Published in final edited form as:

Neuroimage. 2007 October 15; 38(1): 114–123.

Brain tissue segmentation based on DTI data

Tianming Liu^{a,c}, Hai Li^{a,b,c}, Kelvin Wong^{a,c}, Ashley Tarokh^d, Lei Guo^b, and Stephen T.C. Wong^{a,c,*}

^aDepartment of Radiology, The Methodist Hospital, Houston, TX, USA

^bSchool of Automation, Northwestern Polytechnic University, Xi'an, China

^cCornell Weill Medical College, USA

^dFunctional and Molecular Imaging Center, Department of Radiology, Brigham and Women's Hospital, Boston, MA, USA

Abstract

We present a method for automated brain tissue segmentation based on the multi-channel fusion of diffusion tensor imaging (DTI) data. The method is motivated by the evidence that independent tissue segmentation based on DTI parametric images provides complementary information of tissue contrast to the tissue segmentation based on structural MRI data. This has important applications in defining accurate tissue maps when fusing structural data with diffusion data. In the absence of structural data, tissue segmentation based on DTI data provides an alternative means to obtain brain tissue segmentation. Our approach to the tissue segmentation based on DTI data is to classify the brain into two compartments by utilizing the tissue contrast existing in a single channel. Specifically, because the apparent diffusion coefficient (ADC) values in the cerebrospinal fluid (CSF) are more than twice that of gray matter (GM) and white matter (WM), we use ADC images to distinguish CSF and non-CSF tissues. Additionally, fractional anisotropy (FA) images are used to separate WM from non-WM tissues, as highly directional white matter structures have much larger fractional anisotropy values. Moreover, other channels to separate tissue are explored, such as eigenvalues of the tensor, relative anisotropy (RA), and volume ratio (VR). We developed an approach based on the Simultaneous Truth and Performance Level Estimation (STAPLE) algorithm that combines these two-class maps to obtain a complete tissue segmentation map of CSF, GM, and WM. Evaluations are provided to demonstrate the performance of our approach. Experimental results of applying this approach to brain tissue segmentation and deformable registration of DTI data and spoiled gradient-echo (SPGR) data are also provided.

Introduction

Brain tissue segmentation has important applications in studying the structure and function of the brain. A number of methods based on structural MRI data have been proposed for the segmentation problem (Zhang et al., 2001; Wells et al., 1996; Pham and Prince, 1999; Dale et al., 1999). In this work, we propose a robust method for automated brain tissue segmentation based on the multiple-channel fusion in DTI space. Our method can be employed to define accurate tissue maps when dealing with fused structural and diffusion data. This enables us to study the gray matter diffusivity in neurodegenerative and neurological diseases (Liu et al., 2005, 2006). When fusing structural and diffusion information, the imperfect alignment of structural MRI data, e.g., SPGR image, with DTI data results in the problem of heterogeneous voxels when the anatomic information in the structural data is applied to the DTI data. Under

*Corresponding author. Department of Radiology, The Methodist Hospital Research Institute, Houston, TX, USA. E-mail address: wong@crystal.harvard.edu (S.T.C. Wong)..

the problem of heterogeneous voxels, the measurements of the GM diffusivity based on the anatomic information in the SPGR image may fail to reveal the real diffusion in the GM. Figs. 3h and 3i in Liu et al. (2006) illustrate examples of such a problem. Specifically, following nonrigid co-registration using the UCLA AIR tools (Woods et al., 1998), the GM boundaries of SPGR image are crossing CSF of ADC image. Consequently, the GM voxels in the SPGR image correspond to CSF voxels in the ADC image. Such a problem can occur for a variety of reasons, including geometric distortion in DTI imaging (Jezzard and Balaban, 1995), partial volume effect (Helenius et al., 2002; Gonza'lez Ballester et al., 2002), reslicing and interpolation of DTI data, and errors in co-registration.

Recently, we proposed a two-channel fusion method to remove heterogeneous voxels (Liu et al., 2005, 2006). In that work, we performed tissue segmentations in both SPGR space (Zhang et al., 2001) and DTI space, and combined the results to obtain the most conservative definition of GM tissue, which is the consensus of both spaces. For example, Fig. 6 in Liu et al. (2006) shows the conservative definition of GM after heterogeneous voxels removal is performed in Fig. 3 of Liu et al., (2006). In order to perform tissue segmentation in the DTI space, the ADC image was used to distinguish the CSF and non-CSF. The technique takes advantage of the fact that the ADC values in CSF are more than twice as high as the GM and WM values. Meanwhile, the FA image was used to separate the WM from the non-WM tissues, as highly directional white matter structures have much larger fractional anisotropy values.

Our prior approaches employ only ADC and FA channels to classify brain tissues (Liu et al., 2005, 2006), and in this work, our method of tissue segmentation in DTI space is further improved by using the following seven individual channels: ADC, eigenvalues (λ_1 , λ_2 , λ_3), FA, RA, and VR. The Simultaneous Truth and Performance Level Estimation (STAPLE) algorithm (Warfield et al., 2004) is enlisted to combine these two-class maps and obtain a complete segmentation map of CSF, GM, and WM. Extensive evaluations and comparison studies are provided to demonstrate the reasonably good performance of our improved method. Experimental results of applying the proposed method to brain tissue segmentation and deformable registration of DTI data and SPGR data are also presented.

Method

Background

Diffusion-weighted imaging (DWI) and DTI permit in vivo measures of the diffusion of water molecules in living tissues (Le Bihan, 1991). The diffusion of water molecules is typically represented by unrestricted Brownian motion, a particular tissue structure can preferentially restrict the molecular motion, which leads to anisotropic diffusion that is measured by DWI and DTI (Le Bihan, 1991; Bammer, 2003). As an approximation, the measured diffusion can be modeled as an anisotropic Gaussian, parameterized by the diffusion tensor in each voxel (Basser et al., 1994), in order to create a 3-D field of diffusion tensors. Diffusion tensor measurements provide a rich dataset from which a measure of diffusion anisotropy can be obtained in various ways through the application of mathematical formulas and recalculation of the underlying eigenvalues (Bammer, 2003; Moseley et al., 1990; Le Bihan et al., 2001; Basser and Jones, 2002). The ADC, λ_1 , λ_2 , and λ_3 channels are well suited for the measuring of overall diffusivity, whereas anisotropy can be represented by FA, RA, or VR (Sundgren et al., 2004). In Sundgren et al., 2004, it was reported that the diffusivity values of the CSF are more than double the GM and WM values. This is because water diffusion in CSF is much less restricted than those in the GM and WM tissues (Johanna et al., 2002). Therefore, we can use these four channels of images to segment CSF from non-CSF tissues. Also, because highly directional white matter structures have much larger fractional anisotropy values, the FA, RA, and VR images can be used to separate WM from non-WM tissues. Moreover, since all these images intrinsically share the same DTI space, the CSF/non-CSF and WM/non-WM images

can be combined into a complete CSF/GM/WM segmentation map in the DTI space without the need of any registration.

Overview

Our computational framework for tissue segmentation based on DTI data consists of six steps, as summarized in Fig. 1. The first step consists of pre-processing to perform eddy current correction through the use of FSL FDT tools (<http://www.fmrib.ox.ac.uk/fsl/fdt/index.html>), tensor calculation, and channel image generation using DTIStudio (<http://cmrm.med.jhmi.edu/DTIuser/DTIuser.asp>). As a result, seven channels are obtained: ADC, λ_1 , λ_2 , λ_3 , FA, RA, and VR. To reduce noise, we smooth these seven channels through an edge preserving anisotropic diffusion filter (Perona and Malik, 1990). The second step consists of segmenting the brain into CSF and non-CSF compartments by utilizing the tissue contrasts existing in the first four channels, i.e., ADC, λ_1 , λ_2 , λ_3 . The third step consists of segmenting the brain into WM and non-WM compartments using the last three channels, i.e., FA, RA, and VR, separately. In order to generate the final CSF and non-CSF map, the fourth step employs the STAPLE algorithm (Warfield et al., 2004) to fuse the tissue segmentation results of the first four channels (ADC, λ_1 , λ_2 , λ_3). The WM and non-WM map is generated in the fifth step. The final step consists of combining both CSF/non-CSF and WM/non-WM maps to obtain a complete tissue map detailing the CSF, WM, and GM segments.

Tissue segmentation based on DTI data

HMRf-EM tissue segmentation—In each individual channel outlined above, the brain is classified into two classes: either CSF and non-CSF, or WM and non-WM, depending on the channel. The Expectation-Maximization (EM) algorithm, in combination with a Hidden Markov Random Field (HMRf) model (Li, 2001), is used for the two-class tissue segmentation (Zhang et al., 2001; Liu et al., 2005, 2006). The EM model fitting and the MRF Iterated Conditional Modes (ICM) labeling in the HMRf-EM segmentation both require the selection of an initial parameter set (Zhang et al., 2001). In the literature, the k -means clustering has been widely used for automated selection of initial centroids (Zhang et al., 2001; Pham and Prince, 1999). We use this method for the initial estimation.

Multi-channel fusion

Motivation—The application of the HMRf-EM segmentation method to different channels, such as ADC and eigenvalues, results in a discrepancy in the segmentation results. Fig. 2 provides an example where the distribution of the ADC values in the entire brain along with the segmented CSF of ADC channel and the λ_3 channel are presented. Fig. 2 indicates that there are no clear boundaries between different tissues in the ADC distribution of the whole brain (shown in red). Also, the CSF segmentation results of ADC channel (shown in green) and λ_3 channel (shown in blue) are quite different. From this, we postulate that segmentations from multiple channels provide richer information, and an optimal combination of these results will be more desirable.

Due to the difficulty of obtaining or estimating a known true segmentation for real data, the performance of the segmentation from any given channel is difficult to quantify, and each channel could not be assumed to contribute equally to the combined segmentation result. To deal with this problem, we make use of the Expectation-Maximization algorithm for Simultaneous Truth and Performance Level Estimation (STAPLE) algorithm proposed in Warfield et al. (2004). The algorithm considers a collection of segmentations and for each segmentation computes a probabilistic estimate of the true segmentation and a measure of the performance level represented by that segmentation. It then proceeds to estimate the optimal combination of the segmentations in order to obtain a probabilistic estimate of the true segmentation. This is done by weighting each segmentation according to its estimated

performance level while employing a prior model for the spatial distribution of structures being segmented and spatial homogeneity constraints. Because we consider seven two-class tissue segmentation maps, the goal of our multi-channel fusion is to combine the seven binary segmentation maps to construct an improved overall segmentation map and to characterize the segmentation performance level of every channel, simultaneously.

Fusion of CSF/non-CSF segmentations using STAPLE—As outlined above, there are four binary segmentation maps (CSF/ non-CSF) associated with four channels (ADC, λ_1 , λ_2 , λ_3). Let X be the random field defined in the whole volume of N voxels, and x_i denote the configuration of each voxel i . We have

$$X = \{x_i = (x_{i1}, x_{i2}, x_{i3}, x_{i4}) | x_{ij} \in \{0, 1\}, i = 1, L, N, j = 1, L, 4\} \quad (1)$$

where 0 and 1 represent the CSF and non-CSF two tissue types. Let Y be the true segmentation map. The segmentation performance of every channel is characterized by sensitivity and specificity. Sensitivity is the relative frequency of $x_{ij} = 1$ when $Y_i = 1$. We define $p = (p_1, p_2, p_3, p_4)^T$ as a column vector, whose i th element is the sensitivity parameter for the segmentation on the i th channel. Similarly, specificity is the relative frequency of $x_{ij} = 0$ when $Y_i = 0$ for which we define the column vector of specificity values $q = (q_1, q_2, q_3, q_4)^T$ for the four channels considered (Warfield et al., 2004).

We then apply the STAPLE algorithm to estimate the true map via the maximum a posteriori (MAP) method. Specifically, we seek a map Y^* , which is an estimate of the true map Y , according to the MAP criterion:

$$\begin{aligned} Y^* &= \arg \max_{Y \in \Omega} (f(X|Y, p, q) f(Y)) \\ &= \arg \max_{Y \in \Omega} \left(\prod_i \left[\prod_j f(X_{ij}|Y_i, p_j, q_j) \right] f(Y_i) \right) \end{aligned} \quad (2)$$

where $f(Y_i)$ is the prior probability of Y_i , and a voxelwise independence assumption has been made here. Next, assuming that the segmentation maps are mutually independent, the performance level parameter, which maximizes the complete data (X, Y) log likelihood function, is given by:

$$(p_j, q_j) = \arg \max_{p_j, q_j} \ln f(X_j, Y | p_j, q_j) \quad (3)$$

To estimate the solution of Eq. (3), an EM algorithm is used. The iteration of the EM algorithm will be performed as:

$$p_j^{(k)} = \frac{\sum_{i: X_{ij}=1} W_i^{(k-1)}}{\sum_i W_i^{(k-1)}} \quad (4)$$

$$q_j^{(k)} = \frac{\sum_{i: X_{ij}=0} (1 - W_i^{(k-1)})}{\sum_i (1 - W_i^{(k-1)})} \quad (5)$$

In Eqs. (4) and (5), $W_i^{(k-1)}$ indicates the probability of the true segmentation at voxel i being equal to 1:

$$W_i^{(k-1)} \equiv f(Y_i = 1 | X_i, p^{(k-1)}, q^{(k-1)}) = \frac{m_i^{(k-1)}}{m_i^{(k-1)} + n_i^{(k-1)}} \quad (6)$$

where

$$m_i^{(k-1)} = f(Y_i = 1) \prod_{j: X_{ij}=1} p_j^{(k-1)} \prod_{j: X_{ij}=0} (1 - p_j^{(k-1)}), \quad n_i^{(k-1)} = f(Y_i = 0) \prod_{j: X_{ij}=0} q_j^{(k-1)} \prod_{j: X_{ij}=1} (1 - q_j^{(k-1)})$$

). Because the true segmentation is a binary random variable, the posterior probability $f(Y_i = 0 | X_i, p^{(k-1)}, q^{(k-1)})$ is equal to $1 - W_i^{(k-1)}$. According to Eq. (2), the fusion model can be set as:

$$Y_i = \begin{cases} 0 & \text{if } W_i < 0.5 \\ 1 & \text{if } W_i \geq 0.5 \end{cases} \quad (7)$$

In summary, there are two steps in STAPLE for multiple-channel fusion. At each iteration, the first step entails the estimation of the conditional probability of the true segmentation given the segmentation maps of four channels and previous performance parameter estimates. In the second step, estimation values of the performance parameters are updated. More details concerning implementation of the iterative algorithm can be found in Warfield et al. (2004). The fusion result is finally obtained through Eq. (7).

Fusion of WM/non-WM segmentations—First, the three channels of FA, RA, and VR images are separately segmented into WM/non-WM tissue maps using the HMRF-EM method described in the HMRF-EM tissue segmentation subsection. Then, the three tissue segmentation maps are fused into a complete WM/non-WM tissue map by using the STAPLE algorithm described in the Fusion of CSF/non-CSF segmentations using STAPLE subsection. As a result, the FA, RA, and VR channels provide an independent segmentation of the brain tissues into the WM and non-WM compartments.

Experimental results

Datasets

We have applied our tissue segmentation and multiple-channel fusion method to 10 SPGR and DTI datasets. For the SPGR imaging settings, a 1.5-Tesla GE Echospeed system was used with coronal series of contiguous spoiled gradient images. The voxel dimensions in this case are equal to: $0.9375 \times 0.9375 \times 1.5$ mm. For the DTI settings, a 1.5-Tesla GE Echospeed system was used with the following settings. Sequence: maximum gradient amplitudes: 40 mT/M; rectangular FOV: 220×165 mm; 4 mm slice thickness; 1 mm interslice distance; TE: 70 ms; TR: 2500 ms; b value: 1000 s/mm^2 . DWI images along six non-collinear directions were collected. More details about our imaging settings can be found in Kubicki et al. (2005).

The SPGR data and DTI datasets have been processed using the pre-processing methods as described in Liu et al. (2006). In particular, the SPGR image has been co-registered with the DTI images using the non-rigid registration method in the UCLA AIR package (Woods et al., 1998).

Experiment 1

In the fusion of CSF/non-CSF segmentation maps, the initial values $(p_j^{(0)}, q_j^{(0)})$ for four channels are set to (0.45, 0.98), (0.55, 0.98), (0.55, 0.98), and (0.55, 0.98), respectively, according to the volume overlaps of the segmentation maps between the SPGR image and the DTI image (Liu et al., 2006). The prior probability of $f(Y_i)$ is available in the single-channel segmentation using HMRF-EM (Zhang et al., 2001). In this experiment, we average the probability maps of the four channel segmentations with equal weight to obtain $f(Y_i)$.

As an example, Table 1 provides the results for a randomly selected case. It is clear from the results that the CSF and non-CSF segmentation results from the four channels are quite

different, e.g., the CSF volume percentages vary from 11.7% to 22.2% while the CSF percentage on SPGR data is 19.7%. The CSF percentage obtained by multi-channel fusion is 21.4%, which is very close to that of the SPGR segmentation result. Considering that tissue segmentation based on SPGR image is relatively accurate, these results indicate the relatively good performance of the proposed tissue segmentation method.

For the fusion of WM and non-WM segmentations, the WM percentages obtained by different channels are also quite different, ranging from 58.6% to 67.6%. Similarly, the multi-channel fusion renders closer result (49.5%) to the SPGR segmentation result (39.3%). It should be noted that there are relatively large gaps between the fusion result and the SPGR segmentation result. However, using the multiple-channel fusion method, these gaps can be improved.

The results in this section support our point in the Multi-channel fusion subsection that the single-channel segmentation is less reliable, and that an optimal combination of multiple channels, e.g., using the STAPLE algorithm, provides improved performance.

Experiment 2

This experiment provides an example of evaluation by visual inspection. As shown in Fig. 3, it is difficult to obtain an ideal tissue segmentation result by using only one channel in DTI space. For example, in the segmentation of CSF/non-CSF (Fig. 3a), the CSF volume agreement (please refer to Liu et al., 2006 for the definition of volume agreement) with that of SPGR segmentation for the ADC channel is 74.4%. Using our multi-channel fusion method, the CSF volume agreement increases to 95.8%. Another example is that in the segmentation of WM and non-WM regions (Fig. 3b), the WM volume agreement of the FA channel with that of SPGR segmentation is 73.5%. On the other hand, our multi-channel fusion method increases the agreement to 88.5%.

In general, the multi-channel fusion method based on DTI data obtains visually reasonable tissue segmentation results, compared to the tissue segmentation results based on SPGR data (Fig. 3). This supports our claim about the advantage of multi-channel fusion introduced in the Multi-channel fusion subsection. Our experimental result shows that it is possible to obtain a reasonably good tissue segmentation map by combining the multiple-channel segmentation results based on the DTI data only.

To compare the proposed multiple-channel fusion method with the simple voting method, Fig. 3(a) shows the voting result of ADC, λ_1 , λ_2 , and λ_3 as an example. The CSF volume percentage of the segmentation result by voting is 13.5%, which is far away from the SPGR segmentation result (19.7%). However, the multi-channel fusion method obtains more consistent segmentation result, compared to the SPGR channel segmentation.

Experiment 3

We have evaluated the tissue segmentation method based on DTI data for 10 cases. Fig. 4 shows the volume overlaps and agreements between DTI segmentation and SPGR segmentation for these 10 cases (please refer to Liu et al., 2006 for the definitions of volume overlap and volume agreement).

Note that in Fig. 4(a) the volume overlaps for the different channels are quite discrepant, with different cases having different channels that yield the highest level of overlap. For example, in the first case, the λ_3 channel has the highest overlap, whereas in the second case, the λ_2 channel has the highest overlap. This clearly elucidates the value of the multiple-channel fusion method in estimating the performance levels of different channels and in obtaining the final segmentation result based on more reliable channels. In general, the fusion result of tissue segmentation has the highest volume overlap averaged over the 10 cases, as shown in Fig. 4

(a). The mean value and the stand deviation of overlaps for four individual channel segmentations and the fusion result are 0.44 ± 0.08 , 0.48 ± 0.13 , 0.60 ± 0.08 , 0.57 ± 0.07 , and 0.63 ± 0.04 , respectively. This partly shows that the fusion strategy generates more desirable tissue segmentation results than a single channel based on DTI data, considering that tissue segmentation based on SPGR image can serve as a comparison target.

The results of volume agreement for these 10 cases are shown in Fig. 4(b). We see that the channels with maximum volume agreement vary with each case (see Fig. 4b). This again emphasizes the importance of the multiple-channel fusion method in differentiating the performance levels of different channels. Fig. 4(b) shows that the multiple-channel fusion result has the highest volume agreement averaged over the 10 cases. On average, the mean values and the standard deviations of agreements for four individual channel segmentation results and the fusion result are 0.66 ± 0.18 , 0.84 ± 0.23 , 0.88 ± 0.05 , 0.92 ± 0.08 , and 0.94 ± 0.04 , respectively.

Experiment 4

The previous experiments show that the multi-channel fusion algorithm has better results from only one channel. To demonstrate how the different channels affect the results of multi-channel fusion segmentation, we evaluate the contribution from different channels by measuring the change of the posterior probability (reference to Eq. (6)) between W_i and W_{ij} , where W_i is the posterior probability obtained from all channels according to Eq. (6), and W_{ij} is also the posterior probability obtained from all channels except channel j .

We can define the change weight as following:

$$CW_j = \frac{\sum_i |W_{ij} - W_i|}{\sum_j \sum_i |W_{ij} - W_i|} \quad (8)$$

where j indicates the channel we use for the multi-channel fusion, and i indicates the voxels of the image volume. According to Eq. (8), the channel with high change weight (CW) indicates high contribution because the posterior probability varies greatly without this channel, whereas the channel with low change weight indicates low contribution because the posterior probability varies little with or without this channel.

We have done this experiment over the 10 cases. Fig. 5 shows the change weight for four channels (ADC, λ_1 , λ_2 , λ_3). The average change weight of the posterior probability for the four channels are 0.17 ± 0.07 , 0.20 ± 0.04 , 0.43 ± 0.08 and 0.19 ± 0.10 , respectively. Based on the result in Fig. 5, it is apparent that the channel λ_2 has larger contribution to the fusion procedure in most cases. Nevertheless, the other three channels also have import contributions to the segmentation result.

Experiment 5

In this section, we compare the proposed multiple-channel fusion method with our previous method of using two only, for instance, ADC and FA channels (Liu et al., 2006), as well as with the multi-spectral segmentation algorithm in the FSL MFAST (Zhang et al., 2001). It is noted that when performing the multi-spectral segmentation, we used all of the seven channels. To quantify the comparison, Table 2 shows the volume overlaps (O_j) and volume agreements (A_j) between the segmentation maps using different algorithms with that obtained by the SPGR segmentation map. It is apparent that the results using the ADC and FA channels have much lower CSF overlaps and agreements than the multi-spectral segmentation algorithm and the proposed multiple-channel fusion method. It is noted that the volume overlaps and agreements of CSF are much lower than the results in Liu et al. (2006) because the SPGR segmentation result was used to guide the segmentation of CSF in Liu et al. (2006), but not used in this work.

Because low volume overlap and agreement in CSF would cause severe problems of heterogeneous voxels in the measurement of gray matter properties, the result in Table 2 verifies that the proposed multiple-channel fusion method is much more preferred than the tissue segmentation method based on only two channels.

For the multi-spectral segmentation algorithm, although its performance in CSF volume overlap and agreement is comparable to the proposed multiple-channel fusion method (Table 2), its performances in GM and WM volume overlaps and agreements are poor (Table 2). The reason might be that the proposed multiple-channel fusion method exploits the STAPLE algorithm to estimate the optimal combination of the segmentations in seven different channels, whereas the multi-spectral segmentation algorithm does not.

As an example, Fig. 6 shows the segmentation results using the segmentation method based on the two channels of ADC and FA, the segmentation result using the FSL MFAST multi-spectral segmentation method, and the segmentation result using the proposed multi-channel fusion method and the SPGR segmentation result. Clearly, the proposed multiple-channel fusion method generates closer segmentation result with SPGR segmentation result than the other two methods.

Application: alternative brain tissue segmentation method

Tissue segmentation based on structural data, e.g., T2-weighted image, sometimes can produce undesirable segmentation result. For example, tissue segmentation using T2-weighted image cannot accurately distinguish the putamen and the thalamus, as shown in Fig. 7(b), whereas tissue segmentation based on DTI data can achieve better results, as shown in Fig. 7(e) (The data are provided by Dr. Susumu Mori of the Johns Hopkins University. Details about imaging settings are referred to <http://cmrm.med.jhmi.edu/>). White arrows point to putamen and thalamus areas where tissue segmentation in DTI space has better results than that based on T2-weighted image. Therefore, in the absence of structure data or in the case where desirable segmentation result from structure data cannot be obtained, tissue segmentation based on DTI data via the multiple-channel fusion method provides an alternative means to obtain tissue maps of the brain.

Discussion and conclusion

We have demonstrated that the brain tissue segmentation based only on single channel in DTI space produces less reliable result, and that the multiple-channel fusion method presented in this paper can substantially improve the segmentation. There are two important aspects to the final segmentation results: (1) the performance of individual channel segmentation and (2) the assessment of individual channel segmentation and fusion of them. A variety of issues are related to the first aspect, including the performance of individual channel segmentation, e.g., DTI data quality and the segmentation method used. In the future, we will test our multiple-channel fusion method on datasets with various image resolutions and qualities. In addition, we will test different tissue segmentation methods, e.g., the methods in Wells et al. (1996), Pham and Prince (1999), Dale et al. (1999), for individual channel segmentation, and investigate how different segmentation methods affect the final result.

The second issue is addressed by the published STAPLE algorithm, which identifies the performance levels of different segmentation maps and estimates an optimal combination of the segmentation. This is fundamentally different from the voting rule (Warfield et al., 2004). In this paper, we used one fixed set of overlaps (please refer to subsection 3.2.) for the initialization of $(p_j^{(0)}, q_j^{(0)})$. In our future work, we will investigate how the initialization of

$(p_j^{(0)}, q_j^{(0)})$ will influence the final multi-channel fusion results, and how to achieve desirable initializations.

In this work, we use seven channels that are the most frequently used in the DWI/DTI image analysis, as we believe these channels provide complementary important information, e.g., the λ_1 , λ_2 , and λ_3 channels describe the diffusion of three orthogonal directions separately. However, how many channels, or what combination of these channels, is adequate to obtain a satisfactory brain tissue segmentation result would need further investigation. Another interesting issue to be investigated is how the inclusion of channels that are not intrinsically within the same space as DTI, e.g., T1- or T2-weighted channel, into the STAPLE fusion procedure will influence the final multi-channel fusion result.

The STAPLE algorithm assumes that different segmentations are conditionally independent, and then estimates an optimal combination of these segmentations. In this paper, some channels are derived from the three independent eigenvalues and are not mutually independent. How this mutual dependence between different channels influences the STAPLE fusion result remains unclear. This issue will be investigated in our future work. It is also noted that, in our application, all of the seven channels are independently segmented into two class maps, which does add randomness and independence into the final segmentation maps that are used as inputs to the STAPLE algorithm. But what degree of independence this separate two-class segmentation has added remains unclear. Future work to investigate the degree of dependence between the seven channels, as well as between the segmented maps, would be valuable.

In the absence of digital DTI phantoms, currently, we evaluate the proposed tissue segmentation method by measuring the volume agreement and volume overlap between the segmentation results in DWI/DTI space and that in SPGR space. A similar assumption made here as that in Liu et al. (2006) is that the SPGR image provides relatively reliable tissue contrast and segmentation. However, it is noted that SPGR images do not have ideal tissue contrast and cannot serve as a gold standard. Voxel-based comparison using the volume overlap measure might be inaccurate given the possible misalignment between the DTI image and SPGR image caused by a variety of reasons such as the geometric distortion in DTI imaging, the partial volume effect, the reslicing and interpolation of DTI data, and the co-registration error. In the future, evaluation and validation studies could be performed based on comparison of the automated segmentation result with those obtained by expert manual segmentation results.

In summary, we presented a brain tissue segmentation method based on DTI data. This method fuses seven two-class segmentation maps, which are generated by utilizing the tissue contrast existing in the corresponding single channel, to obtain a complete brain tissue segmentation map using the STAPLE algorithm. The tissue segmentation results for 10 test cases show that the single-channel segmentation in DTI space is less reliable, and an optimal combination of seven selected channels produces significantly improved results. The STAPLE algorithm also plays a key role in computing the probabilistic estimate of the true segmentation and a measure of the performance level represented by each segmentation.

Acknowledgments

This research was funded by a research grant to STCW by Harvard Center for Neurodegeneration and Repair, Harvard Medical School. Parts of public DTI and SPGR datasets from NAMIC were provided by the Laboratory of Neuroscience, Department of Psychiatry, Boston VA Healthcare System and Harvard Medical School, which is supported by the following grants: NIMH R01 MH50740 (Shenton), NIH K05 MH01110 (Shenton), NIMH R01 MH52807 (McCarley), NIMH R01 MH40799 (McCarley), VA Merit Awards (Shenton; McCarley), and VA Research Enhancement Award Program (REAP; McCarley). We want to express our thanks to Dr. Susumu Mori for sharing the DTIStudio software and DTI datasets, to STAPLE authors and ITK for sharing the STAPLE filter, and to FSL developers.

References

- Bammer R. Basic principles of diffusion-weighted imaging. *Eur. J. Radiol* 2003;45:169–184. [PubMed: 12595101]
- Basser PJ, Jones DK. Diffusion tensor MRI: theory, experimental design and data analysis—A technical review. *NMR Biomed* 2002;14:456–467. [PubMed: 12489095]
- Basser PJ, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J. Magn. Reson* 1994;103:247–254.
- Dale AM, Fischl, Bruce, Sereno MI. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage* 1999;9(2):179–194. [PubMed: 9931268]
- Gonza'lez Ballester MA, Zisserman A, Brady M. Estimation of the partial volume effect in MRI. *Med. Image Anal* 2002;6(4):389–405. [PubMed: 12494949]
- Helenius J, Soine L, Perkio J, Salonen O, Kangasmaki A, Kaste MR, Carano AD, Aronen HJ, Tatlisumak T. Diffusion weighted MR imaging in normal human brains in various age groups. *AJNR Am. J. Neuroradiol* 2002;23(2):194–199.. [PubMed: 11847041]
- Jezzard P, Balaban RS. Correction for geometrical distortion in echo planar images from Bo field variations. *Magn. Reson. Med* 1995;34:65–73. [PubMed: 7674900]
- Johanna H, Lauri S, et al. Diffusion-weighted MR imaging in normal human brains in various age groups. *AJNR* 2002;23:194–199. [PubMed: 11847041]
- Kubicki M, Park H, Westin CF, Nestor PG, Mulkern RV, Maier SE, Niznikiewicz M, Connor EE, Levitt JJ, Frumin M, Kikinis R, Jolesz FA, McCarley RW, Shenton ME. DTI and MTR abnormalities in schizophrenia: analysis of white matter integrity. *NeuroImage* Jul 15;2005 26(4):1109–1118. [PubMed: 15878290]
- Le Bihan D. Molecular diffusion nuclear magnetic resonance imaging. *Magn. Reson. Q* 1991;7:1–30. [PubMed: 2043461]
- Le Bihan D, Mangin JF, Poupon C, et al. Diffusion tensor imaging: concepts and applications. *Magn. Reson. Imaging* 2001;13:534–546.
- Li, SZ. Markov Random Field Modeling in Image Analysis. Springer-Verlag; New York: 2001.
- Liu, T.; Young, GS.; Chen, NK.; Huang, L.; Wong, STC. MICCAI 2005. Palm Springs; California: 2005. 76-Space analysis of gray matter diffusivity: methods and application..
- Liu T, Young G, Huang L, Chen NK, Wong STC. 76-Space analysis of grey matter diffusivity: methods and applications. *Neuro-Image* May 15;2006 31(1):51–65. [PubMed: 16434215]
- Moseley ME, Kucharczyk J, Mintorovitch J, et al. Diffusion-weighted MR imaging of acute stroke: correlation with T2-weighted and magnetic susceptibility-enhanced MR imaging in cats. *AJNR* 1990;11:423–429. [PubMed: 2161612]
- Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell* 1990;12:629–639.
- Pham DL, Prince JL. Adaptive fuzzy segmentation of magnetic resonance images. *IEEE Trans. Med. Imag* 1999;18(9):737–752.
- Sundgren PC, Dongl Q, Gómez-Hassanl D, Mukherji SK, Malyl P, Welshl R. Diffusion tensor imaging of the brain: review of clinical applications. *Neuroradiology* 2004;46(5):339–350. [PubMed: 15103435]
- Warfield SK, Zou KH, et al. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Trans. Med. Imag* 2004;23:903–921.
- Wells WM, Grimson EL, Kikinis R, Jolesz FA. Adaptive segmentation of MRI data. *IEEE Trans. Med. Imag* 1996;15:429–442.
- Woods RP, Grafton ST, Watson JDG, Sicotte NL, Mazziotta JC. Automated image registration: II. Intersubject validation of linear and nonlinear models. *J. Comput. Assist. Tomogr* 1998;22:153–165. [PubMed: 9448780]
- Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imag* 2001;20(1):45–57.

Fig. 1.

Illustration of the computational framework of tissue segmentation based on DWI/DTI data. There are six steps: (1) Pre-processing; (2) CSF/non-CSF segmentation; (3) WM/non-WM segmentation; (4) multi-channel fusion to obtain CSF/non-CSF map; (5) multi-channel fusion to obtain WM/non-WM map; and (6) combining the CSF/non-CSF and WM/non-WM maps into a complete CSF/GM/WM tissue segmentation.

Fig. 2.
The distribution of the ADC values in the whole brain and in the CSF of single-channel segmentation result using ADC and λ_3 channel separately. ADC scale is $10^{-3} \text{ mm}^2/\text{s}$.

Fig. 3. Multi-channel data fusion. (a) CSF tissue maps obtained by different methods. The left five columns are results by single-channel segmentation. The right top is the segmentation by simple voting strategy. The right bottom is the result by 7-channel fusion. (b) WM and GM tissue maps obtained by different methods. The left four columns are results by single-channel segmentation. The right two bottom columns are results obtained by 7-channel fusion.

Fig. 4.

Evaluation of the tissue segmentation method. (a) Volume overlap for individual channel segmentation result and that for fusion result. The average overlaps of four single-channel segmentations and the fusion result are 0.44 ± 0.08 , 0.48 ± 0.13 , 0.60 ± 0.08 , 0.57 ± 0.07 , and 0.63 ± 0.04 , respectively. (b) Volume agreement for individual channel segmentation result and that for fusion result. The average agreements of four single-channel segmentation results and the fusion result are 0.66 ± 0.18 , 0.84 ± 0.23 , 0.88 ± 0.05 , 0.92 ± 0.08 , and 0.94 ± 0.04 , respectively.

Fig. 5. Change weight of posterior probability of four channels (ADC, λ_1 , λ_2 , λ_3) for CSF segmentation over 10 cases. The average change weight of posterior probability for the four channels are 0.17 ± 0.07 , 0.20 ± 0.04 , 0.43 ± 0.08 , and 0.19 ± 0.10 , respectively.

Fig. 6.

(a) Tissue segmentation map using ADC and FA channels. (b) Tissue segmentation map using FSL MFAST multi-spectral segmentation algorithm. (c) Tissue segmentation map using multiple-channel fusion algorithm. (d) Tissue segmentation map using SPGR image.

Fig. 7.

Tissue segmentation using T2-weighted image and DTI images (seven channels). (a) T2-weighted image. (b) Tissue segmentation using T2 image. (c). ADC image. (d). FA image. (e). Tissue segmentation map through multi-channel fusion. White arrows point to areas where tissue segmentation in DWI/DTI space has better results than that in T2-weighted image.

Table 1

Multiple-channel fusion result

SPGR	CSF (%)		GM (%)		WM (%)		O_j (%)	A_j (%)	P_j, q_j (%)
	19.7		41.0		39.3				
DWI/DTI	CSF (%)		Non-CSF (%)						
*	ADC	11.7	88.3				45.2	74.4	50.4, 99.9
	λ_1	18.2	81.8				56.3	96.0	67.1, 99.6
	λ_2	15.5	84.5				54.3	88.0	65.6, 99.9
	λ_3	22.2	77.8				60.7	94.0	96.1, 99.9
	γ	21.4	78.6				70.5	95.8	
#	WM (%)	Non-WM (%)							
	FA	67.6	32.4				90.1	73.5	100, 93.4
	RA	62.4	37.6				87.4	77.4	100, 93.9
	VR	58.6	41.4				85.2	80.3	100, 94.3
	γ	49.5	50.5				86.3	88.5	

* marks the single-channel segmentation result, and

marks the fusion result. The third and fourth columns show the volume percentages of each tissue types in SPGR space or DWI/DTI space. The fifth column shows the volume overlap (O_j), which is determined as the tissue volume overlap ratio of each DWI/DTI channel against the corresponding tissue type in SPGR image. The sixth column shows volume agreement (A_j , please refer to Liu et al., 2006 for the definition of volume agreement). The seventh column shows the segmentation performance level of every single channel.

Table 2
Volume overlap (O_j) and volume agreement (A_j) with different algorithms

	WM		GM		CSF	
	O_j (%)	A_j (%)	O_j (%)	A_j (%)	O_j (%)	A_j (%)
ADC+FA	0.83	0.84	0.56	0.94	0.28	0.56
Multi-spectral	0.46	0.74	0.43	0.84	0.80	0.57
Multi-channel	0.68	0.89	0.64	0.85	0.63	0.94

The result is averaged over 10 cases. The third, fourth, and fifth rows show the volume overlaps and volume agreements of the segmentation results using the ADC+FA channel, using the FSL MFAST multi-spectral segmentation algorithm, and using the proposed multi-channel fusion algorithm, respectively.