Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Three-dimensional brain magnetic resonance imaging segmentation via knowledge-driven decision theory

Nishant Verma
Gautam S. Muralidhar
Alan C. Bovik
Matthew C. Cowperthwaite
Mark G. Burnett
Mia K. Markey

SPIE.

Three-dimensional brain magnetic resonance imaging segmentation via knowledge-driven decision theory

Nishant Verma,[a,b,*] Gautam S. Muralidhar,[a,c] Alan C. Bovik,[d] Matthew C. Cowperthwaite,[b] Mark G. Burnett,[b] and Mia K. Markey[a,e]
[a]University of Texas at Austin, Department of Biomedical Engineering, Austin, Texas 78712, United States
[b]St. David's HealthCare, NeuroTexas Institute, Austin, Texas 78705, United States
[c]University of Texas MD Anderson Cancer Center, Department of Diagnostic Radiology, Houston, Texas 77030, United States
[d]University of Texas at Austin, Department of Electrical and Computer Engineering, Austin, Texas 78712, United States
[e]University of Texas MD Anderson Cancer Center, Department of Imaging Physics, Houston, Texas 77030, United States

Abstract. Brain tissue segmentation on magnetic resonance (MR) imaging is a difficult task because of significant intensity overlap between the tissue classes. We present a new knowledge-driven decision theory (KDT) approach that incorporates prior information of the relative extents of intensity overlap between tissue class pairs for volumetric MR tissue segmentation. The proposed approach better handles intensity overlap between tissues without explicitly employing methods for removal of MR image corruptions (such as bias field). Adaptive tissue class priors are employed that combine probabilistic atlas maps with spatial contextual information obtained from Markov random fields to guide tissue segmentation. The energy function is minimized using a variational level-set-based framework, which has shown great promise for MR image analysis. We evaluate the proposed method on two well-established real MR datasets with expert ground-truth segmentations and compare our approach against existing segmentation methods. KDT has low-computational complexity and shows better segmentation performance than other segmentation methods evaluated using these MR datasets. [DOI: 10.1117/1.JMI.1.3.034001]

Keywords: magnetic resonance imaging; tissue segmentation; Bayesian decision theory; level set formulation; Markov random field.

Paper 14072R received Jun. 3, 2014; revised manuscript received Aug. 21, 2014; accepted for publication Sep. 10, 2014; published online Oct. 1, 2014.

1 Introduction

Magnetic resonance (MR) imaging is routinely used to obtain detailed anatomical information about patients' brains. Structural changes observed on MR imaging are clinically significant for diagnostic and treatment planning purposes for several neurological diseases.[1–3] However, because of the large amount of data collected in MR imaging, manual structural measurements (such as cortical thickness) are tedious and time intensive. This has motivated the development of computer-based tools to quantify structural changes on MR volumes that are caused by neurological disorders.

MR tissue segmentation is an important, and often prerequisite, component of any comprehensive MR image analysis. This involves classifying brain MR voxels into four classes: white matter (WM), gray matter (GM), cerebrospinal fluid (CSF), and background (BG). However, automatic tissue segmentation in brain MR images is difficult because of the presence of image corruptions such as partial volume effects and intensity inhomogeneities (or bias field). Accurate segmentation of MR images requires incorporating the contributions from such image corruptions while classifying MR voxels into the four classes.

The most common segmentation methods are probabilistic formulations that represent an MR volume with a parametric model such as finite mixture model with four Gaussian components.[4–7] Thereafter, a classification rule attributes class labels to every voxel in the MR volume. However, the presence of image corruptions greatly skews the distribution of voxel intensities in MR volumes. As a result, tissue classes have arbitrarily shaped and variable density functions of intensities in MR volumes, which are difficult to represent using an *a priori*, assumed parametric model. To overcome this, the use of more flexible parametric models has been suggested.[8–12] While flexible modeling of intensity density functions yields improved segmentation performance,[8,10,13] such methods still suffer from the specification bias of the assumed parametric models.[14] Most of the modeling errors are concentrated along the tails of the intensity distributions, which are the regions of intensity overlap between the tissue classes. Therefore, the specification bias of the assumed parametric models directly translates to errors in voxel classification. Several nonparametric approaches, such as kernel density estimation, have also been employed for modeling tissue intensity distributions in MR volumes.[15–17] They provide better flexibility in modeling arbitrary intensity distributions and show improved tissue segmentation performance.[15–17]

Besides producing arbitrarily shaped intensity density functions, the presence of intensity inhomogeneities also results in significant overlap between the intensity density functions of tissue classes. Most segmentation errors are the result of inaccurate classification of MR voxels that reside in this spectrum of intensity overlap and produce similar likelihoods of membership to multiple tissue classes. To minimize such errors, preprocessing methods are typically employed to reduce the effect of intensity inhomogeneities in MR volumes.[18,19] However, the performance of subsequent tissue segmentation becomes sensitive to the accuracy of the preprocessing methods used to

*Address all correspondence to: Nishant Verma, E-mail: vnishant@utexas.edu

0091-3286/2014/$25.00 © 2014 SPIE

remove intensity inhomogeneities from MR volumes. Moreover, the computation complexity associated with such methods is generally very high.

Accurate MR tissue segmentation requires precise modeling of tissue classes and a classification rule that takes the effects from image corruptions. In this study, we present a new three-dimensional (3-D) knowledge-driven decision theory (KDT) approach toward handling the intensity overlap across tissue classes. The approach is motivated by an observation that tissue class pairs have different relative extents of intensity overlap in MR volumes. In the presence of image corruptions (such as bias field), the intensity overlap between tissue classes increases; however, the relative proportions stay approximately the same across different MR volumes. The incorporation of intensity overlap knowledge in the segmentation model enables more accurate classification of voxels residing in the intensity overlap spectrum. In KDT, a decision theory-based objective function is minimized using a variational level-set-based approach. Variational segmentation methods have gained popularity for brain MR segmentation;[13,17] however, their performance on well-known datasets is poorly documented. This makes it difficult to establish their potential in comparison with other energy minimization strategies, such as graph cuts. In this study, we evaluate our approach using two well-established datasets from the Internet brain segmentation repository (IBSR) and compare against segmentation methods that used different energy minimization techniques. There are four main technical contributions of this work:

- We demonstrate that the relative extents of intensity overlap between tissue classes are different and follow a common trend across MR volumes.
- We propose a Bayesian decision theory framework that incorporates the knowledge on relative intensity overlaps between tissue classes to improve tissue segmentation performance.
- We present a modified methodology of incorporating Markov random field-(MRF) based class priors in tissue segmentation that illustrates better performance than the traditional methodology in proposed decision theory framework.
- We validate the potential of level-set-based energy minimization for MR tissue segmentation on two standardized datasets.

A more detailed discussion on the technical contributions of this work is provided in the Discussion section (Sec. 5).

The paper is organized as follows: Sec. 2 discusses the relationships of KDT with existing segmentation methods. Section 3 describes the proposed KDT algorithm for tissue segmentation and provides details on its numerical implementation. Section 4 evaluates the segmentation performance of KDT, compares the performance with existing methods, illustrates the significance of KDT's individual components, and performs the computational complexity analysis. Finally, Sec. 5 summarizes the technical contributions of this work and discusses the advantages and limitations of KDT.

2 Related Work

The novelty of the proposed KDT approach lies in its use of prior knowledge on relative extents of intensity overlaps between tissue classes to improve tissue segmentation performance. KDT combines several individual components in order to incorporate such prior knowledge in the segmentation framework. Some of these components have been previously employed or have similarities to methodologies adopted in other segmentation methods. Therefore, in this section, we discuss some related tissue segmentation approaches and lay down their similarities and differences with KDT. Specifically, we consider existing segmentation approaches that involved some of the following components: (a) Bayesian energy formulations for tissue segmentation, (b) modeling arbitrary intensity distributions of tissue classes, (c) tissue class priors based on probabilistic atlases and MRFs, and (d) level-set-based energy minimization.

Most existing methods have used a Bayesian maximum *a-posteriori* (MAP) formulation for tissue segmentation and minimized it using the expectation maximization (EM) algorithm. Wells et al.[20] proposed an adaptive MAP method for simultaneous MR tissue segmentation and intensity inhomogeneity estimation. Leemput et al.[6] extended this approach by using probabilistic atlases for automatic modeling of tissue classes. Marroquin et al.[4] also presented a Bayesian MAP formulation for tissue segmentation along with a variant of the EM algorithm for more efficient energy minimization. Adaptive pixon represented segmentation (APRS) method by Lin et al.[16] used a MAP formulation, but their formulation involved clusters of connected pixels (pixons) rather than individual pixels. Several other well-known segmentation approaches have also used a MAP formulation for driving tissue segmentation.[10,12,21] In contrast, the proposed KDT approach minimizes a Bayesian decision theory-based energy function with a loss matrix optimized to encode prior information on relative extents of intensity overlap between tissue classes. However, Bayesian decision theory becomes equivalent to the MAP formulation if the relative extents of intensity overlap are assumed to be equal across tissue class pairs. More comparisons between Bayesian decision theory and MAP are performed in Sec. 4.6.2.

Most of the MAP formulations of tissue segmentation have assumed a parametric Gaussian distribution of intensities within each tissue class.[6,12,16,20] As discussed earlier and also noted by Prastawa et al.,[21] intensity distributions of tissue classes show significant overlap and modeling with Gaussian distributions results in degenerate decision boundaries. As a result, some segmentation approaches have considered use of alternate parametric models for tissue intensities. Marroquin et al.[4] assumed a parametric model of spline models with a Gibbsian prior for modeling tissue classes. The constrained Gaussian mixture model (CGMM) framework by Greenspan et al.[10] utilized a mixture of a large number of Gaussian components to represent individual tissue classes. However, the intensity parameters of all Gaussian components representing each tissue class were constrained to be equal, which limit the ability of the CGMM method to model arbitrary intensity distributions of tissue classes. The proposed KDT approach does not make any model assumptions about intensity distributions within the tissue classes and models them using the nonparametric method of kernel density estimation (KDE). KDE or parzen-window estimation has been previously used in other segmentation approaches to model arbitrary intensity distributions inside tissue classes.[21–24] Awate et al.[22] developed an unsupervised tissue segmentation method that adaptively learns image-neighborhood

Markov statistics and entails estimation of intensity distributions using parzen-window estimation. KDE has also been utilized in two mean shift inspired approaches of the adaptive mean-shift (AMS) method by Mayer et al.[23] and the mean shift method with edge confidence maps (MSECM) by Jiménez-Alaniz et al.[24] Similar to KDT, these segmentation approaches require precise modeling of the arbitrary intensity distributions of tissue classes for good segmentation performance.

MRF-based contextual priors and probabilistic tissue atlases are often used for defining prior anatomical information and guide tissue segmentation.[25] Leemput et al.[26] proposed an approach that combined tissue atlases with MRF priors to define tissue priors and illustrated significant improvement in segmentation performance. A similar approach was also followed by the MPM-MAP[4] and APRS[16] segmentation methods for defining tissue class priors. Rivera et al.[12] used a modified MRF methodology involving quadratic potentials, which allowed for computation of probability estimates for voxels belonging to all tissue classes. The segmentation approach by Awate et al.[22] used tissue atlases only for initialization purposes. However, since their segmentation framework implicitly incorporated MRF-based smoothness constraints, their approach also utilizes atlases and MRF contextual priors for guiding tissue segmentation. The subvolume probabilistic atlas segmentation (SVPASEG) method by Tohka et al.[27] also utilized a MRF-based framework with tissue atlases used for dividing MR volumes into different domains. KDT uses a similar approach as the ones utilized in KVL, MPM-MAP, and APRS for defining tissue priors. The only difference between tissue priors of KDT and these studies lies in the methodologies adopted for combining tissue atlases with MRF contextual priors. KVL, MPM-MAP, and APRS combined atlases and MRF priors using fixed weightings throughout the segmentation. As a result, accurate alignment of atlas maps with MR volumes using nonrigid registration methods is essential for achieving good segmentation performance. KDT uses a slightly different methodology that initializes tissue priors using atlas maps and then iteratively superimposes them with MRF contextual priors in subsequent iterations. Because the contribution from tissue atlases reduces with progressing segmentation stages, KDT does not require accurate alignment between atlas maps and MR volumes. More comparisons between these methodologies are performed in Sec. 4.6.3.

While a level-set-based approach for energy minimization has been extensively used for segmentation of natural scene images,[28,29] its application in MR tissue segmentation has been relatively scarce. The level-set-based approach is highly flexible and enables representation of energy functions containing wide varieties of energy terms, such as local region, smoothness, and area terms. The ease of implementation also makes it an attractive framework for representing brain MRI tissue segmentation models. Some level-set-based methods have been developed for brain tissue segmentation that illustrated impressive results.[13,30–32] However, the relative value of level sets in comparison with alternate energy minimization strategies is difficult to appreciate because of poor documentation of level-set-based methods on well-established segmentation datasets. KDT minimizes a Bayesian decision theory energy function using a level-set-based framework similar to previous studies [13,30–32] and evaluates its performance on two well-established datasets.

3 Methods

Notations: we define some notations that are frequently used in this paper. Given an MR volume V defined as a function $V:\Omega \rightarrow \mathbb{R}$ on a continuous 3-D domain Ω, the goal of tissue segmentation is to partition Ω into four disjoint classes $C \in \{\mathrm{WM}, \mathrm{GM}, \mathrm{CSF}, \mathrm{BG}\}$. Any MR voxel is hence defined by its spatial location (or coordinates) $x \in \Omega$ and associated MR signal/intensity value $V(x)$. Besides the spatial image domain, we also interpret KDT in the intensity range domain. The intensity range domain for a given MR volume V is defined by the space of all possible voxel intensities $\mathcal{I} \in \mathbb{I}$, where $\mathbb{I} = [\min_{x\in\Omega}[V(x)], \max_{x\in\Omega}[V(x)]]$.

Motivation: the motivation behind our approach is the observation that the relative extents of intensity overlap between different tissue class pairs are not equal and follow a consistent trend across MR volumes. We illustrate this fact by calculating the intensity overlap areas between all tissue class pairs $k, j \in \{\mathrm{WM}, \mathrm{GM}, \mathrm{CSF}\}$ using the expert ground-truth segmentations. Intensity overlap area $\mathrm{Overlap}(C_k, C_j)$ between tissue classes C_k and C_j is defined as

$$\mathrm{Overlap}(C_k, C_j) = \int_{R_k} P(\mathcal{I}, C_j)d\mathcal{I} + \int_{R_j} P(\mathcal{I}, C_k)\mathrm{d}\mathcal{I} \quad (1)$$

where $\mathcal{I}$ denotes the voxel intensities in the MR volume; $P(\mathcal{I}, C)$ denotes the likelihood of voxel intensity $\mathcal{I}$ belonging to class C; and R_k and R_j represent the intensity ranges defined as $R_k = \{\mathcal{I}: P(\mathcal{I}, C_k) > P(\mathcal{I}, C_j), \mathcal{I} \in \mathbb{I}\}$ and $R_j = \{\mathcal{I}: P(\mathcal{I}, C_j) > P(\mathcal{I}, C_k), \mathcal{I} \in \mathbb{I}\}$, respectively [as illustrated in Fig. 1(a)]. Figure 1(b) shows the overlap areas between tissue pairs WM, GM, and CSF relative to the overlap areas between WM and CSF. The scatterplot is generated using expert ground-truth segmentations of 18 real MR volumes from the IBSR. The consistent pattern across MR volumes suggests that the extents of intensity overlap are different among tissue class pairs. In terms of magnitude, the overlap area between WM and GM is higher than the overlap area between GM and CSF and between WM and CSF. Some of the MR volumes in Fig. 1(b) contain high levels of intensity inhomogeneities, which show increased overlap areas between the tissue class pairs, such as MR volumes 3 and 10. While we have simply combined the partial overlap areas between tissue classes [Areas A and B in Fig. 1(a)] for illustrating that the intensity overlap areas are not equal, asymmetry may exist between the partial overlap areas and has been considered for investigation in our experiments. The relative magnitude of overlap areas in MR volumes is a combined effect of several factors such as the lengths of boundaries between tissue types, extent of intensity inhomogeneities, partial volume effects, and contrast between the tissue types.

3.1 *Knowledge-Driven Decision Theory*

Noting this observation, we now formally present the KDT algorithm for MR tissue segmentation. We use a Bayesian decision theory framework for integrating knowledge of the relative extents of intensity overlap between tissue class pairs. A loss matrix L is defined, where each element $L_{k,j}$ represents the loss incurred if a voxel from tissue class C_k is classified as belonging to class C_j. Therefore, the total expected loss $\mathbb{E}$ from classification of voxels $x \in \Omega$ can be defined as

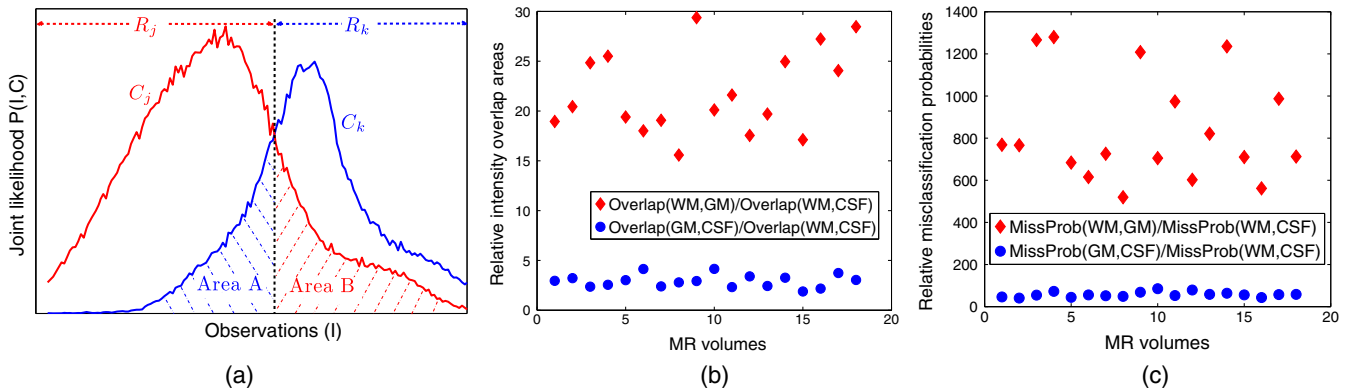

Fig. 1 Relative extents of intensity overlap between tissue classes: (a) schematic illustration showing the partial intensity overlap areas $\int_{R_j} P(\mathcal{I}, C_k)d\mathcal{I}$ = Area A and $\int_{R_k} P(\mathcal{I}, C_j)d\mathcal{I}$ = Area B, (b) scatterplot of overlap areas Overlap (WM, GM) and Overlap (GM, CSF) across MR volumes relative to Overlap (WM, CSF), and (c) scatterplot comparing the aggregate misclassification probabilities between WM and GM and GM and CSF across MR volumes [same as in (b)] relative to the total misclassification probabilities between WM and CSF. The aggregate misclassification probability between two classes C_j and C_k is defined as: $\text{MissProb}(C_j, C_k) = \int_{x\in C_j} P(x, C_k)\text{d}x + \int_{x\in C_k} P(x, C_j)\text{d}x$.

$$\mathbb{E} = \sum_k \sum_j \int_{x\in C_j} L_{k,j} \times P(x, C_k)\text{d}x, \tag{2}$$

where $P(x, C_k)$ denotes the joint likelihood of voxel, x belonging to class C_k. Decision theory has been traditionally used to determine optimum decision boundaries incurring the least expected loss in the class likelihood space based on the loss values ($L_{k,j}$ and $L_{j,k}$) and the overlap between the distributions $P(x, C_k)$ and $P(x, C_j)$.[33] Because the class distributions $P(x, C)$ for MR volumes are unknown *a priori*, decision theory has been rarely applied for MR tissue segmentation.[34]

We utilize the expected loss $\mathbb{E}$ to iteratively influence the decision boundaries such that the final voxel classification produces an intensity overlap similar to Fig. 1(b). The energy function (2) can be interpreted as a weighted sum of the intensity overlap areas between the tissue class pairs. To understand this, it is important to note the relationship between the two terms: (1) $\int_{x\in C_j} P(x, C_k)\text{d}x$ in Eq. (2) measuring the aggregate probability of misclassification of voxels belonging to class C_k into class C_j, and (2) $\int_{\mathcal{I}\in R_j} P(\mathcal{I}, C_k)\text{d}\mathcal{I}$ in Eq. (1) measuring the partial overlap area (in likelihood space) between classes C_k and C_j in the intensity range of C_j. While the aggregate probability term is calculated over the image region C_j and the partial overlap term is calculated over the intensity range R_j, they both intrinsically measure the same underlying effect. The aggregate probability term is simply the value of partial intensity overlap area scaled with the number of voxels belonging in the overlap area. This relationship can be observed in the scatterplot [Fig. 1(c)] that shows the aggregate misclassification probabilities between class pairs WM and GM and GM and CSF relative to the aggregate misclassification probabilities between WM and CSF for the same 18 MR volumes. A comparison with the relative intensity overlaps in Fig. 1(b) shows that the relative misclassification probabilities follow a very similar trend across all MR volumes.

MR tissue segmentation based solely on intensity overlaps is sensitive to the presence of image corruptions, such as MR noise. Therefore, in KDT, we define the joint voxel likelihoods $P(x, C_k)$ using an intensity term $P(V(x)|C_k)$ and a spatial prior term $P_{C_k}(x)$,

$$\mathbb{E} = \sum_k \sum_j \int_{x\in C_j} L_{k,j} \times P(V(x)|C_k) \times P_{C_k}(x)\text{d}x, \tag{3}$$

where $V(x)$ denotes the intensity value of the MR volume at voxel location $x \in \Omega$; $P(V(x)|C_k)$ is the likelihood of MR intensity value $V(x)$ belonging to class C_k; and $P_{C_k}(x)$ denotes the prior probability of class C_k at a location x in the MR volume. In Eq. (3), tissue segmentation is primarily driven by the intensity term $P(V(x)|C_k)$ that controls the relative extents of intensity overlap between tissue class pairs. The spatial priors help to identify the tissue types and reduce KDTs sensitivity to image corruptions. The following sections provide detailed descriptions of the likelihood $P(V(x)|C_k)$ and the prior $P_{C_k}(x)$ terms.

3.1.1 *Modeling arbitrary intensity distributions of tissue classes*

$P(V(x)|C)$ is estimated by modeling the arbitrarily shaped density functions of intensities inside the classes $C \in \{\text{WM}, \text{GM}, \text{CSF}, \text{BG}\}$. Assuming parametric models for intensities results in inaccurate modeling of the tissue classes. The estimation errors are mostly concentrated along the tails of the intensity density functions, which are the major regions of intensity overlaps between the classes. Therefore, accurate modeling of the arbitrarily intensity density functions inside tissue classes is essential for KDT. We use a nonparametric method of adaptive KDE based on linear diffusion processes[35] to model the intensity distributions inside the tissue classes. Adaptive KDE is specifically selected over other KDE methods because adaptive KDE has better local adaptivity, lower sensitivity to outliers, lower boundary bias, and can handle data that are not normally distributed.[35–38] These properties become significant in MR volumes because of the nonnegative nature of intensity data, presence of outliers (such as noise and artifacts), and intensity distributions that are not normally distributed.

3.1.2 *Adaptive tissue class priors*

A combination of probabilistic atlas maps and MRF-based contextual priors is used for defining tissue class priors $P_C(x)$ in

KDT. Such tissue class priors are commonly employed to guide MR tissue segmentation and reduce sensitivity to image corruptions.[4,6] In this study, we use adaptive class priors that are initialized with atlas maps and iteratively superimposed with MRF contextual priors:

$$P_C(x,n+1) = (1-w)\times P_C(x,n) + w \times P_C^{\mathrm{MRF}}(x,n) \quad (4)$$

where $P_C^{\mathrm{MRF}}(x,n)$ denote the MRF contextual priors computed at iteration n and w is an adaptive weight that controls the contribution of $P_C^{\mathrm{MRF}}(x,n)$ in class priors $P_C(x)$ at iteration $n+1$. The MRF contextual priors are characterized using Pott's model:[39–41]

$$P_C^{\mathrm{MRF}}(x) = \exp\left\{-\sum_{p\in P}\delta[C,C(p)]\right\}/Z, \quad (5)$$

where Z is a normalizing constant; $p \in \mathcal{P}$ represent all possible cliques (set of voxels) of size two in a six-neighborhood system (in 3-D) around voxel location x; $\delta(.)$ represents the Dirac delta function; and $C(p)$ denotes the classes of voxels contained in clique p. The adaptive class priors are initialized with tissue atlas maps $P_C(x,n=0) = P_C^{\mathrm{Atlas}}(x)$ spatially aligned with the MR volume using affine registration.

The expected loss function $\mathbb{E}$ in Eq. (3) is minimized iteratively by drawing decision boundaries in the likelihood space based on the loss matrix values and the tissue class distributions $P(x,C)$. Any perturbations in the decision boundaries change the voxel classification, which in turn, change the tissue class distributions. The loss matrix is determined such that the final segmentation produces an intensity overlap profile as observed in Fig. 1(c). We can relate the energy function in Eq. (3) with the MAP classification, which is often used for MR segmentation. In MAP, the objective function is minimized by choosing the tissue classes with maximum posterior probabilities for MR voxels. This decision rule is equivalent to minimizing $\mathbb{E}$ in Eq. (3) when the same loss values are considered for all tissue misclassifications: $L_{i,j} = k$ (constant) $\forall\ i, j \neq i$ and $L_{i,i} = 0$. The equal misclassification loss values imply that all overlap areas are equally penalized, which would result in equal overlap areas between all tissue class pairs. In Sec. 4.6.2, we quantitatively evaluate the effect of unequal loss values by comparing against MAP for voxel classification.

3.2 *Energy Minimization using Level Sets Framework*

The energy function Eq. (3) is difficult to minimize in terms of the evolving image regions $C_j \in \{\mathrm{WM},\mathrm{GM},\mathrm{CSF},\mathrm{BG}\}$. A level set formulation enables representation of the regions C_j in terms of higher dimensional level set functions $\Phi{:}\Omega \to \mathbb{R}$. Each level set Φ partitions the image domain Ω into two disjoint subdomains $\Omega_1 = \{x \in \Omega{:}\Phi(x) > 0\}$ and $\Omega_2 = \{x \in \Omega{:}\Phi(x) < 0\}$. Therefore, two level sets Φ_1, Φ_2 can be simultaneously used to represent the four classes:

$$\Omega = \begin{cases} C_1(\mathrm{WM}) & \Phi_1 > 0, \Phi_2 > 0 \\ C_2(\mathrm{GM}) & \Phi_1 < 0, \Phi_2 > 0 \\ C_3(\mathrm{CSF}) & \Phi_1 > 0, \Phi_2 < 0 \\ C_4(\mathrm{BG}) & \Phi_1 < 0, \Phi_2 < 0 \end{cases}.$$

The energy function in Eq. (3) for the four classes can be written as

$$\mathbb{E} = \sum_{j=1}^{4}\int_{x\in C_j}\left\{\sum_{k=1}^{4} L_{k,j}\times P[V(x)|C_k]\times P_{C_k}(x)\right\}\mathrm{d}x. \quad (6)$$

For notational simplicity, we represent the total expected loss from classification of voxels into class C_j by $E_j = \sum_{k=1}^{4} L_{k,j}\times P(V(x)|C_k)\times P_{C_k}(x)$. Using the Heaviside function $H(\Phi)$ and the Dirac delta function $\delta(\Phi)$,

$$H(\Phi) = \begin{cases} 1 & \text{if } \Phi \geq 0 \\ 0 & \text{if } \Phi < 0 \end{cases}, \delta(\Phi) = \frac{d}{d\Phi}H(\Phi),$$

the energy function in Eq. (6) can be represented as

$$\begin{aligned}\mathbb{E}(\Phi_1,\Phi_2) = \int_{x\in\Omega}[&E_1H(\Phi_1)H(\Phi_2) + E_2H(-\Phi_1)H(\Phi_2) \\ &+ E_3H(\Phi_1)H(-\Phi_2) + E_4H(-\Phi_1)H(-\Phi_2)]\mathrm{d}x.\end{aligned}$$

$\mathbb{E}(\Phi_1,\Phi_2)$ is used as the data term in the level set energy functional $\mathbb{F}(\Phi_1,\Phi_2) = \mathbb{E}(\Phi_1,\Phi_2) + \mu\times\mathbb{R}(\Phi_1,\Phi_2)$, where $\mathbb{R}(\Phi_1,\Phi_2)$ is a regularization term with weight μ on the evolving level set functions $\Phi_1(x)$ and $\Phi_2(x)$. $\mathbb{R}(\Phi_1,\Phi_2)$ ensures smoothness of the level set functions Φ_1 and Φ_2 by penalizing the arc length of their zero level contours (tissue boundaries):

$$\mathbb{R}(\Phi_1,\Phi_2) = \int_{x\in\Omega}\delta(\Phi_1)|\nabla\Phi_1|\mathrm{d}x + \int_{x\in\Omega}\delta(\Phi_2)|\nabla\Phi_2|\mathrm{d}x.$$

Hence, the energy functional $\mathbb{F}(\Phi_1,\Phi_2)$ can be represented as

$$\begin{aligned}\mathbb{F}(\Phi_1,\Phi_2) = \int_{x\in\Omega}[&E_1H(\Phi_1)H(\Phi_2) + E_2H(-\Phi_1)H(\Phi_2) \\ &+ E_3H(\Phi_1)H(-\Phi_2) + E_4H(-\Phi_1)H(-\Phi_2) \\ &+ \mu\times\delta(\Phi_1)|\nabla\Phi_1| + \mu\times\delta(\Phi_2)|\nabla\Phi_2|]\mathrm{d}x. \quad (7)\end{aligned}$$

Minimizing the energy functional $\mathbb{F}(\Phi_1,\Phi_2)$ with respect to $\Phi_1(x)$, $\Phi_2(x)$ yields the associated Euler–Lagrange equations. Parameterizing by an artificial time $t \geq 0$ leads to the following update equations in $\Phi_1(x)$, $\Phi_2(x)$ in the steepest gradient descent direction:

$$\begin{aligned}\frac{\partial\Phi_1}{\partial t} = \delta(\Phi_1)[&\mu\times\mathrm{div}\left(\frac{\nabla\Phi_1}{|\nabla\Phi_1|}\right) + (E_2-E_1)H(\Phi_2) \\ &+ (E_4-E_3)H(-\Phi_2)], \quad (8)\end{aligned}$$

$$\begin{aligned}\frac{\partial\Phi_2}{\partial t} = \delta(\Phi_2)\Bigg[&\mu\times\mathrm{div}\left(\frac{\nabla\Phi_2}{|\nabla\Phi_2|}\right) + (E_3-E_1)H(\Phi_1) \\ &+ (E_4-E_2)H(-\Phi_1)\Bigg], \quad (9)\end{aligned}$$

where ∇ and div are the gradient and divergent operators, respectively. For more details on the derivation of the level set update equations, refer to Appendix A.

To summarize, the energy function Eq. (3) is minimized iteratively. At every iteration, the arbitrary intensity density functions of tissue classes are modeled using KDE and tissue priors are updated by the MRF spatial contextual prior calculated on the previous iteration's segmentation. The flowchart in Fig. 2(a)

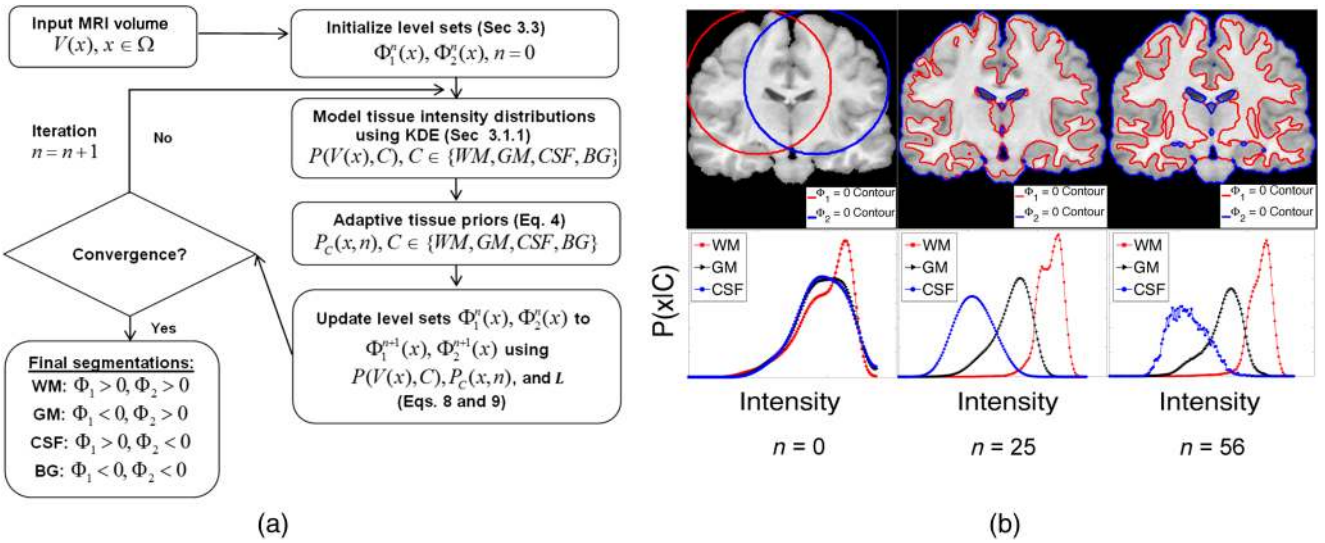

Fig. 2 Knowledge-driven decision theory (KDT) segmentation summary. (a) Flowchart summarizing the main steps of KDT tissue segmentation algorithm. (b) An illustration on the update of class intensity density functions (second row) and corresponding tissue segmentations (first row) at iterations $n = 0$, $n = 25$, and $n = 56$ (convergence). The red and blue outlines show the zero contours of the level set functions Φ_1 and Φ_2, respectively (as described in Sec. 3.2).

summarizes these main steps of KDT for energy minimization. Additionally, an example of updating intensity density functions of tissue classes is shown in Fig. 2(b) at different iterations to ultimately produce an overlap profile similar to the one observed in Fig. 1(b).

3.3 *Numerical Implementation*

For the numerical implementation of level sets, we use $C^\infty(\bar{\Omega})$ regularized versions of the Heaviside function and the Dirac delta function, denoted H_ε and δ_ε, respectively:[29] $H_\varepsilon(\Phi) = 1/2 + 1/\pi \tan^{-1}(\Phi/\epsilon)$, $\delta_\varepsilon(\Phi) = \partial H_\varepsilon(\Phi)/\partial\Phi = \epsilon/\pi(\epsilon^2/\Phi^2)$. This regularization has the tendency to compute a global minimizer without being affected by the initialization of level sets.[29] An implicit finite difference scheme is used to discretize and linearize the update Eqs. (8) and (9).[29,42] As frequently recommended in previous level set implementations,[29,43] the space step in the finite difference scheme is chosen as $h = 1$ and $\epsilon = 1$ is used to obtain the regularized functions H_ε and δ_ε. Similar to previous level set implementations, the regularization weight μ is set to the standard value of 0.1×255^2.[29,42] Because adaptive tissue priors iteratively superimpose MRF contextual prior on the atlas maps, the performance of adaptive priors is expected to be partially dependent on the time step Δt used in Eqs. (8) and (9). Therefore, besides optimizing MRF weight w, we also consider optimization of the time step to Δt obtain the optimum adaptive prior performance (Sec. 4.4). The initial level set functions $\Phi_1^0(x), \Phi_2^0(x)$ were defined as the signed distance transforms of two intersecting spherical surfaces randomly selected on the image domain Ω. The diameter of the spherical surfaces was defined to be one-eighth of the smallest dimension in the image domain Ω. As recommended,[44] the criteria for convergence of level set evolution are set as $|\Delta C(x)|/\Delta n < \tau$, where $|\Delta C(x)|$ denotes the number of voxels where the class labels change during a span of Δn iterations. In our implementation, we use the threshold $\tau = 1$ and iteration span $\Delta n = 15$.

4 Experiments and Results

4.1 *Data*

We consider two real brain MR datasets obtained from the IBSR to evaluate the segmentation performance of KDT. The first dataset (IBSR-20) contains MR volumes from 20 normal subjects along with expert ground-truth tissue segmentations. The data were collected using 1.5 Tesla T1-weighted spoiled gradient echo MRI scans on two different imaging systems with a slice thickness of 3.1 mm. The second dataset (IBSR-18) contains MR volumes from 18 normal subjects under IBSR V2.0. The data have higher spatial resolution in comparison with IBSR-20 and were collected using three Tesla T1-weighted MRI scans with a slice thickness of 1.5 mm. These datasets are established references for brain segmentation algorithm evaluation because they contain images with varying levels of difficulty, such as low contrast and high intensity inhomogeneity, to comprehensively evaluate automatic segmentation methods. We only consider real datasets in this study because simulated datasets often implicitly assume a normal distribution of tissue intensities and, therefore, exclude any analysis on intensity overlaps from arbitrary intensity density functions.

As a preprocessing step, all MR volumes in IBSR-20 and IBSR-18 datasets underwent automatic skull stripping using the brain extraction tool (BET).[45] The outputs from skull stripping were visually inspected, and any skull stripping errors were manually corrected before tissue segmentation using KDT. For adaptive tissue class priors, we used the International Consortium for Brain Mapping atlas maps provided by the Laboratory of Neuroimaging, University of California at Los Angeles.[46] The spatial alignment of atlas maps with subject MR data was performed using the FMRIB's linear image registration tool (FLIRT).[47]

4.2 *Evaluation Metrics*

We quantify the segmentation accuracy of KDT by comparing against the expert ground truth segmentations. The Dice

similarity coefficient $D(A,B)$ and the Jaccard index $J(A,B)$ are the two most commonly reported metrics in the literature for calculating the overlap between an obtained segmentation and the ground truth of each class. However, these metrics are interrelated as $J = D/(2-D)$. Therefore, we only use the Jaccard index $J(A,B)$ to assess performance in this study, as it is more intuitive for both quantitative evaluation and comparison purposes. The indices are given by $J(A,B) = |A \cap B|/|A \cup B|$ and $D(A,B) = 2|A \cap B|/(|A| + |B|)$, where A and B are the sets of voxels labeled as tissue class in KDT and the ground-truth segmentations, respectively. $|.|$ represents the cardinality of the voxel sets. The results from studies reporting only Dice coefficients were converted into their equivalent Jaccard indices to equitably compare the performances of the methods. We use a second-order Taylor expansion to approximate the mean and standard deviation of $J(A,B)$,

$$\mu_{J(A,B)} \approx \frac{\mu_{D(A,B)}}{2-\mu_{D(A,B)}} + \frac{2\sigma^2_{D(A,B)}}{(2-\mu_{D(A,B)})^3}$$

$$\sigma_{J(A,B)} \approx \frac{2\sigma_{D(A,B)}}{(2-\mu_{D(A,B)})^2}.$$

4.3 *Statistical Comparisons*

We statistically compare the segmentation performance of KDT with other competitive methods that reported accuracies on IBSR-20 and IBSR-18 datasets. Because of the paired nature of segmentation accuracies, we perform a two-sided Wilcoxon signed rank test with methods that reported the subject-wise segmentation accuracies. While such a comparison would be ideal, most of the studies did not report the subject-wise accuracies and only reported the summary statistics (mean and standard deviation) of the overlap metric, which excludes any paired statistical comparisons.

4.4 *Parameter Optimization*

The parameters that need to be optimized are the loss matrix L, the adaptive class prior weight w, and the time step Δt in the level set implementation. We randomly select a set of three MR volumes each from IBSR-18 and IBSR-20 datasets to find the optimum parameter values. We consider a range of possible values for each parameter and select the values that produce the best segmentation performance. The following ranges for the parameters are considered: {0–10} for every $L_{i,j}$ in the loss matrix L (with a step size of 1), {0, 0.05, 0.1, 0.2, 0.4, 0.5, 0.7, 1} for w and {0–1} for time step Δt (with a step size of 0.1). To simultaneously measure the segmentation performance across all tissue classes, we use voxel misclassification rate (VMR) defined as VMR where $\text{VMR} = \sum_i \sum_{j,j\neq i} |G_i \cap S_j|/|G_i|$ where G_i denotes the set of ground-truth voxels for the i-th class, S_j denotes the set of voxels classified by KDT as belonging to the j'th class, $|.|$ denotes the cardinality of the set, and $i, j \in \{\text{WM}, \text{GM}, \text{CSF}\}$.

The effects of prior weight w and time step Δt values on the segmentation performance are expected to be interdependent, whereas the loss matrix is expected to be independent from the other two variables. The independence assumption is justified because the loss matrix values in theory should be solely determined by the relative extents of intensity overlap between tissue classes. While the inclusion of spatial information using adaptive priors helps to improve segmentation accuracy, its omission should not affect the loss matrix values.

The loss matrix elements are optimized first using a small time step $\Delta t = 0.1$ and $w = 0.2$. The cost of classification into the correct class is considered as zero ($L_{i,i} = 0$), and the cost of misclassification into BG is set to 20 (very high loss since the background has already been removed using BET). While Fig. 1(c) simply combined the partial overlap areas between any two tissue classes into a single overlap area value, we consider asymmetric loss matrix to investigate any differences in partial intensity overlap areas. For optimizing the loss matrix, one of the elements ($L_{\text{WM,GM}}$) is set to 1, and others are estimated relative to this value. We further assume that $L_{\text{WM,CSF}}, L_{\text{CSF,WM}} > L_{\text{GM,CSF}}, L_{\text{CSF,GM}} > L_{\text{GM,WM}}$ based on the pattern of intensity overlap areas observed in Fig. 1(c). We find that the following asymmetric loss matrix produces the best segmentation performance:

$$L = \begin{array}{c} \\ \text{WM} \\ \text{GM} \\ \text{CSF} \end{array} \begin{array}{c} \begin{array}{ccc} \text{WM} & \text{GM} & \text{CSF} \end{array} \\ \begin{pmatrix} 0 & 1 & 10 \\ 1 & 0 & 6 \\ 10 & 9 & 0 \end{pmatrix} \end{array},$$

While GM and CSF show the differences between their partial overlap areas (GM → CSF overlap > CSF → GM overlap), the overlap distributions corresponding to WM − GM and WM − CSF tissue pairs are symmetric.

The prior weight w and time step Δt are optimized simultaneously by considering all possible combinations. We find that a combination of $w = 0.05$ and $\Delta t = 0.2$ produces the best segmentation performance across all tissue classes. Figure 3 shows the dynamics of VMR for different values of Δt and w around the optimum combination of $w = 0.05$ and $\Delta t = 0.2$. To test the validity of independence assumption, a subset of the loss matrix elements are again optimized using $w = 0.05$ and $\Delta t = 0.2$. No changes in the optimum loss matrix elements are observed that confirm that the loss matrix optimization is independent of other parameters in KDT.

4.5 *Segmentation Performance on Internet Brain Segmentation Repository Datasets*

4.5.1 *Internet brain segmentation repository-20 dataset*

Table 1 compares the overlap metrics between KDT and other methods that have reported segmentation results using IBSR-20 dataset. Some studies combined GM and CSF into a single class GM + CSF and evaluated their segmentation algorithms using overlap metrics of WM and GM + CSF classes. For a fair comparison, we compare KDT with methods belonging to both the categories: segmentation into WM, GM, CSF [Table 1(a)] and segmentation into WM, GM + CSF [Table 1(b)]. Besides Rivera et al.,[12] none of the other studies reported the subject-wise accuracies, which exclude any pairwise statistical comparisons. From comparing summary overlap statistics among the existing methods, Rivera et al.[12] produce better WM segmentation accuracy (74.20 ± 3.90) than all other existing methods.[4,10,11,13,16,23,24,27,48–52] KDT produces statistically significant improvements in WM and GM + CSF segmentation accuracies over Rivera et al.[12] ($p = 2.19 \times 10^{-4}$ for WM and $p = 8.9 \times 10^{-5}$ for GM + CSF) and, therefore, has better WM accuracy than other competitive methods as well. KDT also

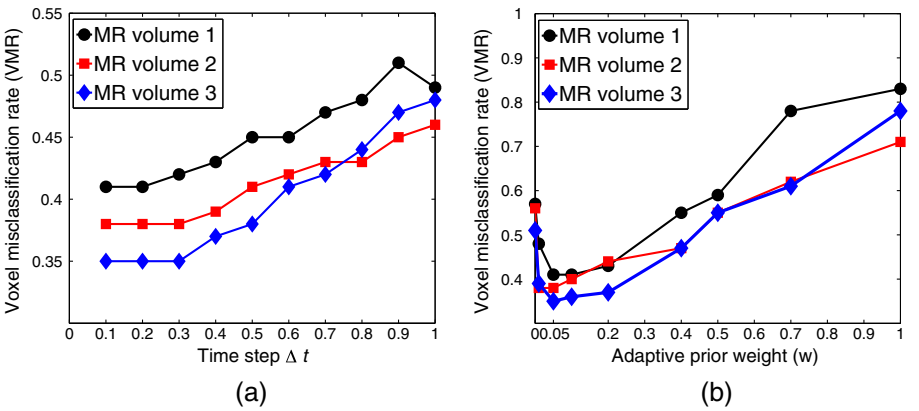

Fig. 3 Parameter optimization: plots showing the dynamics of segmentation performance against different values of (a) time step Δt (using $w = 0.05$) and (b) adaptive tissue prior weighting w (using $\Delta t = 0.2$). For clarity, we have only shown the results on three out of the six MR volumes considered for optimization.

Table 1 Segmentation performance on IBSR-20 dataset.

(a) segmentation into WM, GM, CSF

Method	J^{WM}	J^{GM}	J^{CSF}
KDT	76.98 ± 3.44	83.68 ± 2.58	72.94 ± 2.17
APRS[16]	74.10 ± 2.92	82.60 ± 2.53	70.80 ± 5.65
SPM,[48,a]	71.50 ± 3.75	79.80 ± 4.10	70.50 ± 4.32
DMC-EM[49]	69.00 ± 12.00	71.00 ± 8.00	71.00 ± 7.00
Rueda et al.[50]	70.10 ± 4.20	70.80 ± 4.50	—
Zheng et al.[51]	70.79	65.02	5.10
Dual-front[13]	67.00	73.90	—
MPM-MAP[4]	68.30	66.20	22.70
Akselrod-Ballin et al.[52]	66.85 ± 5.56	75.65 ± 6.16	28.13 ± 9.74
AMS[23]	69.10 ± 4.20	68.30 ± 3.50	—
SVPASEG[27]	68.50	69.80	—
CGMM[10]	66.00 ± 6.00	68.00 ± 4.00	—
MSECM[24,b]	62.80	59.40	21.0

(b) segmentation into WM, GM + CSF

Method	J^{WM}	J^{GM+CSF}
KDT	76.98 ± 3.44	85.86 ± 2.12
Rivera et al.[12]	74.20 ± 3.90	81.90 ± 2.80
Ibrahim et al.[11]	66.83	77.43

[a]Reported by Lin et al.[16]
[b]Mean shift with edge confidence map method by Jiménez-Alaniz et al.[24]

produces significantly better GM segmentation accuracy than most of the existing methods.[4,10,13,23,24,27,48–51] The methods by Akselrod-Ballin et al.[52] and Lin et al.[16] produce similar GM segmentation accuracy; however, KDT produces significantly better WM and CSF segmentation accuracies. When compared for CSF segmentation, KDT performs significantly better than all existing methods. If the three MR volumes selected for parameter optimization are excluded, the WM, GM, and CSF segmentation accuracies on the remaining 17 MR volumes are 76.75 ± 3.44, 83.66 ± 2.78, and 72.54 ± 1.14, respectively. The negligible differences in the summary segmentation accuracies after removal of the three MR volumes suggest that the parameter values are not biased toward the volumes suggested for parameter optimization.

4.5.2 *Internet brain segmentation repository-18 dataset*

Table 2 shows the segmentation performance of KDT on the IBSR-18 dataset. Similar to IBSR-20, we report segmentation results for both the cases when brain tissue is segmented into all three tissue types WM, GM, CSF [Table 2(a)], and when CSF and GM classes are combined into one class GM + CSF [Table 2(b)]. KDT produces better GM and CSF segmentation accuracies than all other methods in Table 2(a). Besides three methods (Local–linear, KVPASEG, and Awate et al.[22]), Rivera et al. [12] produce better WM segmentation accuracy (78.82 ± 2.83) than all other existing methods.[5–7,10,15,48,49,53] KDT produces statistically significant improvements in the segmentation accuracies of WM and GM + CSF over Rivera et al. [12] ($p = 0.02$ for WM and $p = 2 \times 10^{-4}$ for GM + CSF) and, therefore, has better WM segmentation performance than the rest of the methods. While the local–linear,[17] KVPASEG,[27] and Awate et al.[22] methods produce similar WM segmentation, KDT produces much better average GM and CSF segmentation accuracies. When compared with methods that combined GM and CSF [Table 2(b)], KDT produces better average GM + CSF segmentation accuracy than all other methods.[5,7,12,48,53] The summary WM, GM, and CSF segmentation accuracies are 79.72 ± 2.78, 88.59 ± 1.23, and 74.39 ± 6.43 when the three MR volumes used for parameter optimization are removed, which suggests that there is no significant bias of the selected parameter values on the segmentation results.

Table 2 Segmentation performance on IBSR-18 dataset.

(a) segmentation into WM, GM, CSF

Method	J^{WM}	J^{GM}	J^{CSF}
KDT	79.93 ± 2.58	88.62 ± 1.32	74.55 ± 5.86
Akselrod-Ballin et al.[15]	76.99	75.44	70.94
DCM-EM[49]	77.00 ± 6.00	73.00 ± 13.00	62.00 ± 11.00
Local-linear[17]	79.53	84.84	20.77
RCM++[17]	77.62	81.82	17.37
KVPASEG[27]	80.31 ± 2.14	71.92 ± 3.15	—
Awate et al.[22]	79.71 ± 2.89	67.91 ± 5.99	—
CGMM[10]	73.71 ± 6.62	65.56 ± 8.18	12.65 ± 6.31
Leemput et al.[6,a]	75.04 ± 3.21	65.02 ± 6.79	9.05 ± 3.56

(b) segmentation into WM, GM + CSF

Method	J^{WM}	J^{GM+CSF}
KDT	79.93 ± 2.58	89.71 ± 1.74
Rivera et al.[12]	78.82 ± 2.83	86.17 ± 2.30
FAST[7,b]	76.77 ± 1.64	86.43 ± 1.89
RiCE[5]	76.28 ± 2.62	88.09 ± 1.36
SURFER-FCM[53,b]	76.40 ± 2.35	87.63 ± 1.34
SPM[48,b]	74.90 ± 4.32	84.08 ± 3.67

[a]Reported by Greenspan et al.[10]
[b]Reported by Roy et al.[5]

4.5.3 *Performance comparison between Internet brain segmentation repository datasets*

In comparison with IBSR-20, KDT produces better segmentation accuracies for IBSR-18 MR volumes. This is because of the higher resolution of MR volumes in IBSR-18 (less slice thickness and higher magnetic field strength) with less partial volume effects than in IBSR-20. Figure 4 shows variations in J^{WM}, J^{GM}, and J^{CSF} across subjects in IBSR-20 and IBSR-18 datasets. Some MR volumes in IBSR-18 dataset (such as subjects 15 and 16) have significantly fewer CSF voxels (smaller ventricles), which result in relatively lower J^{CSF} for those volumes (same number of misclassified voxels produce much higher reduction in Jaccard overlap values). As a result, we observe higher variability of J^{CSF} in the IBSR-18 dataset as compared with the IBSR-20 dataset. Although KDT produces consistent segmentations in both datasets with small variations, the segmentation performance slightly declines in MR volumes that contain high levels of intensity inhomogeneities (such as subjects 2, 3 in IBSR-20 and subjects 11, 13 in IBSR-18). This suggests that while KDT may have better ability in handling intensity overlaps between tissue classes, it is still sensitive to the presence of high levels of intensity inhomogeneities.

4.6 *Significance of Individual Components*

In this section, we evaluate the significance of individual components in KDT by comparing against the most commonly used alternatives. Besides the component being evaluated, all other aspects of KDT are kept exactly the same to ensure that the results truly reflect the significance of that particular component.

4.6.1 *Significance of modeling arbitrarily shaped intensity distributions*

We illustrate the significance of modeling arbitrarily shaped intensity density functions by comparing with the case when the most common parametric assumption of normal distribution of intensities is assumed inside each class. Tables 3 and 4 show comparisons between the overlap scores, sensitivity, and specificity when modeling arbitrary distributions (first row) and assuming normal distribution (second row). The tissue segmentation accuracies are significantly higher when arbitrary distributions inside tissue classes are modeled ($p = 1.53 \times 10^{-5}$ for WM, $p = 1.53 \times 10^{-5}$ for GM, and $p < 7.63 \times 10^{-6}$ for CSF). In the normal distribution case, we observe an improvement in SN^{WM}; however, both SC^{WM} and J^{WM} decrease. On the other hand, SN^{GM} and J^{GM} decrease while SC^{GM} improves. This indicates overclassification of voxels as WM and, therefore, an improvement in sensitivity is produced, although the overlap

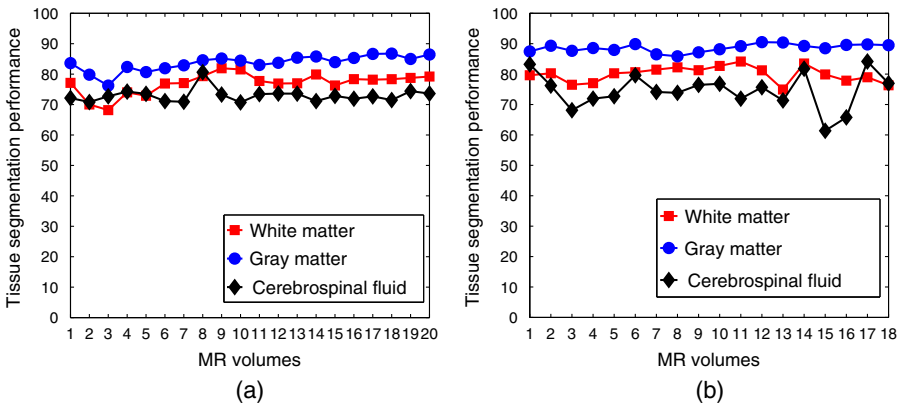

Fig. 4 Variation in KDT's segmentation performance: Plots showing the variations in J^{WM}, J^{GM}, and J^{CSF} across subjects in (a) IBSR-20 and (b) IBSR-18 datasets.

Table 3 Significance of individual components.

	Jaccard overlap indices		
	J^{WM}	J^{GM}	J^{CSF}
KDT	79.93 ± 2.58	88.62 ± 1.32	74.55 ± 5.86
Normal distance	75.11 ± 5.19	82.67 ± 2.64	69.06 ± 7.31
MAP	79.70 ± 4.08	86.88 ± 1.82	59.34 ± 7.74
Atlas	74.48 ± 5.02	83.71 ± 2.43	59.43 ± 8.55
MRF	77.72 ± 4.80	86.82 ± 2.74	72.34 ± 7.15

score and specificity suffers. Overclassification of voxels as WM is from the inaccurate estimation of intensity distributions of the tissue classes, which is crucial for the analysis of overlap areas. Figure 5 visually illustrates overclassification of voxels as WM using normal distribution (fourth column) when compared with segmentation results from KDT (third column) and ground-truth segmentations (second column).

4.6.2 *Significance of incorporating intensity overlap knowledge*

We illustrate the significance of incorporating knowledge of the relative extents of intensity overlap between tissue class pairs (loss matrix) by comparing with the case when equal loss values are considered for all tissue misclassifications (equivalent to the MAP model). Table 3 compares the performance between using optimum loss matrix (first row) and the MAP model (third row). While similar J^{WM} is observed, J^{GM} and J^{CSF} of MAP are significantly lower than KDT ($p = 3 \times 10^{-3}$ for GM, and $p < 7.63 \times 10^{-6}$ for CSF). MAP penalizes the overlap areas between tissue class pairs with equal costs and results in voxel misclassification between GM and CSF. Figure 5 visually illustrates the GM – CSF voxel misclassification in MAP (fifth column) in comparison with segmentations produced by using optimum loss matrix (third column). This illustrates the importance of incorporating knowledge regarding the relative extents of intensity overlap between tissue class pairs.

4.6.3 *Significance of adaptive class priors*

Adaptive class priors combine atlas maps with MRF contextual information to incorporate spatial information in voxel

Table 4 Significance of individual components: comparison of sensitivities and specificities.

	Sensitivity and Specificity					
	SN^{WM}	SN^{GM}	SN^{CSF}	SC^{WM}	SC^{GM}	SC^{CSF}
KDT	89.6 ± 2.3	92.0 ± 2.5	72.5 ± 5.9	92.2 ± 2.5	89.3 ± 2.3	99.98 ± 0.0
Normal Dist.	95.0 ± 1.4	83.5 ± 2.8	68.5 ± 5.5	83.9 ± 2.7	94.2 ± 1.3	99.90 ± 0.1
MAP	89.7 ± 2.4	91.6 ± 2.7	62.2 ± 5.5	91.9 ± 2.6	88.7 ± 2.5	99.84 ± 0.1
Atlas	89.6 ± 2.3	86.2 ± 3.4	64.3 ± 4.8	86.6 ± 3.3	89.1 ± 2.4	99.79 ± 0.2
MRF	89.7 ± 3.5	89.9 ± 3.9	77.7 ± 7.3	90.1 ± 3.9	89.9 ± 3.4	99.84 ± 0.1

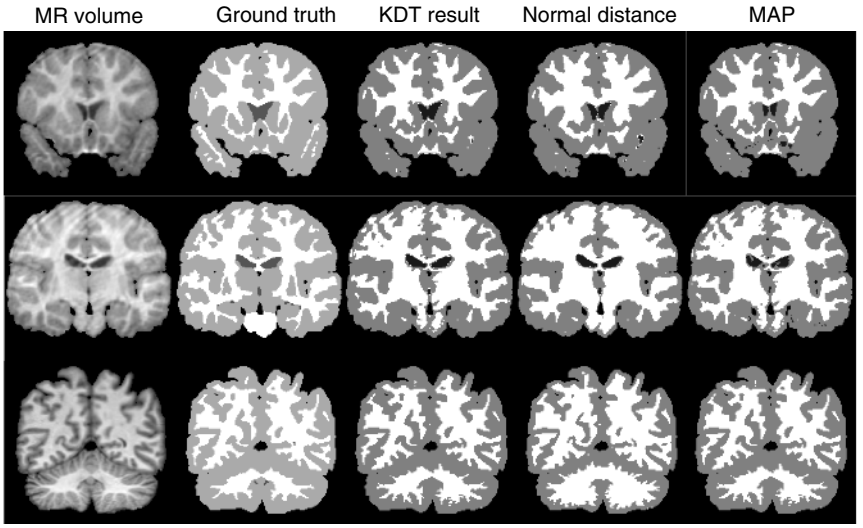

Fig. 5 Significance of individual components: visual comparisons between ground truths (second column), KDT segmentations (third column), segmentations obtained from assuming normal distribution for tissue classes (fourth column), and maximum *a-posteriori* classification (fifth column).

classification. First, we illustrate the significance of MRF contextual information in class priors by comparing with the case when atlas maps are solely utilized as class priors (fourth row). The atlas maps are aligned to the MR volumes using the 3-D nonrigid demon registration method.[54,55] The sole use of atlas maps results in significantly lower WM, GM, and CSF segmentation performance ($p < 7.63 \times 10^{-6}$ for WM, GM, and CSF) than using adaptive class priors. This reduction in segmentation performance is because of errors in alignment, which directly translates to segmentation errors. The inclusion of MRF contextual information helps to reduce the impact of errors made during the alignment of atlas maps with the MR volumes. In adaptive class priors, the atlas maps are aligned with the MR volumes using simple linear registration.[47] The application of nonrigid registration methods did not produce any statistically significant differences in the final segmentation performance. This further illustrates the significance of MRF contextual information in the class priors.

In this study, the methodology for combining MRF contextual priors with atlas maps is slightly different from the traditional way of defining MRF class priors.[4,6] Traditional methodology combined atlas maps and MRF contextual priors with fixed weightings throughout the segmentation process. As a result, accurate alignment of atlas maps with the MR volumes is essential for obtaining good tissue segmentation performance. Any alignment errors between atlas maps and MR volumes directly translate to errors in tissue segmentation. On the other hand, the methodology used in this study initializes tissue priors with atlas maps and keeps superimposing MRF contextual priors at every iteration on the tissue priors. As a result, the contribution of atlas maps reduces over the course of the segmentation iterations. Therefore, while tissue atlases still provide important prior anatomical information in the early stages of segmentation, any alignment errors do not result in final segmentation errors. In proposed decision theory framework, we found that the modified methodology is more efficient in incorporating spatial information and produces better segmentation results. We illustrate this by comparing the segmentation performance obtained using adaptive class priors with the traditional MRF class priors (fifth row in Tables 3 and 4). The atlas maps in the case of traditional MRF class priors were spatially aligned with the MR volumes using the 3-D nonrigid demon registration method.[54,55] The use of adaptive tissue class priors produces significantly better WM and GM segmentation performance than the traditional MRF class priors ($p = 1.39 \times 10^{-2}$ for WM and $p = 2.68 \times 10^{-2}$ for GM). These differences in WM and GM segmentation performances are from errors made during the alignment of atlas maps with the MR volumes, which persist throughout the segmentation. No significant difference in CSF segmentation performance is observed between traditional MRF priors and adaptive class priors. This is because of the significantly higher contrast between CSF and other tissues, which results in good segmentation performance even if the priors are inaccurately defined.

4.6.4 *Impact of loss matrix elements on segmentation*

In comparison with equal loss values for all tissue pairs (MAP model), we illustrated that the optimum loss matrix produces significantly better segmentation performance across all tissue classes (Sec. 4.6.2). Based on the relative loss values assigned to different tissue misclassification types, decisions are taken for voxels that have similar posterior probability of belonging to multiple classes. However, certain applications require higher sensitivity in segmentation of specific tissue types. Here, we take the example of computer-based support systems for diseases such as multiple sclerosis, which require high sensitivity in WM segmentation. As shown in Fig. 6, significant improvement in SN^{WM} can be obtained by simply increasing the relative loss values associated with WM misclassification.

4.7 *Robustness to Initialization of Level Set Functions*

The level set framework for energy minimization is robust to the initialization of functions Φ_1, Φ_2. To illustrate this, we evaluated the variation in KDTs segmentation performance across all tissue classes (using VMR) on a MR volume for 20 random initializations (as discussed in Sec. 3.3). We observe the mean and standard deviation of VMR to be 0.3748 and 0.021, respectively. The small variation in misclassification rate (~0.7% per brain tissue) shows that the level set framework is robust to the initialization of level set functions Φ_1, Φ_2. Similar results have also been reported by previous level-set-based segmentation methods.[13,17,29]

4.8 *Computational Complexity*

We analyze the computational complexity of KDT and compare it with other segmentation methods. Adaptive kernel density estimation [35] involves the use of fast Fourier transform for calculation of cosine and inverse cosine transforms on MR voxel intensities. Therefore, modeling class distribution has a complexity of $O(K \log K)$, where K denotes the number of MR voxels in the tissue class. In practice, $K < N/2$, where N is the total number of voxels in a MR volume. MRF calculation using the standard belief propagation for a clique size of two has a complexity of $O(NL^2)$, where L is the number of class labels ($L = 4$). The level set evolution using Eqs. (7) and (8) has a linear complexity $O(N)$. Therefore, the overall complexity of KDT is $O(N \log N)$. The number of iterations required for convergence in our numerical implementation on the IBSR data is typically around $T \sim 50 - 60$. The average physical run time for

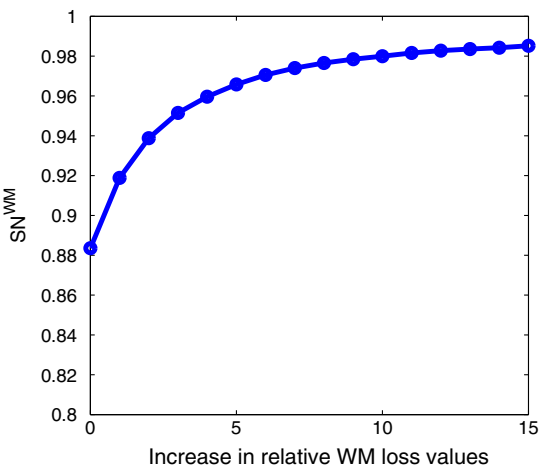

Fig. 6 Significance of individual components: plot showing the effect on SN^{WM} when the relative loss values associated with WM misclassification are increased.

segmenting a MR volume from IBSR-20 dataset (typical size $256 \times 256 \times 60$) on a Intel Core 2.7 Ghz desktop machine was 4.72 min. On IBSR-18, the average running time per volume increased to 9.27 min because of the higher resolution MR data (typical size $256 \times 256 \times 120$).

Among other segmentation methods that performed complexity analysis, KDT is one of the least computationally intensive. Local-linear method [17] has a complexity of $O(nM^2S)$ for two-dimensional (2-D) segmentation framework, where $M \sim 71 - 91$ is the window size, $n \sim 256 \times 256$ is the number of voxels per MR slice, and $S \sim 60 - 120$ is the number of slices in a MR volume. This results in physical run times of 30 min per MR volume in IBSR-18 using a Intel Core 2 Ghz machine. Ibrahim et al. [11] also reported a complexity of $O(N^2\mathrm{GL})$, where G is the number of Gaussian components and L is the sequence length. Rivera et al.[12] did not perform complexity analysis but reported physical running times of 3.2 h per MR volume for 2-D framework and 4.1 h per MR volume for their 3-D framework on a 3 GHz machine. We also compared the physical run times of KDT with the state-of-the-art FMRIB's automated segmentation tool (FAST)[7] included in the FMRIB software library. The physical run times of FAST were 7.11 and 12.38 min for IBSR-20 and IBSR-18 volumes, respectively, using the same desktop machine used for all experiments in this study. Therefore, both KDT and FAST have comparable physical run times; however, KDT produces significantly better segmentation results [Table 2(b)].

5 Discussions and Conclusions

MR tissue segmentation is a difficult task because of significant overlaps in the intensity distributions of the tissue classes. Most of the voxel classification errors occur in these regions of intensity overlap, where voxels have similar likelihoods of belonging to multiple tissue classes. To address this, the most common approach has been to correct for image corruptions that reduce the intensity overlap between tissue classes prior to tissue segmentation.[18,19] In this study, we proposed a new strategy to better deal with intensity overlaps between tissue classes without separately accounting for image corruptions. We illustrated that such a strategy produces more accurate classification of voxels belonging to intensity overlap regions in comparison with the existing methods, several of which employed methods for correction of image corruptions.

There are four main technical contributions of this work. First, we demonstrated that the relative extents of intensity overlap between tissue classes are different. The incorporation of this knowledge of the relative intensity overlaps significantly improves the tissue segmentation performance. We illustrated this (Sec. 4.6.2) by comparing the tissue segmentation performance of KDT (with optimal loss matrix) against the segmentation performance obtained using MAP, which is a specific case of KDT when all intensity overlaps are penalized with the same cost. Second, we presented a Bayesian decision theory framework KDT to incorporate the knowledge on relative intensity overlaps between tissue classes in tissue segmentation. Decision theory has been traditionally utilized to make decisions on new observations, once the class likelihood distributions are known. Because tissue distributions are unknown prior to segmentation, we utilize the Bayesian decision theory in a different manner. We exploit its ability to draw decision boundaries iteratively, such that the final location of decision boundaries produces class distributions that conform to the overlap profile as observed in Fig. 1(c).

Third, we presented a modified approach of adaptive MRF class priors for tissue segmentation. The adaptive MRF priors show better adaptivity than the traditional MRF class priors.[4,6] Adaptive MRF priors also have lower computational complexity because they do not require the use of time consuming nonrigid image registration methods for aligning patient MR volumes with the atlas maps (Sec. 4.6.3). We illustrated these benefits by comparing the tissue segmentation performances obtained using adaptive MRF priors and traditional MRF class priors, while keeping all other components of the segmentation framework the same. While adaptive class priors show significant improvements in WM and GM segmentation performances, these improvements might be specific only for the proposed decision theory framework. Therefore, further investigation of adaptive class priors incorporated in different segmentation frameworks is required to establish their significance in MR tissue segmentation.

Fourth, we illustrated that the level set approach for energy minimization is highy promising for MR segmentation. While level-set-based methods have become popular in computer vision, their application in MR tissue segmentation still remains to be validated because of lack of evaluation on standardized datasets. We evaluated the performance of KDT on two very popular datasets of real MR volumes, which have been extensively utilized for evaluating tissue segmentation methods. In comparison with methods that employed other energy minimization techniques (such as EM and graph cuts), our method using a level-set framework produced significantly better segmentation results. This demonstrates that the level-set-based framework is quite promising as a tool for minimizing complicated energy functions.

KDT performs better than most existing segmentation methods for simultaneously segmenting brain MR images into WM, GM, and CSF.[4–7,10–12,15,22,27,48,51,53] Some methods report similar segmentation performance on certain tissue types; however, they fail to perform as well on other brain tissue types.[16,17,22,27,52] KDT also performs better than the popular segmentation method FAST, which is widely used by the neuroimaging community.[7] Several of these methods involve minimizing image corruptions as part of their segmentation framework. Therefore, KDT illustrates better ability in handling intensity overlaps between tissue classes without the use of any preprocessing method to reduce MR corruptions. Besides improved segmentation performance, KDT also has one of the best computational complexities $O(N \log N)$ in comparison with other segmentation methods. For applications that require higher sensitivity in segmentation of specific tissue types, KDT also provides a very convenient framework for adapting the segmentation method by simply increasing the relative values of loss matrix elements.

Besides illustrating advantages, KDT also suffers from certain limitations. While KDT better handles intensity overlaps between the tissue classes, it is still affected by the presence of high levels of intensity inhomogeneities. This can be observed in Fig. 3, where the segmentation performance of KDT declines in MR volumes that suffer from high levels of intensity inhomogeneities. Another associated limitation of KDT is its sensitivity to the presence of partial volume effects in MR volumes. This is the reason behind the lower segmentation performance of KDT on MR volumes of IBSR-20 as compared with MR volumes of IBSR-18. The use of preprocessing steps for reducing

the effects of intensity inhomogeneities and partial volume effects can help to improve the segmentation performance in MR volumes that contain high levels of MR artifacts. However, inclusion of preprocessing steps will increase the overall computational complexity of tissue segmentation task. Moreover, KDTs segmentation performance will become highly sensitive to the performance of preprocessing steps. The need for skull and background extractions in MR volumes prior to segmentation is another limitation of KDT. The skull and other background structures often present with very similar intensity distributions as the brain tissues, which result in erroneous segmentations using KDT.

This study is limited by its strategy for comparing segmentation accuracy of KDT against the segmentation accuracies of the existing segmentation methods. Because IBSR datasets were developed to contain MR volumes with varying level of difficulties, a paired statistical test is ideal for comparing performance between segmentation methods. However, most studies only reported the summary statistics of overlap metrics, which make it impossible to perform statistical comparisons with the existing methods. Noting this limitation, we provide the subject-wise segmentation accuracies for IBSR-20 and IBSR-18 datasets in Table 5 to facilitate paired statistical comparisons in future studies.

Table 5 IBSR-20 and IBSR-18 subject-wise tissue segmentation accuracies (Jaccard indices).

IBSR20 Subjects	WM	GM	CSF	IBSR18 Subjects	WM	GM	CSF
5_8	77.12	83.61	72.13	01	79.62	87.44	83.21
4_8	70.02	79.81	70.83	02	80.26	89.30	76.21
2_4	68.17	76.23	72.71	03	76.49	87.64	68.11
6_10	74.01	82.34	74.26	04	77.01	88.56	71.92
15_3	72.89	80.67	73.53	05	80.27	87.94	72.71
16_3	76.92	81.93	71.13	06	80.62	89.86	79.62
17_3	77.05	82.89	70.91	07	81.46	86.51	74.12
8_4	79.34	84.55	80.63	08	82.22	85.85	73.83
7_8	81.92	85.13	73.29	09	81.28	87.16	76.45
110_3	81.54	84.43	70.71	10	82.68	88.17	76.81
111_2	77.74	83.01	73.43	11	84.13	89.18	71.91
112_2	76.91	83.72	73.61	12	81.23	90.50	75.67
100_23	76.98	85.39	73.58	13	74.94	90.39	71.34
202_3	79.89	85.79	71.13	14	83.49	89.25	81.77
191_3	76.27	83.93	72.84	15	79.89	88.51	61.36
12_3	78.35	85.31	71.91	16	77.85	89.59	65.76
13_3	78.18	86.64	72.68	17	78.96	89.76	84.15
1_24	78.34	86.74	71.48	18	76.32	89.50	76.96
205_3	78.76	84.98	74.41				
11_3	79.21	86.44	73.67				

Appendix A: Detailed Derivation of KDT Level-Set Framework

In MR volumes, the image domain Ω is a 3-D Cartesian grid where any location $x \in \Omega$ is defined by coordinates $x = \{x_1, x_2, x_3\}$ along the three orthogonal axes. The energy functional Eq. (7) can be rewritten as

$$\mathbb{F}(\Phi_1, \Phi_2) = \int_{x\in\Omega} G(\Phi_1, \Phi_2, \Phi_{1,1}, \Phi_{2,1}, \Phi_{1,2}, \Phi_{2,2}, \Phi_{1,3}, \Phi_{2,3}) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3, \tag{10}$$

where G is a real-valued function of level sets Φ_1, Φ_2 and their derivatives. The partial derivative of Φ_i with respect to x_j is denoted as $\Phi_{i,j}$ in the above equation.

$$\begin{aligned} G = {} & E_1 H(\Phi_1) H(\Phi_2) + E_2 H(-\Phi_1) H(\Phi_2) \\ & + E_3 H(\Phi_1) H(-\Phi_2) + E_4 H(-\Phi_1) H(-\Phi_2) \\ & + \mu \times \delta(\Phi_1) |\nabla \Phi_1| + \mu \times \delta(\Phi_2) |\nabla \Phi_2|. \end{aligned}$$

The energy functional $\mathbb{F}$ is minimized using a gradient descent method with t as an artificial time parameter:

$$\frac{\partial \Phi_i}{\partial t} = -\nabla_{\Phi_i} \mathbb{F}(\Phi_1, \Phi_2), \quad i = 1,2. \tag{11}$$

The partial derivatives of energy functional $\mathbb{F}(\Phi_1, \Phi_2)$ with respect to level sets Φ_1, Φ_2 are obtained by writing the Euler–Lagrange equations of (10):

$$\begin{aligned} & \nabla_{\Phi_1} \mathbb{F}(\Phi_1, \Phi_2) \\ & = \frac{\partial G}{\partial \Phi_1} - \frac{\partial}{\partial x_1}\left(\frac{\partial G}{\partial \Phi_{1,1}}\right) - \frac{\partial}{\partial x_2}\left(\frac{\partial G}{\partial \Phi_{1,2}}\right) - \frac{\partial}{\partial x_3}\left(\frac{\partial G}{\partial \Phi_{1,3}}\right) \\ & = \delta(\Phi_1) \times [E_1 H(\Phi_2) - E_2 H(\Phi_2) + E_3 H(-\Phi_2) \\ & \quad - E_4 H(-\Phi_2)] - \left(\frac{\partial}{\partial x_1}\left[\frac{\Phi_{1,1}}{(\Phi_{1,1}^2 + \Phi_{1,2}^2 + \Phi_{1,3}^2)^{1/2}}\right]\right. \\ & \quad + \frac{\partial}{\partial x_2}\left[\frac{\Phi_{1,2}}{(\Phi_{1,1}^2 + \Phi_{1,2}^2 + \Phi_{1,3}^2)^{1/2}}\right] \\ & \quad \left. + \frac{\partial}{\partial x_3}\left[\frac{\Phi_{1,3}}{(\Phi_{1,1}^2 + \Phi_{1,2}^2 + \Phi_{1,3}^2)^{1/2}}\right]\right) \times \mu \times \delta(\Phi_1) \\ & = \delta(\Phi_1) \times \left[(E_1 - E_2) \times H(\Phi_2) \right. \\ & \quad \left. + (E_3 - E_4) \times H(-\Phi_2) - \mu \times \mathrm{div}\left(\frac{\nabla \Phi_1}{|\nabla \Phi_1|}\right)\right] \end{aligned}$$

Similarly for the level set function Φ_2,

$$\nabla_{\Phi_2}\mathbb{F}(\Phi_1,\Phi_2) = \frac{\partial G}{\partial \Phi_2} - \frac{\partial}{\partial x_1}\left(\frac{\partial G}{\partial \Phi_{2,1}}\right) - \frac{\partial}{\partial x_2}\left(\frac{\partial G}{\partial \Phi_{2,2}}\right) - \frac{\partial}{\partial x_3}\left(\frac{\partial G}{\partial \Phi_{2,3}}\right)$$
$$= \delta(\Phi_2)\times\left[(E_1-E_3)\times H(\Phi_1) + (E_2-E_4)\times H(-\Phi_1) - \mu\times \mathrm{div}\left(\frac{\nabla\Phi_2}{|\nabla\Phi_2|}\right)\right]$$

Using these partial derivatives in Eq. (11) gives the level set update equations:

$$\frac{\partial\Phi_1}{\partial t} = \delta(\Phi_1)[\mu\times\mathrm{div}\left(\frac{\nabla\Phi_1}{|\nabla\Phi_1|}\right) + (E_2-E_1)\times H(\Phi_2) + (E_4-E_3)\times H(-\Phi_2)]$$
$$\frac{\partial\Phi_2}{\partial t} = \delta(\Phi_2)\left[\mu\times\mathrm{div}\left(\frac{\nabla\Phi_2}{|\nabla\Phi_2|}\right) + (E_3-E_1)\times H(\Phi_1) + (E_4-E_2)\times H(-\Phi_1)\right]$$

Acknowledgments

The authors would like to thank Texas 4000 for providing financial support for this study. We would also like to thank Mehul Sampat for providing valuable feedback and suggestions on the experimental evaluation performed in this study. The IBSR brain MR datasets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at: http://www.cma.mgh.harvard.edu/ibsr/.

References

1. P. A. Brex et al., "A longitudinal study of abnormalities on MRI and disability from multiple sclerosis," *N. Engl. J. Med.* **346**(3), 158–164 (2002).
2. G. B. Frisoni et al., "The clinical use of structural MRI in Alzheimer disease," *Nat. Rev. Neurol.* **6**(2), 67–77 (2010).
3. D. J. Gelb, E. Oliver, and S. Gilman, "Diagnostic criteria for Parkinson disease," *Arch. Neurol.* **56**(1), 33 (1999).
4. J. L. Marroquin et al., "An accurate and efficient Bayesian method for automatic segmentation of brain MRI," *IEEE Trans. Med. Imaging* **21**(8), 934–945 (2002).
5. S. Roy et al., "Consistent segmentation using a Rician classifier," *Med. Image Anal.* **16**(2), 524–535 (2012).
6. K. Van Leemput et al., "Automated model-based tissue classification of MR images of the brain," *IEEE Trans. Med. Imaging* **18**(10), 897–908 (1999).
7. Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Trans. Med. Imaging* **20**(1), 45–57 (2001).
8. M. B. Cuadra et al., "Comparison and validation of tissue modelization and statistical classification methods in T1-weighted MR brain images," *IEEE Trans. Med. Imaging* **24**(12), 1548–1565 (2005).
9. A. R. Ferreira da Silva, "A Dirichlet process mixture model for brain MRI tissue classification," *Med. Image Anal.* **11**(2), 169–182 (2007).
10. H. Greenspan, A. Ruf, and J. Goldberger, "Constrained Gaussian mixture model framework for automatic segmentation of MR brain images," *IEEE Trans. Med. Imaging* **25**(9), 1233–1245 (2006).
11. M. Ibrahim et al., "Hidden Markov models-based 3D MRI brain segmentation," *Image Vis. Comput.* **24**(10), 1065–1079 (2006).
12. M. Rivera, O. Ocegueda, and J. L. Marroquin, "Entropy-controlled quadratic Markov measure field models for efficient image segmentation," *IEEE Trans. Image Process.* **16**(12), 3047–3057 (2007).
13. H. Li, A. Yezzi, and L. D. Cohen, "Fast 3D brain segmentation using dual-front active contours with optional user-interaction," in *Proc. Computer Vision for Biomedical Image Applications*, pp. 335–345, Springer-Verlag, Berlin, Heidelberg (2005).
14. E. Lehmann, "Model specification: the views of Fisher and Neyman, and later developments," *Stat. Sci.* **5**(2), 160–168 (1990).
15. A. Akselrod-Ballin et al., "Prior knowledge driven multiscale segmentation of brain MRI," in *Proc. Medical Image Computing and Computer-Assisted Intervention–MICCAI*, pp. 118–126, Springer-Verlag, Berlin, Heidelberg (2007).
16. L. Lin et al., "Adaptive pixon represented segmentation (APRS) for 3D MR brain images based on mean shift and Markov random fields," *Pattern Recognit. Lett.* **32**(7), 1036–1043 (2011).
17. D. Rivest-Hénault and M. Cheriet, "Unsupervised MRI segmentation of brain tissues using a local linear model and level set," *Magn. Reson. Imaging* **29**(2), 243–259 (2011).
18. J. C. Rajapakse and F. Kruggel, "Segmentation of MR images with intensity inhomogeneities," *Image Vis. Comput.* **16**(3), 165–180 (1998).
19. M. Y. Siyal and L. Yu, "An intelligent modified fuzzy c-means based algorithm for bias estimation and segmentation of brain MRI," *Pattern Recognit. Lett.* **26**(13), 2052–2062 (2005).
20. W. M. Wells, III et al., "Adaptive segmentation of MRI data," *IEEE Trans. Med. Imaging* **15**(4), 429–442 (1996).
21. M. Prastawa et al., "Automatic segmentation of MR images of the developing newborn brain," *Med. Image Anal.* **9**(5), 457–466 (2005).
22. S. P. Awate et al., "Adaptive Markov modeling for mutual-information-based, unsupervised MRI brain-tissue classification," *Med. Image Anal.* **10**(5), 726–739 (2006).
23. A. Mayer and H. Greenspan, "An adaptive mean-shift framework for MRI brain segmentation," *IEEE Trans. Med. Imaging* **28**(8), 1238–1250 (2009).
24. J. R. Jiménez-Alaniz, V. Medina-Bañuelos, and O. Yáñez-Suárez, "Data-driven brain MRI segmentation supported on edge confidence and a priori tissue information," *IEEE Trans. Med. Imaging* **25**(1), 74–83 (2006).
25. M. Cabezas et al., "A review of atlas-based segmentation for magnetic resonance brain images," *Comput. Meth. Prog. Biomed.* **104**(3), e158–e177 (2011).
26. K. Van Leemput et al., "Automated model-based bias field correction of MR images of the brain," *IEEE Trans. Med. Imaging* **18**(10), 885–896 (1999).
27. J. Tohka et al., "Brain MRI tissue classification based on local Markov random fields," *Magn. Reson. Imaging* **28**(4), 557–573 (2010).
28. D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape," *Int. J. Comput. Vis.* **72**(2), 195–215 (2007).
29. T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.* **10**(2), 266–277 (2001).
30. L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *Int. J. Comput. Vis.* **50**(3), 271–293 (2002).
31. C. Li et al., "Minimization of region-scalable fitting energy for image segmentation," *IEEE Trans. Image Process.* **17**(10), 1940–1949 (2008).
32. C. Li et al., "A variational level set approach to segmentation and bias correction of images with intensity inhomogeneity," in *Proc. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, pp. 1083–1091, Springer-Verlag, Berlin, Heidelberg (2008).
33. J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York (1985).
34. J. L. Prince, D. Pham, and Q. Tan, "Optimization of MR pulse sequences for Bayesian image segmentation," *Med. Phys.* **22**, 1651–1656 (1995).
35. Z. Botev, J. Grotowski, and D. Kroese, "Kernel density estimation via diffusion,"*Annal. Stat.* **38**(5), 2916–2957 (2010).
36. J. S. Marron and M. P. Wand, "Exact mean integrated squared error," *Annal. Stat.* **20**(2), 712–736 (1992).

37. B. Park et al., "Adaptive variable location kernel density estimators with good performance at boundaries," *J. Nonparametr. Stat.* **15**(1), 61–75 (2003).
38. G. R. Terrell and D. W. Scott, "Variable kernel density estimation," *Annal. Stat.* **20**(3), 1236–1265 (1992).
39. J. Ashkin and E. Teller, "Statistics of two-dimensional lattices with four components," *Phys. Rev.* **64**(5–6), 178–184 (1943).
40. J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. R. Stat. Soc. Series B* **36**(2), 192–236 (1974).
41. Y. Boykov, O. Veksler, and R. Zabih, "Markov random fields with efficient approximations," in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 648–655, IEEE, Santa Barbara, California (1998).
42. C. Li et al., "Distance regularized level set evolution and its application to image segmentation," *IEEE Trans. Image Process.* **19**(12), 3243–3254 (2010).
43. B. C. Vemuri et al., "Image registration via level-set motion: applications to atlas-based segmentation," *Med. Image Anal.* **7**(1), 1–20 (2003).
44. K. N. Chaudhury and K. Ramakrishnan, "Stability and convergence of the level set method in computer vision," *Pattern Recognit. Lett.* **28**(7), 884–893 (2007).
45. S. M. Smith, "Fast robust automated brain extraction," *Hum. Brain Mapp.* **17**(3), 143–155 (2002).
46. J. Mazziotta et al., "A probabilistic atlas and reference system for the human brain: international consortium for brain mapping (ICBM)," *Philos. Trans. R. Soc. B* **356**(1412), 1293–1322 (2001).
47. M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Med. Image Anal.* **5**(2), 143–156 (2001).
48. J. Ashburner and K. J. Friston, "Voxel-based morphometry–the methods," *Neuroimage* **11**(6), 805–821 (2000).
49. M. Wels et al., "A discriminative model-constrained EM approach to 3D MRI brain tissue classification and intensity non-uniformity correction," *Phys. Med. Biol.* **56**(11), 3269 (2011).
50. A. Rueda et al., "Topology-corrected segmentation and local intensity estimates for improved partial volume classification of brain cortex in MRI," *J. Neurosci. Methods* **188**(2), 305–315 (2010).
51. B. Zheng and Z. Yi, "A new method based on the CLM of the LV RNN for brain MR image segmentation," *Digit. Signal Process.* **22**(3), 497–505 (2012).
52. A. Akselrod-Ballin et al., "Atlas guided identification of brain structures by combining 3D segmentation and SVM classification," in *Proc. Medical Image Computing and Computer-Assisted Intervention–MICCAI*, pp. 209–216, Springer-Verlag, Berlin, Heidelberg (2006).
53. A. M. Dale, B. Fischl, and M. I. Sereno, "Cortical surface-based analysis: I. Segmentation and surface reconstruction," *Neuroimage* **9**(2), 179–194 (1999).
54. X. Pennec, P. Cachier, and N. Ayache, "Understanding the demons algorithm: 3D non-rigid registration by gradient descent," in *Proc. Medical Image Computing and Computer-Assisted Intervention–MICCAI*, pp. 597–605, Springer-Verlag, Berlin, Heidelberg (1999).
55. J.-P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Med. Image Anal.* **2**(3), 243–260 (1998).

Nishant Verma is a PhD candidate and and received his MS degree in biomedical engineering at The University of Texas (UT) at Austin. He received his BTech degree in biological sciences and bioengineering from the Indian Institute of Technology (IIT) Kanpur, India, in 2010. He has been a recipient of the George J. Heuer, Jr. PhD Endowed Graduate Fellowship, Carol Lewis Heideman Endowed Presidential Fellowship, Tata Consultancy Services Award, and Proficiency Medal.

Gautam S. Muralidhar received his BE degree in electronics and communications engineering from the Visveswaraya Technological University, Belgaum, India, in July 2002, and his MSE and the PhD degrees in biomedical engineering from The University of Texas at Austin (UT), in May 2009 and December 2012, respectively. Prior to joining UT, he worked as a software engineer at Philips Healthcare, Bangalore, India, from July 2003 to July 2007.

Alan C. Bovik is the Curry/Cullen Trust Endowed Chair professor at The University of Texas at Austin and director of the Laboratory for Image and Video Engineering. He has published more than 650 articles and two U.S. patents. He was SPIE/IS&T imaging scientist of the year in 2011 and has received several IEEE signal processing society awards: Best paper award, education award, technical achievement award, and meritorious service award.

Matthew C. Cowperthwaite is director of the Center for Computational Neuroscience at NeuroTexas Institute (NTI), St. David's HealthCare, where he has been building a clinical computational biology and biomedical informatics research group for NTI. He received his PhD degree from The University of Texas at Austin in 2008, working under the supervision of Dr. Lauren A. Meyers.

Mark G. Burnett is a neurosurgeon and executive medical director of the NeuroTexas Institute, St. David's HealthCare in Austin, Texas. He graduated from Yale University and is a Fulbright scholar. He received his MD degree from the University of Maryland School of Medicine and completed his neurosurgical residency at the University of Pennsylvania. Following residency, he completed a fellowship in complex spine surgery at the Barrow Neurological Institute.

Mia K. Markey is an associate professor of biomedical engineering at The University of Texas at Austin and an adjunct associate professor of imaging physics at The University of Texas MD Anderson Cancer Center. A 1994 graduate of the Illinois Mathematics and Science Academy, she received her BS degree in computational biology (1998) from Carnegie Mellon University and her PhD degree in biomedical engineering (2002), along with a certificate in bioinformatics, from Duke University.