphia, PA 19104-6021, USA

human help. For example, parameters for the distribution of voxel intensities for a tissue was estimated from the small neighborhood of a few voxels that were specified by the user. This poses at least two problems. First, the reliability of the parameters is questionable, since it heavily depends on where the point was selected by the user. In case of fine, thin structures, such as gray matter, selecting reliable regions may be difficult. Also, this way, several user sessions are required within the entire segmentation process: one for the parameter estimation, and another for verification. The latter may take place hours or even days after the first session.

Techniques have been developed to overcome these issues. A technique for standardizing MR image intensities has been devised. It allows parameter setting without the need for user input. Also, fast algorithms have been developed for computing fuzzy connectedness at interactive speeds. If images having intensities with tissue specific meaning, such as standardized MR images, are utilized, most of the parameters for the segmentation method can be fixed once for all, all intermediate data can be computed before the user interaction is needed and the user can be provided with more information at the time of interaction. The reliable recognition (assisted by human operators) and the accurate, efficient, and sophisticated delineation (automatically performed by the computer) can be effectively incorporated into a single interactive process. This, we believe, is a major improvement in segmentation.

In this paper, we describe an approach to the segmentation of the whole brain parenchyma (BP), as well as to the precise quantification of gray matter (GM) and white matter (WM) on several routinely acquired MR images of MS patients, such as fast spin-echo (FSE) proton density (PD) and T2-weighted, T1-weighted, and spoiled-gradient (SPGR) images. We do not attempt, however, to segment abnormal tissue regions, such as MS lesions, as separate objects in this paper.

A brief description of fuzzy connectedness is given in Section 2. Details of the new segmentation method are given in Section 3. Section 4 describes the methods of evaluation used in this work. Results and discussion are given in Section 5, and concluding remarks in Section 6.

## 2. FUZZY CONNECTEDNESS PRINCIPLES

Before delving into the details of the method, we explain the main idea behind fuzzy connectedness. We refer to an  $n$ -dimensional digital image as a *scene* and represent it by a pair  $\mathbf{C} = (C, \mathbf{f})$ , where  $C$  is a rectangular  $n$ -dimensional array of spatial elements (*spels* for short) and  $\mathbf{f}$  is a function (called *scene intensity*) which assigns to every spel a vector of  $m$  integer intensity values. The spels are pixels for  $n = 2$  and voxels for  $n = 3$ . We use the notation  $c = (c_1, c_2, \dots, c_n)$  to denote the coordinates of any spel  $c \in C$  in the array.

Independent of any image data, we think of the digital grid system defined by the spels as having a *fuzzy adjacency relation*. This relation assigns to every pair  $(c, d)$  of spels a value between zero and one. The closer  $c$  and  $d$  are spatially to each other, the greater is this number. This is intended to be a “local” phenomenon. How “local” it ought to be should depend on the blurring property of the imaging device. We denote the fuzzy adjacency relation by  $\alpha$  and the degree of adjacency assigned to any spels  $(c, d)$  by  $\mu_\alpha(c, d)$ .

Now consider the grid points (spels) as having scene intensities assigned to them. That is, we are given a scene  $\mathbf{C} = (C, \mathbf{f})$ . We define another local fuzzy relation called *affinity* on spels denoted by  $\kappa$ . The strength of this relation between any spels  $c$  and  $d$ , denoted  $\mu_\kappa(c, d)$ , lies between zero and one, and indicates how the spels “hang together” locally in the scene.  $\mu_\kappa(c, d)$  is determined based on  $\mu_\alpha(c, d)$ , as well as on how similar the intensities or intensity-based properties at  $c$  and  $d$  are. The properties of fuzzy affinity relations are studied extensively and guidance as to how to setup fuzzy affinities in practical applications is given in.<sup>1</sup> One practical recommendation is *scale-based affinity*. It has been shown that using scale-based affinity leads to significantly improved image segmentation.<sup>1</sup>

The idea of “scale” used here, is to determine the size of local structures under a prespecified region-homogeneity criterion. Scale in a scene  $\mathbf{C}$  at any voxel  $c \in C$  is defined as the radius  $r(c)$  of the largest ball centered at  $c$  which lies entirely within the same object region (defined under a prespecified region-homogeneity criterion). This adaptive scale concept has evolved in connection with fuzzy connectedness principles<sup>2</sup> and represents the local size of the object at the voxel. In,<sup>1</sup> a simple and effective algorithm has been described which estimates  $r(c)$  at every location  $c \in C$  in any scene  $\mathbf{C}$  without explicit object segmentation but based on continuity of region homogeneity.

Generally, both the homogeneity- and object-feature-based components should be considered in the design of fuzzy spel affinities, although in some applications each component on its own may be adequate. The general formulation for

affinity is  $\mu_{\kappa}(c, d) = \mu_{\alpha}(c, d)g(\mu_{\psi}(c, d), \mu_{\phi}(c, d))$ , where  $\mu_{\psi}$  and  $\mu_{\phi}$  represent the homogeneity-based and object-feature-based components of affinity, respectively, and  $g$  is a function that combines the two fuzzy values. Both  $\mu_{\psi}$  and  $\mu_{\phi}$  are computed using the scale region around the spels.<sup>1</sup>

Our aim is to capture the global phenomenon of “hanging togetherness” in a global fuzzy relation on spels called *fuzzy connectedness*, denoted  $K$ . The strength of this relation  $\mu_K(c, d)$  between any spels  $c$  and  $d$ , indicating the strength of their connectedness, lies between zero and one, and is determined as follows: There are numerous possible “paths” within the scene domain  $C$  between  $c$  and  $d$ . Each path for our purposes is a sequence of spels, starting from  $c$  and ending in  $d$ , with the successive spels being nearby. (In general, “nearby” means that the successive spels  $v$  and  $v'$  are such that  $\mu_{\alpha}(v, v') \neq 0$ .) We think of each pair of successive spels as constituting a link and the whole path to be a chain of links. We assign a strength (between zero and one) to every path which is simply the smallest pairwise spel affinity along the path. This indicates the weakest link in the chain. Finally, the strength of connectedness between  $c$  and  $d$  is the strength associated with the strongest of all paths between  $c$  and  $d$ .

A *fuzzy connected object*  $\mathbf{O}$  of  $\mathbf{C}$  of strength  $\theta_x = [x, 1]$ ,  $0 \leq x \leq 1$ , and containing a spel  $o$  consists of a pool  $O \subset C$  of spels together with a value indicating “objectness” assigned to every spel.  $O$  is such that  $o \in O$ , for any spel  $c$  and  $d$  in  $O$ , the strength of connectedness between them  $\mu_K(c, d) \geq x$ , and for any spels  $c \in O$  and  $e \notin O$ , the strength  $\mu_K(c, e) < x$ . The value of objectness assigned to spels in  $O$  varies between zero and one and the value assigned to spels outside  $O$  is zero. See<sup>3</sup> for a precise mathematical definition. There are several choices as to how to assign objectness to spels in  $O$ , see.<sup>1,3</sup> A *K-connectivity scene* of a scene  $\mathbf{C} = (C, \mathbf{f})$  with respect to a spel  $o \in C$  and fuzzy affinity  $\kappa$  in  $C$  is a scene  $\mathbf{C}_o = (C_o, \mathbf{f}_o)$  such that  $C_o = C$ , and for any spel  $c \in C$ ,  $\mathbf{f}_o(c) = \mu_K(o, c)$ .

*Relative fuzzy connectedness*<sup>4</sup> eliminates the need for selecting a threshold value  $\theta_x$  on the  $K$ -connectivity scene by having the objects “compete” for the spels. A spel is assigned to that object to which (meaning, with respect to the seeds of which) its fuzzy connectedness value is the highest. This is the strategy that is employed in this paper.

### 3. INTERACTIVE SEGMENTATION METHOD

First, we give an overview of the segmentation method. It is followed by more detailed description of each step in subsequent sections.

#### 3.1. Overview of the Segmentation Method

We distinguish two phases in this work. First, the segmentation procedure is set up, parameters are trained for, and fine tuned. This is performed only once for each protocol. Then, each dataset is segmented into the desired objects (i.e., BP, WM and GM). Below are the main steps of the segmentation method. We give details of each step in subsequent sections.

- TS1-TS5: Training and set up
- SS1: Correcting for RF field inhomogeneity
- SS2: Standardizing MR image intensities
- SS3: Computing fuzzy affinity
- SS4: Specifying seed points and volume of interest
- SS5: Computing fuzzy connectedness
- SS6: Creating the brain mask
- SS7: Verifying the brain mask
- SS8: Creating masks for each object
- SS9: Computing volume

Most steps are automatic, their parameters are determined and fixed in the training/setup phase. For those steps, where user input is needed, we shall mention the amount of input and the way it is required. Depending on what acquisition protocol was used, small differences may be present as to how fuzzy affinity is computed. For example, when using dual-echo FSE images, both PD and T2 scenes, as well as a third, derived scene are utilized in the affinity computation, while for T1, for example, only one input scene is used. This difference may be considered hidden if we formulate it with vector valued scenes and affinities. However, in our implementation, affinities for each image are separately computed and then combined. See<sup>5</sup> for a fully vectorial formulation of fuzzy connectedness.

### 3.2. Training

For the training phase, for each protocol, a few datasets are selected and used to extract the values for the parameters. These images are also used for testing the flow of operations and the control scripts step-by-step. The training procedure mostly requires continuous user control, since the fine tuning of parameters is a modify-and-verify iteration. The major steps of the training phase are the followings.

TS1: Perform the initial steps of the segmentation algorithm (SS1-SS2). SS4 (seed and VOI specification) may also be performed in advance. The reason for it being listed in the segmentation procedure after SS3 (affinity computation) is to make it possible to require only one user session. Otherwise, since affinity computation is a relatively slow operation, seed specification and brain mask verification would be possible only in two separate interactive sessions.

TS2: Determine threshold intervals to roughly segment the different tissues (CSF, WM, GM) in the intensity images. Threshold values for different input images should now be close to each other due to intensity standardization; thus, a common threshold can be determined. Compute the distribution of intensities within these segmented regions for each of the training data sets and compute a common value (average) for their parameters (mean and standard deviation). These values are good initial values for setting up the fuzzy affinity relations.

TS3: Perform the segmentation steps of the algorithm (SS3-SS8).

TS4: Evaluate the segmented objects. If necessary, modify the parameters and go to TS3. Otherwise, go to TS5.

TS5: Save the found best parameters for future use.

### 3.3. Correcting for RF Field Inhomogeneity

One of the well-known problems in MR imaging is intensity variation due to the RF field inhomogeneity. Most image processing and analysis methods are affected by this variation and their performance can be improved considerably if this variation is corrected. We applied a general, entirely image-based strategy for correcting this variation. The method is completely automatic, acquisition-protocol-independent, and requires no prior knowledge.<sup>6</sup>

It is an iterative algorithm in which at each iteration a set of homogeneous tissue regions are obtained, no matter what each tissue region is. A smooth (2nd order polynomial) function is fit to the intensity variation within each tissue region and then combined into a single “gain field” for which the image is corrected. To determine homogeneous tissue regions, local scale values<sup>1</sup> are computed in intensity standardized images.<sup>7</sup> Then, for the  $k$  largest mutually exclusive scale regions in the scale image, the mean  $\mu_k$ , and standard deviation  $\sigma_k$  of the standardized intensities is determined within the corresponding regions of the standardized image. The standardized image is thresholded depending on these parameters to yield homogeneous regions. Standardization is effected at the end to make sure that the corrected images are not biased by the fitting strategy and that the intensity gray scale variations are also accounted for.

### 3.4. Standardizing MR Image Intensities

The image intensities in MRI do not have a fixed meaning, not even within the same protocol for the same body region obtained on the same scanner for the same patient. This poses problems in image display, segmentation, and quantification. The results of segmentation by two physicians are likely to differ because of the differences in their training. Setting values for the parameters for the non-manual methods becomes more difficult without the same protocol-specific intensity meaning.

The method described in<sup>7</sup> offers a simple way of transforming the images so that there is a significant gain in similarity of the resulting images. It is a two-step process consisting of a training step and a transformation step. This new transformation results in a standard scale for each protocol and body region. Standardized images can be displayed with fixed windows without the need of per case adjustment. The method facilitates improving the degree of automation of image segmentation. More importantly, extraction of quantitative information about healthy organs or about abnormalities can be considerably simplified.

### 3.5. Computing Fuzzy Affinity

For different acquisition protocols fuzzy affinity was computed slightly differently due to the differences of the input scenes (for example, there are single- and dual-echo protocols). However, the main scheme was the same for all protocols, scale-based affinity with Gaussian functions was used.<sup>1</sup> For computational simplicity, we used the six-adjacency relation for  $\alpha$ , weighted by the appropriate voxel dimensions.

For the dual-echo FSE protocol, we consider the given vector-valued scene to be equivalent to two scenes  $\mathbf{C}_{PD} = (C, f_{PD})$  and  $\mathbf{C}_{T2} = (C, f_{T2})$ . The *angle scene*  $\mathbf{C}_{\text{angle}} = (C, f_{\text{angle}})$  is derived from  $\mathbf{C}_{PD}$  and  $\mathbf{C}_{T2}$  such that  $f_{\text{angle}}(c) = \tan^{-1} \frac{f_{T2}(c)}{f_{PD}(c)}$ . This scene emphasizes CSF with bright intensities while suppressing other tissues.<sup>2</sup> For each of GM and WM, affinity was computed from both  $\mathbf{C}_{PD}$  and  $\mathbf{C}_{T2}$  scenes separately, and then combined into a single fuzzy affinity for the object using a simple min operator. For CSF, affinity was computed from  $\mathbf{C}_{\text{angle}}$ .

For the T1 protocol, affinity was computed for GM and WM. However, a separate affinity for CSF was not used because CSF has similar intensities to that of bone, and peripheral CSF is indistinguishable from the enclosing bone, thus CSF was treated as background for the brain. For SPGR, affinity was computed for GM and WM, as well as for CSF.

### 3.6. Specifying Seed Points and Volume of Interest

At this point, user help is needed in the form of specifying seed points for each of WM, GM, and ventricular CSF. Four points for each of WM, and CSF, and 8 points for GM were input for each scene. During this same interactive session, the user is also asked to specify the first and last slices in which brain tissue is present. This volume of interest is utilized to exclude those parts of the image (mostly the bottom slices) that also contain regions which have similar affinity values to those of the objects of interest and may get connected to them during computation of fuzzy connectedness.

### 3.7. Computing Fuzzy Connectedness

Relative fuzzy connectedness was used to determine the fuzzy objects for different tissues. A voxel  $c$  is assigned to that object with respect to whose seeds  $c$ 's fuzzy connectedness value is the strongest.

### 3.8. Creating the Brain Mask

If we were using absolute fuzzy connectedness, the K-connectivity scenes would need to be thresholded to give binary masks for the objects. However, relative fuzzy connectedness directly outputs binary masks for the objects without requiring a threshold.

In this step, we combine GM, WM, and CSF masks to yield a brain mask (the intracranial volume). Then we apply some morphological operators. This is necessary for several reasons. In patient data, abnormal tissues may lead to very low (zero) affinity values within the normal tissue volume (e.g., WM lesions) and no object would claim those voxels in the "competition" of relative fuzzy connectedness. By morphological operation, these holes get filled and misclassification error reduced. Also, small, isolated clusters of voxels classified as one object get reassigned labels.

### 3.9. Verifying the Brain Mask

The "retouched" brain mask is presented to the operator for verification. The correction performed by the operator is mainly removing extracranial regions that may get connected to the brain via weak links of fuzzy connectedness.

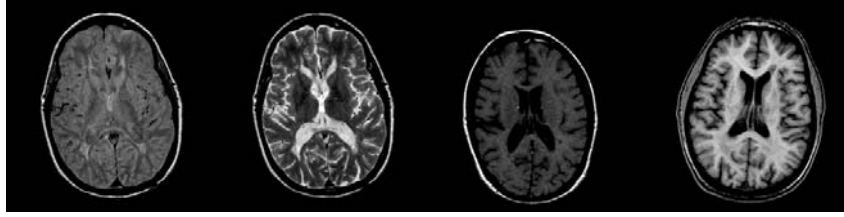
The previous two steps (computing fuzzy connectedness and creating the brain mask) are so rapid (a few seconds per scene) that this and the previous interaction (seed and VOI specification) may be included in the same user session.

### 3.10. Creating Masks for Each Object

GM, WM objects are recalculated within the corrected brain mask.

### 3.11. Computing Volume

Volume of each object is simply computed by multiplying the number of voxels within its binary mask by the voxel volume.



**Figure 1:** Slices of corrected and intensity standardized datasets. FSE PD, T2, T1, and SPGR.

## 4. EVALUATION

### 4.1. Image Data

MRI head datasets of over 40 different MS patients and 10 normal controls acquired with several routinely applied MRI protocols were used in the experiments described below. The patients and datasets were randomly selected from our database that contains about 100 patients with about 4000 3D scenes comprising fast-spin-echo dual-echo (FSE) T2-weighted (T2) and proton-density (PD) weighted studies, T1-weighted studies without (T1) and with Gadolinium enhancement (T1E), magnetization transfer (MT) studies without and with saturation pulse, as well as spoiled-gradient (SPGR) studies.

The selected datasets were visually inspected and found to be of good quality having no severe artifacts. Although a few datasets contained visible movement artifact or MRI shading, we kept them for the experiment. Unfortunately, artifacts occur in real situations and we wanted to evaluate the segmentation method in non-idealistic, real situations. Most datasets were of MS patients, some of them with severe lesion load. We have not treated abnormal tissue (such as abnormal WM and MS lesions) as separate objects; these have been segmented as part of WM or GM. Their separation can be accomplished in a subsequent step within each mask (e.g., separation of normal WM, abnormal WM, and lesions within the WM mask).

All datasets were acquired on the same GE Signa 1.5T scanner with a quadrature transmitter/receiver head coil and consisted of contiguous axial slices covering the entire brain with slice dimensions  $256 \times 256$  and pixel size of  $0.86 \text{ mm}^2$ , and FOV = 22 cm. The FSE PD-T2 pairs were acquired with interleaved slices. All images were acquired at 3 mm slice thickness. The following protocol parameters were used: for FSE, TR = 2500 ms, with TE = 16–18 and 80–96 ms, NEX = 1; for T1, TR = 600 ms and TE = 27 ms, NEX = 1; for SPGR, TI = 500 ms and flip angle = 20, NEX = 2. For each patient, the 3 mm datasets consisted of about 50–60 slices. Slices of images of several protocols are displayed in Figure 1.

Repeated MR studies of several patients were selected for each of FSE, T1, and SPGR protocols. The two scans in the pairs for each patient were obtained with a short time gap in between during which the patients were removed from the table and then positioned again.

### 4.2. Accuracy

Assessing accuracy of delineation is a difficult task due to the lack of a ground truth. Almost always some surrogates of truth, such as mathematical or physical phantoms, and manual tracing are used. In all cases, the authenticity of the surrogate becomes questionable. Additionally, the precision with which the surrogate itself can be measured also becomes an important issue.

For the FSE datasets segmentation was readily available from previous research experiments. These segmentations were generated by a method and inspected by a neuroradiologist who manually corrected them where it was necessary to become the surrogate.

The segmentation results of our method were compared to the surrogate truth by using three accuracy measures<sup>8</sup>: *true positive volume fraction* (TPVF), *false positive volume fraction* (FPVF), and *false negative volume fraction* (FNVF).

### 4.3. Precision (Reproducibility)

Precision may be measured with respect to several factors in the segmentation method. Both intra- and inter-operator precision, as well as repeat-scan precision were measured. For each test, volume difference and overlap agreement is measured. For repeat-scan overlap measurement, registration of the two scenes is necessary. Registration transformation was found between the original intensity scenes by using a method using mutual information.<sup>9</sup> The binary objects were transformed by this transformation.

### 4.4. Efficiency

We used an internal version of the 3DVIEWNIX<sup>10</sup> software system and PCs with Intel Pentium and AMD Athlon processors to do all image processing. Following the philosophy of 3DVIEWNIX, the software was written as a combination of several smaller programs, controlled by a script, instead of being a single, large, self-contained application. Due to these separate implementations, the programs communicate via files, that is, several file I/O operations are performed during the process, which adds to the processing time. If all steps are combined in a single application, these I/O operations are eliminated and the running time is reduced by several minutes in the automatic part, and by valuable seconds in the interactive step (on a modest performance PC).

Several types of efficiency measures may be considered. The running time (wall clock time) may be measured, although, it is highly dependent on what type of hardware the program is running on. Simply by switching CPU architecture (Intel, AMD, Sparc) or CPU clock frequency, wall clock running time may change substantially. We rather measure efficiency by the amount of the necessary human interaction during the process and also by the way this human input is required.

Our method has two phases. In the training phase, the user needs to interactively find the proper values of the parameters for the algorithms. This is a one time requirement for each protocol. In the real segmentation phase, user interaction is necessary when seed points and volume of interest are required and when the output of the algorithm needs to be verified and/or corrected. This is per-study interaction.

Seed points are specified by simple mouse clicks on the displayed images. The user may need to adjust the display window settings to better visualize the different tissues, although, when displaying intensity standardized images with pre-set windows, adjustments may not be necessary at all. The user slides through the axial slices for two purposes: to find good locations for the seed points, and, to specify the volume of interest, that is, to find the first and last slices containing brain tissue. During verification/correction, the user again slides through the axial slices of the original image. At this time, the corresponding slice of the binary brain mask is overlaid, and the user may cut misclassified regions from the mask or add missed regions to it, if found necessary.

## 5. RESULTS AND DISCUSSION

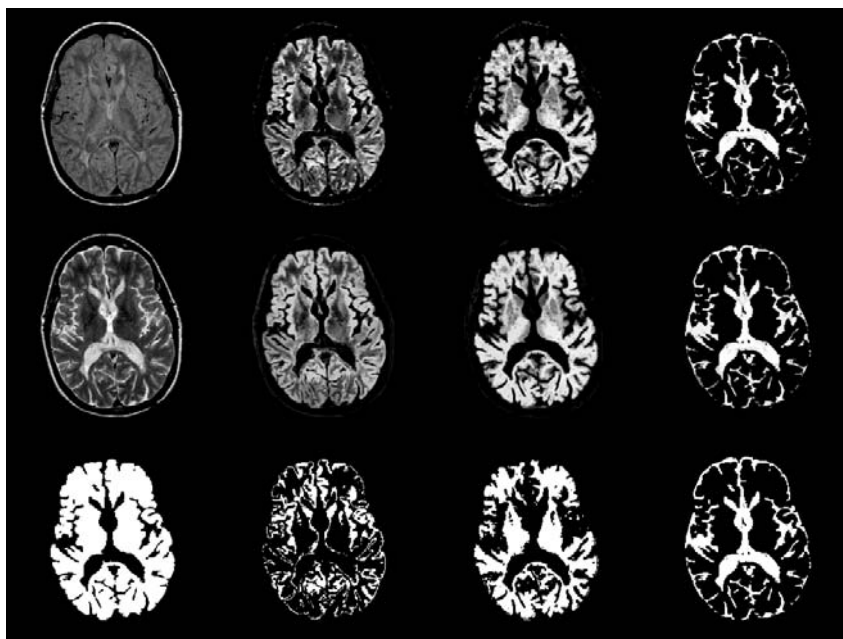
### 5.1. Images

Slices for original intensity scenes, fuzzy affinity and connectivity scenes, and binary segmentation results for different objects are displayed for each protocol in Figures 2 through 4.

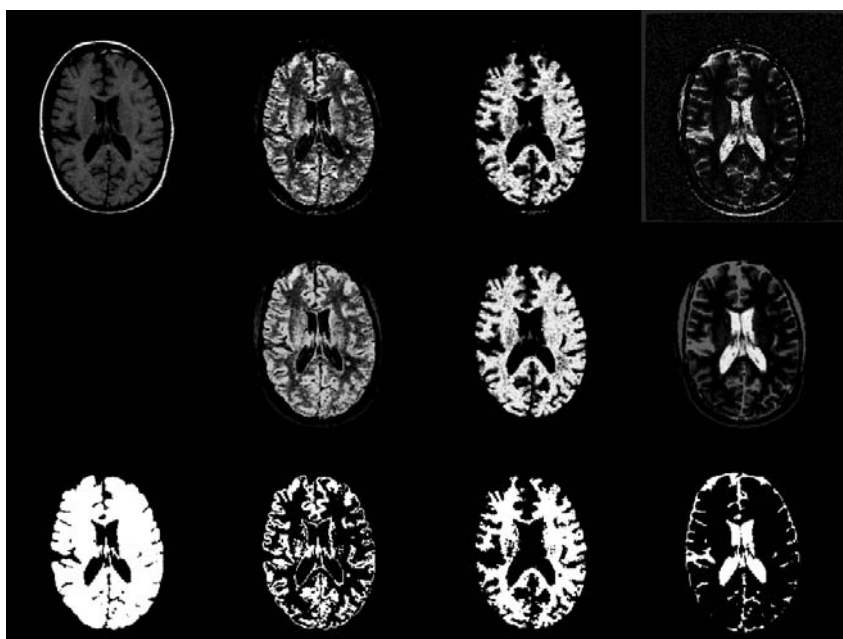
### 5.2. Accuracy

Table 1 shows FPVF, FNVF, and TPVF for segmenting brain parenchyma (BP), gray matter (GM), and white matter (WM) from FSE images. TPVF, FPVF, and FNVF are with respect to the surrogate truth and expressed in percent. Mean and standard deviation over the image population is displayed.

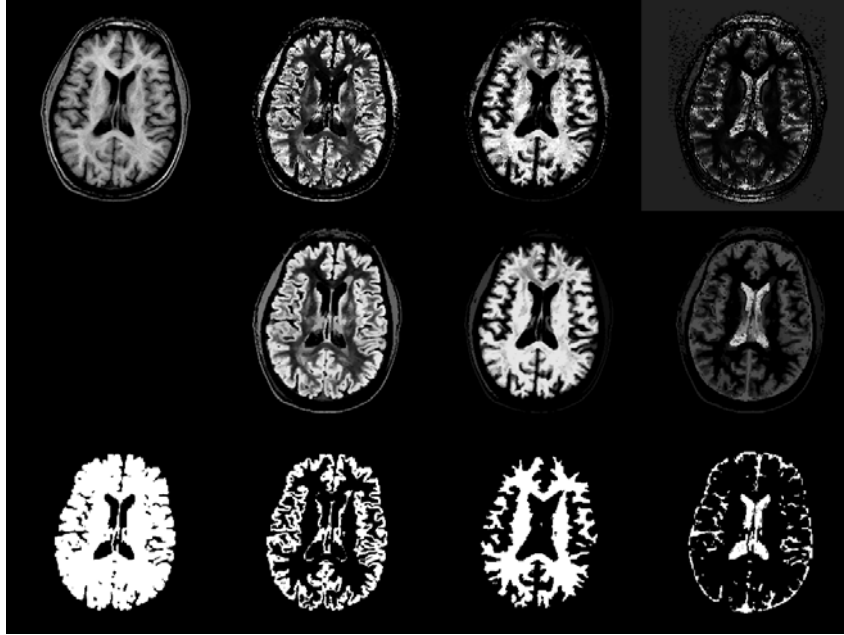
The method is highly accurate with over 99% TPVF and less than 1% FPVF for the brain parenchyma, and similar TPVF for the GM and WM objects. The slightly higher FPVF and FNVF for the WM and GM objects is due to misclassifications within the white matter near the ventricles. There is also some misclassification of GM/CSF partial volume voxels as WM around the brain parenchyma. Since this affects most slices and CSF volume is much less than that of BP, this results in a FNVF of around 3.5% for CSF.



**Figure 2.** Slices from a standardized FSE PD, T2 study pair (left column, top two images), the corresponding slices from the scenes depicting the fuzzy affinity relations for the GM, WM, and CSF objects (first row), the same slices from the scenes depicting the connectedness values (second row), and the hard (binary) segmented objects (third row). Binary mask for brain parenchyma is shown in the bottom left image.



**Figure 3.** Slices from a standardized T1 study (top left image), the corresponding slices from the scenes depicting the fuzzy affinity relations for the GM, WM, and CSF objects (first row), the same slices from the scenes depicting the connectedness values (second row), and the hard (binary) segmented objects (third row). Binary mask for brain parenchyma is shown in the bottom left image.



**Figure 4.** Slices from a standardized SPGR study (top left image), the corresponding slices from the scenes depicting the fuzzy affinity relations for the GM, WM, and CSF objects (first row), the same slices from the scenes depicting the connectedness values (second row), and the hard (binary) segmented objects (third row). Binary mask for brain parenchyma is shown in the bottom left image.

**Table 1.** Accuracy of the method for segmenting brain parenchyma (BP), gray matter (GM), and white matter (WM) from FSE images. TPVF, FPVF, and FNVF are expressed in percent. Mean and standard deviation over the image population is displayed.

		FSE	
BP	TPVF	99.99±0.01	
	FPVF	0.67±0.44	
	FNVF	0.01±0.01	
GM	TPVF	99.55±0.36	
	FPVF	0.52±0.57	
	FNVF	0.45±0.36	
WM	TPVF	99.45±0.70	
	FPVF	1.78±1.19	
	FNVF	0.55±0.70	
CSF	TPVF	96.54±2.63	
	FPVF	0.82±0.58	
	FNVF	3.46±2.63	

### 5.3. Precision (Reproducibility)

Tables 2 through 4 show repeat-scan, intra-, and inter-operator precision for segmenting brain parenchyma (BP), gray matter (GM), and white matter (WM) from images of different imaging protocols. Precision is measured between the segmentations from two scans by the same operator, two segmentations by the same operator, and segmentations by two operators. Volume difference and overlap agreements are expressed in percent. Mean and standard deviation over the image population is displayed.

Repeat-scan precision for the brain parenchyma is better than that for the separate objects. Volume disagreement is less than half percent for BP segmented from FSE and between 1 and 2% for T1 and SPGR protocols. As expected,

**Table 2.** Repeat-scan precision of the method for segmenting brain parenchyma (BP), white matter (WM), gray matter (GM), and CSF from images of different imaging protocols. Volume difference (top) and overlap agreement (bottom). Volume difference and overlap agreements are expressed in percent. Mean and standard deviation over the image population is displayed.

		FSE	T1	SPGR
volume	BP	0.36±0.34	1.66±1.58	1.16±0.42
	GM	1.84±1.41	4.46±4.30	3.84±1.76
	WM	2.62±1.23	3.34±1.31	4.37±2.63
	CSF	2.45±1.68	5.72±3.58	3.90±4.47
overlap	BP	95.53±1.75	94.56±1.11	95.25±0.31
	GM	81.73±4.87	71.52±8.12	83.42±2.90
	WM	81.23±4.09	73.05±5.84	88.37±0.29
	CSF	77.99±3.80	70.21±4.00	67.90±2.56

**Table 3.** Intra-operator precision of the method for segmenting brain parenchyma (BP), white matter (WM), gray matter (GM), and CSF from images of different imaging protocols. Volume difference (top) and overlap agreement (bottom). Volume difference and overlap agreements are expressed in percent. Mean and standard deviation over the image population is displayed.

		FSE	T1	SPGR
volume	BP	0.02±0.02	0.09±0.06	0.11±0.06
	GM	0.03±0.02	0.14±0.08	0.19±0.14
	WM	0.03±0.03	0.04±0.04	0.03±0.03
	CSF	0.60±0.53	1.64±1.16	0.72±0.34
overlap	BP	99.96±0.03	99.91±0.05	99.92±0.04
	GM	99.96±0.03	99.86±0.09	99.86±0.08
	WM	99.96±0.04	99.97±0.02	99.98±0.02
	CSF	99.50±0.39	97.96±0.76	98.96±0.27

the volume agreement is a few percent lower for GM, WM, and CSF than that for the whole brain. When segmenting FSE datasets, two scenes (proton density and T2-weighted) utilized in the segmentation, therefore, more (complementary) information is used that results in higher reproducibility. This is reflected in volume disagreements of the tissues as well as in the overlap measures.

The repeat-scan overlap for BP is around 95%. For GM, WM, and CSF this overlap is lower. It is around 80% for FSE, around 70% for T1, above 80% for GM and WM, and about 70% for CSF from the SPGR protocol. We explain the lower values in part by the registration required to measure overlap. Peripheral CSF has a fine, mostly one voxel thin structure. These voxels, when registered (even with subvoxel accuracy), will likely fall just next to their corresponding voxels in the other scene. Therefore, the overlap measure will drop.

For FSE and SPGR, GM and WM overlap is about 10% lower than that for the whole brain. This is, again, likely due to the registration error that occurs on the interface between GM and WM. In the T1 images, GM/WM contrast is much worse than that in images of the other protocols. This results in GM and WM masks that are less homogeneous and more “spotty”. After registration, these misclassified “spots” will not likely match just like the CSF structures and the GM/WM interface. Thus the even lower overlap values.

Intra- and inter-observer precision is very high and comparable, 99% and above overlap and almost negligible volume difference between compared segmentations. The only exception is CSF for which volume disagreement is slightly higher and the overlap is “only” between 97.3% and 99.5%. This is, again, explainable with the small CSF volume, compared to the other tissues, and also that the very thin peripheral CSF structures are much more than other tissues subjected to the brain mask correction that is the manual step of the method.

**Table 4.** Inter-operator precision of the method for segmenting brain parenchyma (BP), white matter (WM), gray matter (GM), and CSF from images of different imaging protocols. Volume difference (top) and overlap agreement (bottom). Volume difference and overlap agreements are expressed in percent. Mean and standard deviation over the image population is displayed.

		FSE	T1	SPGR
volume	BP	0.03±0.03	0.05±0.04	0.27±0.19
	GM	0.04±0.04	0.11±0.10	0.45±0.37
	WM	0.03±0.03	0.07±0.04	0.10±0.04
	CSF	0.47±0.37	1.94±1.73	1.88±1.04
overlap	BP	99.94±0.05	99.81±0.05	99.84±0.09
	GM	99.95±0.04	99.73±0.12	99.73±0.17
	WM	99.94±0.06	99.90±0.04	99.94±0.02
	CSF	99.28±0.41	97.32±1.13	98.48±0.24

#### 5.4. Efficiency

The following time measurements were performed on an AMD Athlon XP processor running at 1400 MHz. These measurements are highly dependent on the CPU architecture and clock frequency used. The values given below are average values, individual studies may require a little more or less, depending on the noise, artifacts in the images, as well as the severity of abnormalities in the tissues. Running time (wall clock time) for step SS1 (inhomogeneity correction) is 2 minutes per scene for FSE PD, T2, and SPGR, and it is 3 minutes per scene for T1. Running time for steps SS2-SS3 for a FSE dataset is 3 minutes, that for a T1 or SPGR dataset is less than 2 minutes. Running time for step SS6 is a few seconds and it is included within the interactive session (SS5-SS7). Running time for steps SS8-SS9 is 5 seconds (again, dependent on computer power).

Next, we give time measurements needed for the interactive sessions. The one time training time for each protocol is a few hours. It is somewhat longer for the dual-echo FSE protocol than for the others. Specifying 4 seed points for each of CSF and WM, and 8 points for GM, and selecting the top and bottom slices of the volume of interest takes 90 seconds per study. Verifying and correcting the brain mask takes about 3 minutes per study. Most corrections take place at the bottom and the top of the brain.

## 6. CONCLUDING REMARKS

We presented a new method for segmenting brain MRI that utilizes standardized MR images and scale-based relative fuzzy connectedness with a fast algorithm. The user interaction is reduced to a single interactive session. Once the appropriate parameters have been fixed, the same procedure is applicable regardless of the MRI protocol.

Segmentation of brain tissues as 3D fuzzy connected objects from MRI data is feasible at interactive speeds. Further, there is no need to devise protocol-specific segmentation strategies. Utilizing the robust fuzzy connectedness principles and fast algorithms, it is possible to interactively select fuzzy affinity, seed point, threshold parameters and perform an efficient, precise, and accurate segmentation.

## ACKNOWLEDGMENTS

The authors are thankful to Dr Robert I. Grossman for the MRI data sets utilized in this research. The first author also thanks Ms Emese Balogh for her invaluable help during the preparation of this manuscript. The authors' work is supported by DHHS Grants NS37172 and LMO-3502-MOD3.

## REFERENCES

1. P. K. Saha, J. K. Udupa, and D. Odhner, "Scale-based fuzzy connected image segmentation: Theory, algorithms, and validation," *Computer Vision and Image Understanding* **77**, pp. 145–174, 1999.

2. J. K. Udupa, L. Wei, S. Samarasekera, Y. Miki, M. A. van Buchem, and R. I. Grossman, "Multiple sclerosis lesion quantification using fuzzy-connectedness principles," *IEEE Transactions on Medical Imaging* **16**, pp. 598–609, 1997.
3. J. K. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation," *Graphical Models and Image Processing* **58**, pp. 246–261, 1996.
4. P. K. Saha and J. K. Udupa, "Relative fuzzy connectedness among multiple objects: Theory, algorithms, and applications in image segmentation," *Computer Vision and Image Understanding* **82**, pp. 42–56, 2001.
5. Y. Zhuge, J. K. Udupa, and P. K. Saha, "Vectorial fuzzy connected image segmentation," *SPIE Proceedings* **4684**, 2002. paper #170.
6. Y. Zhuge, J. K. Udupa, J. Liu, P. K. Saha, and T. Iwanaga, "Scale-based method for correcting background intensity variation in acquired images," *SPIE Proceedings* **4684**, 2002. paper #125.
7. L. G. Nyúl, J. K. Udupa, and X. Zhang, "New variants of a method of MRI scale standardization," *IEEE Transactions on Medical Imaging* **19**, pp. 143–150, 2000.
8. J. K. Udupa, V. R. LaBlanc, H. Schmid, C. Imielinska, P. K. Saha, G. J. Grevera, Y. Zhuge, P. Molholt, Y. Tin, and L. M. Currie, "A methodology for evaluating image segmentation algorithms," *SPIE Proceedings* **4684**, 2002. paper #28.
9. L. G. Nyúl, J. K. Udupa, and P. K. Saha, "Task-specific comparison of 3-D image registration methods," *SPIE Proceedings* **4322**, pp. 1588–1598, 2001.
10. J. K. Udupa, D. Odhner, S. Samarasekera, R. J. Goncalves, K. Iyer, K. Venugopal, and S. Furuie, "3DVIEWNIX: An open, transportable, multidimensional, multimodality, multiparametric imaging software system," in *SPIE Proceedings*, **2164**, pp. 58–73, 1994.