

A Novel Rotationally Invariant Region-Based Hidden Markov Model for Efficient 3-D Image Segmentation

Albert Huang, *Student Member, IEEE*, Rafeef Abugharbieh, *Senior Member, IEEE*, and Roger Tam

Abstract—We present a novel 3-D region-based hidden Markov model (rbHMM) for efficient unsupervised 3-D image segmentation. Our contribution is twofold. First, rbHMM employs a more efficient representation of the image data than current state-of-the-art HMM-based approaches that are based on either voxels or rectangular lattices/grids, thus resulting in a faster optimization process. Second, our proposed novel tree-structured parameter estimation algorithm for the rbHMM provides a locally optimal data labeling that is invariant to object rotation, which is a highly valuable property in segmentation tasks, especially in medical imaging where the segmentation results need to be independent of patient positioning in scanners in order to minimize methodological variability in data analysis. We demonstrate the advantages of our proposed technique over grid-based HMMs by validating on synthetic images of geometric shapes as well as both simulated and clinical brain MRI scans. For the geometric shapes data, our method produced consistently accurate segmentation results that were also invariant to object rotation. For the brain MRI data, our white matter and gray matter segmentation resulted in substantially higher robustness and accuracy levels with improved Dice similarity indices of 4.60% ($p = 0.0022$) and 7.71% ($p < 0.0001$), respectively.

Index Terms—Brain segmentation, hidden Markov models (HMMs), rotationally invariant segmentation, 3-D HMM, 3-D image segmentation.

I. INTRODUCTION

HIDDEN MARKOV models (HMMs) have over the past years been successfully used in numerous applications such as speech recognition [1], [2], document classification [3], [4], recognition of characters [5], gestures [6] and facial features [7], [8], segmentation of video [9], color images [10] and semantic images [11], [12], regional climate forecast [13], target tracking [14], protein sequence comparisons [15], and many others [16], [17]. The theory of HMMs was originally developed by Baum [18] with an underlying basic 1-D Markov chain [19].

Manuscript received July 25, 2008; revised January 15, 2010; accepted March 24, 2010. Date of publication April 26, 2010; date of current version September 17, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Lina J. Karam.

A. Huang and R. Abugharbieh are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: alberth@ece.ubc.ca; rafeef@ece.ubc.ca).

R. Tam is with the Department of Radiology, University of British Columbia, Vancouver, BC V6T 1Z3, Canada (e-mail: roger@msmri.medicine.ubc.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2010.2048965

In simple terms, a HMM is a Markov source—a conditionally independent process on a Markov chain or a Markov chain on a memoryless noisy channel [20], [21]. In a first-order system, the Markovian assumption states that the current system state at any given time depends only on the immediate preceding state with fixed transitional probabilities. In an HMM, an observation is generated as a random function of the current state probability distribution; therefore, it is not possible to determine the current state by relying only on the current observation. To estimate the model parameters, the Baum–Welch maximum-likelihood algorithm [22] is usually used. An efficient approximation, the Viterbi training [23], [24], uniquely assigns the observation to the state with the maximum *a posteriori* (MAP) probability.

To extend HMM to 2-D images, a pseudo 2-D HMM [25] has been proposed. The model is pseudo in the sense that, instead of a fully connected 2-D HMM, a hierarchy of 1-D HMMs is used. A truly 2-D HMM was developed by representing images as Markov meshes in [26]. Further extensions to 3-D images include the pseudo 3-D HMM [8], grid-based 3-D HMM [27], and block-based 3-D HMM [28]. Another popular model for 2-D and 3-D Markov processes is the Markov random field (MRF) [29], [30]. As an undirected graphical model based on a predefined local neighborhood system of sites, an MRF is a probabilistic stochastic relaxation approach that converges approximately to the MAP estimate. MRF-based approaches have been adopted for labeling of watershed segments [31], [32], region adjacency graph processing [33], rotationally invariant texture feature classification [34], and hidden MRF expectation-maximization (HMRF-EM)-based segmentation of medical images [35].

Two difficulties with the HMRF model have generally been identified. First, an HMRF requires the estimation of a normalization constant that is typically intractable to compute in order to achieve a balanced solution that fits both the observed data and the given prior distribution [30], [36]. Second, an HMRF is computationally intensive. In a typical N -ary segmentation problem, MAP estimation can only be achieved through other optimization strategies such as the iterated conditional modes algorithm [37]. HMM representation, in comparison, allows for exact MAP decoding without parameters and neighborhood system that are trained for the specific problem, and requires less cost increases when feature dimensions are increased to incorporate larger neighborhood samples [28], additional image features, or multiple imaging modality data. That said, as image resolutions increase with continuous advances in imaging technologies, there is a need for greater efficiency

in 3-D HMM-based representations and algorithms as some computational tasks would otherwise still be prohibitively expensive. Recently, Joshi *et al.* [38] derived a computationally efficient 3-D HMM parameter estimation method that iteratively decodes each row of voxels using neighboring rows in a 3-D rectangular lattice or grid. Though this approach is computationally efficient, the method is sensitive to object rotations as a striping artifact is observed in the horizontal image direction regardless of object orientation. Such pose dependency of segmentation results is undesirable for applications such as in medical imaging, where patient positioning in the scanners could potentially have a significant impact on subsequent analysis tasks.

In this paper, we propose a novel 3-D region-based HMM (rbHMM). Our approach employs a new tree-structured parameter estimation algorithm for efficient, accurate, as well as rotationally invariant 3-D image segmentation. Our 3-D rbHMM organizes volumetric images into homogeneous subregions akin to the concept of superpixels [40]. Our parameter estimation algorithm then iteratively estimates the model parameters and the MAP labeling of each region using a nonuniform representation of voxel data in a 3-D space. The representation is nonuniform in the sense that we do not make any *a priori* positioning or directionality assumptions on the data. We validate our proposed segmentation approach on synthetic 3-D geometric shapes as well as simulated and clinical T1-weighted brain magnetic resonance imaging (MRI) volumes. We demonstrate, using quantitative results, the superior performance of rbHMM in both accuracy and efficiency measures over current grid-based approach. Some early results were presented in [39], and here we present the proposed rbHMM approach with more in-depth experimental validation using real image data and comparison with the grid-based HMM.

The remainder of this paper is organized as follows. In Section II, we introduce the proposed 3-D rbHMM model. We then present an unsupervised, computationally efficient algorithm for estimating the MAP states using 1-D tree-structured Markov chains. Finally, we present extensive experimental results in Section III and conclude in Section IV.

II. METHODS

Here, we present the motivation for grouping voxels into local regions, from which our 3-D region-based hidden Markov model (rbHMM) is derived. The term “regional” has previously been used in the context of HMMs [13], [14] but generally in a geographical sense rather than based on image properties. We describe the proposed model in its general form along with the statistical assumptions made. We then present an efficient extension of the Viterbi algorithm for estimating the 3-D rbHMM parameters and the MAP states.

A. Computing Image Regions

The proposed algorithm first groups image voxels into contiguous local regions. The motivations behind this grouping is that voxels are typically a discrete representation of a continuous scene and, as such, are not independent; similar regions exhibit similar properties such as intensity or texture and therefore can be grouped so that the complexity and computational

cost are reduced from the number of voxels to the number of regions observed in the image. In this study, we apply the watershed transform [40] for simplicity. Alternative methods such as normalized cuts [41] can be employed to further incorporate texture cues.

Determining the watershed depth based on image gradient magnitude is nontrivial due to the wide variability in image data and image quality. Here, we adopt a procedure similar to the threshold selection process in most edge detectors, namely by examining the cumulative distribution function of gradient magnitudes. Instead of defining a specific gradient value, a percentage threshold T_w on the cumulative distribution is chosen to ensure oversegmentation and preservation of image structures. This ensures a threshold value that is stable with respect to the image intensity scale and is insensitive to small changes in the threshold value, making it applicable to a wider range of images.

B. rbHMM

We propose a 3-D HMM governed by the immediate neighborhood of each delineated region in a nonuniform 3-D space. The representation does not follow any *a priori* directionality assumption, and the ordering of the model nodes is data-driven. Unlike that of Li *et al.* [27], which follows a strict lexicographic order of voxels from left-to-right, top-to-bottom, and front-to-back, or that of Li *et al.* [42], which follows a center-to-boundary ordering, the proposed region-based model fundamentally allows for improved computational efficiency and invariance to object rotation. We describe the model as follows.

Given a 3-D image U of dimensions $W \times H \times D$, we denote a voxel with 3-D coordinates (x, y, z) along the X -, Y -, and Z -axes. U can therefore be described as a collection of the intensity values $I(x, y, z)$ of all of the voxels in the 3-D grid: $U = \{I(x, y, z) : 1 \leq x \leq W, 1 \leq y \leq H, 1 \leq z \leq D\}$. By dividing an image into regions, each voxel in U is assigned a region label $L(x, y, z)$. A region r_i is then defined as $r_i = \{(x, y, z) : L(x, y, z) = i\}$ with $N_i = \|r_i\|$ reflecting the size of the region. Consequently, we can describe U as a set of regions $R = \{r_i : 1 \leq i \leq N_R\}$, where $N_R = \|R\|$ is the total number of regions. In addition, we denote the neighboring adjacent regions of r_i in 3-D space as $R'_i = \{r_j : j \neq i, \exists (x, y, z) \in r_i \text{ s.t. } (x \pm 1, y, z) \text{ or } (x, y \pm 1, z) \text{ or } (x, y, z \pm 1) \in r_j\}$ and $N'_i = \|R'_i\|$ is the number of neighbors of r_i .

We can represent each r_i by using a mean regional feature vector f_i . For grayscale images such as MRIs, we can define f_i simply as the mean of the observed voxel intensities

$$f_i = \frac{\sum_{(x,y,z) \in r_i} I(x,y,z)}{N_i}, \quad 1 \leq i \leq N_R. \quad (1)$$

If the image is ideal and the image acquisition is perfect, i.e., noiseless and artifact-free, all r_i with the same true underlying state s_i will share the same observed feature f_i . However, images are often corrupted by noise (often approximated as white Gaussian in MR images with signal-to-noise ratio (SNR) > 2 [43]) such that representing f_i by mean intensities provides noise resiliency, and, given f_i , the true state s_i can be optimally estimated in the MAP sense by using the proposed 3-D

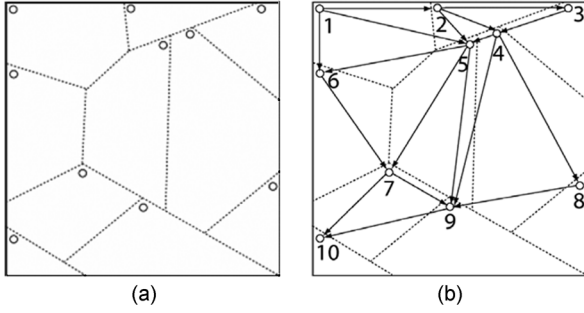


Fig. 1. A 2-D example of rbHMM. (a) Predelineated region map; each region is represented as a node in a HMM. (b) Representation of connected nodes based on the lexicographic order (indicated by the node number) of regions from top-left to bottom-right.

rbHMM. We inherit the following model assumptions from the classical HMM.

Assumption 1: The observed features follow a conditional Gaussian distribution given a hidden model state $l \in \{1, 2, \dots, N_s\}$ for N_s true underlying states such that the distribution parameters (mean μ_l , and variance σ_l^2) can be readily estimated from the sample observation. The choice of N_s is generally determined *a priori*. For example, in MRI data, N_s is typically the number of distinct tissues present. If N_s is not apparent from the data, model selection methods such as information criterion approaches [44], [45] can be applied.

Assumption 2: For any region r_i , the observed features f_i given the hidden states s_i are conditionally independent.

Assumption 3: s_i is governed by an irregular Markov mesh with transitional probabilities: $P(s_i = l | n_i) = \alpha_{n_i, l}$, where $n_i = \{(s''_j, f''_j) : r''_j \in R''_i\}$ denotes a set of states and features of the preceding neighbors, R''_i , of r_i in the irregular mesh. R''_i is a subset of R'_i , where states s''_j are known, and $\|R''_i\| = N''_i$ is the total number of preceding neighbors.

Fig. 1 shows a 2-D representation of rbHMM based on the lexicographic order of regions from top-left to bottom-right. This model shows that each node transition is affected only by the immediate preceding neighbors, in this case, $R''_i = \{r_j : j < i\}$. However, this representation is not unique for any given region map as one can traverse the image regions in many different ways, e.g., from bottom-left to top-right, depending on the node number assignment. For real images, the ordering of regions and the transitional dependencies are based on the image features and cannot be predetermined. Section II-C explains the computation of this ordering in detail.

C. Model Parameter Estimation

Unlike a grid-based 3-D HMM [27], [38], where a total of $N_s(W \times H \times D)$ possible state combinations exist, the proposed 3-D rbHMM greatly reduces the state search space to $N_s^{N_R}$. We adopt the Viterbi algorithm used in 1-D HMM estimation to iteratively search for the optimal MAP states $\{s_i^* : 1 \leq i \leq N_R\}$, given the current model parameters by locally optimizing the likelihood of the observation. The unknown states are estimated, and the model parameters are then updated assuming that the MAP states s_i^* are the true underlying states.

We initialize s_i^* using K -means clustering [46] on f_i for simplicity. Other alternative unsupervised learning algorithms such as mixture modeling [47] can also be employed. To compute the model parameters $[\mu_l, \sigma_l^2, \alpha_{n_i, l}]$ for each state $l \in \{1, 2, \dots, N_s\}$, the mean μ_l and variance σ_l^2 are estimated as the sample mean and the sample variance of all observations, and the transition probabilities $\alpha_{n_i, l}$ are computed based on the empirical frequencies and subject to the constraint $\sum_{l=1}^{N_s} \alpha_{n_i, l} = 1$ for $1 \leq i \leq N_R$. These model parameters are then iteratively updated for unsupervised estimation without priors or training. By assuming that the estimated states are the true underlying states, the parameter estimation procedure is identical to an expectation-maximization step by finding the maximum-likelihood estimates of parameters in a discrete Gaussian mixture HMM [22]. The key computational issue in applying the Viterbi algorithm to the 3-D rbHMM is to solve for the MAP states s_i^* given a set of parameters with no dependencies or directionality assumption.

Thus, we propose the following iterative algorithm to search for the MAP states, where, to reiterate, R'_i denotes the set of all neighbors of region r_i , R''_i denotes the set of preceding neighbors of r_i , and s_i and f_i denote the state and observed feature of r_i .

- Step 1) Initialize iteration number $t = 1$.
- Step 2) Initialize $s_i^{(t-1)}$ using K -means for $1 \leq i \leq N_R$.
- Step 3) Initialize $\mu_l, \mu_l^2, n_i, \alpha_{n_i, l}$, for $1 \leq i \leq N_R, 1 \leq l \leq N_s$.
- Step 4) Initialize tree branch depth $d = 0$.
- Step 5) Initialize current and processed region sets:

$$C^d = C^0 = \{r_i\} \text{ where } i = \text{rand}(\{1, 2, \dots, N_R\}).$$

$$P^d = P^0 = \{\emptyset\}.$$

- Step 6) Update the processed region set by including the current region set:

$$P^{(d+1)} = \{P^d \cup C^d\}.$$

Expand the current region set with neighbors of the current region set that have not yet been processed:

$$C^{(d+1)} = \{\cup R'_i\} \setminus P^{(d+1)}, \quad r_i \in C^d.$$

- Step 7) If $(\|C^{(d+1)}\| > 0)d++$ and go to Step 6), else $t++$.
- Step 8) For each tree branch given r_i from $d = 0$ to end

$$S_i = \{s_j^t : r_j \in R''_i\} \cup \{s_j^{(t-1)} : r_j \in R'_i \setminus R''_i\}$$

$$F_i = \{f_j : r_j \in R'_i\}.$$

Search for MAP states s_i^* given S_i and F_i .

End

- Step 9) Update $s_i^t = s_i^*$ for $1 \leq i \leq N_R$.
- Step 10) Update $\mu_l, \mu_l^2, n_i, \alpha_{n_i, l}$ for $1 \leq i \leq N_R, 1 \leq l \leq N_s$.
- Step 11) If $(t < \text{max iterations})$ go to Step 4).
- Step 12) Return s_i^t for $1 \leq i \leq N_R$.

As can be seen above, the algorithm starts at a random region in each iteration t . A nested process [Steps 6)–7)] iteratively constructs a Markov tree by extending one step outward from all nodes that have already been formed towards the immediate neighboring regions. The direction of the tree branching at each node follows a path of the most influential neighbor, which is determined based on the sum of two ratios defined as region adjacency $a_R(j, i)$ and edge adjacency $a_E(j, i)$ between neighboring regions r_j and r_i :

$$a_R(j, i) = \frac{\text{area}(r_j)}{\sum_{r_k \in R'_i} \text{area}(r_k)} \quad (2)$$

$$a_E(j, i) = \frac{\text{edge}(r_j, r_i)}{\sum_{r_k \in R'_i} \text{edge}(r_k, r_i)} \begin{cases} 1 \leq i, j \leq N_R \\ r_j \in R'_i \end{cases} \quad (3)$$

where $\text{area}(r_j)$ denotes the area of region r_j and $\text{edge}(r_j, r_i)$ denotes the number of edge points between regions r_i and r_j . Both region and edge adjacencies are defined as region and edge ratios of the neighborhood of the region r_i . Under the same initial conditions, the direction of the tree branching remains the same under image rotations since the regional neighborhood relationship has not changed; thus, allowing a rotationally invariant description of the image data.

For each tree branch, 1-D Viterbi is used to search for the optimal MAP states, s_i^* , given in (4)

$$\begin{aligned} s_i^* &= \arg \max_{s_i} P(s_i | f_i, \bar{s}_j, f_j : r_j \in R'_i) \\ &= \arg \max_{s_i} P(s_i, f_i, \bar{s}_j, f_j \bar{s}_k, f_k : r_j \in R''_i, r_k \in R'_i \setminus R''_i) \\ &= \arg \max_{s_i} P(f_i | s_i) \times P(s_i | \bar{s}_j, : r_j \in R'_i) \\ &\quad \times P(\bar{s}_k | s_i, \bar{s}_j, : r_j \in R'_i, r_k \in R'_i \setminus R''_i). \end{aligned} \quad (4)$$

The first term on the right-hand side of the last equality sign for (4) is the likelihood of the observation belonging to the state, whereas the second and third terms indicate the transition probabilities. This derivation is based on [38], where \bar{s} is used to distinguish the conditioned states of the neighboring regions from the state, s , of the current region to be optimized. The difference is that instead of locally optimizing 1-D rows on a 3-D grid s_i^* is calculated along each 1-D model branch. For nodes that lie on multiple branches, a majority voting rule is applied after all the branches have been optimized.

Since we work on an irregular representation rather than a grid mesh, $\alpha_{n_i, l}$ is generated and iteratively updated as a codebook of neighboring states given state l . Each region in the image often yields a unique codebook entry with the probability determined based on the region area ratio. In addition, to improve lookup efficiency, we define an overall neighborhood adjacency overlap ratio $a_O(i, j)$ in (5) such that any pair of entries i and j with significant overlap ($>70\%$) will have their corresponding region and edge adjacencies merged and the corresponding transition probabilities updated by using a weighted

sum based on the size of regions i and j normalized over the entire image or volume

$$\begin{aligned} a_O(i, j) &= \frac{1}{2} \sum_{k=1}^{N_s} \left(\min \left\{ \sum_{m \in R'_i} a_R(m, i), \right. \right. \\ &\quad \left. \left. \times \sum_{n \in R'_j} a_R(n, j) | s_m = s_n = k \right\} \right. \\ &\quad \left. + \min \left\{ \sum_{m \in R'_i} a_E(m, i), \right. \right. \\ &\quad \left. \left. \times \sum_{n \in R'_j} a_E(n, j) | s_m = s_n = k \right\} \right). \end{aligned} \quad (5)$$

III. RESULTS AND DISCUSSION

We applied our proposed rbHMM to 3-D image segmentation, and demonstrated in the following sections: 1) the robustness of the proposed 3-D rbHMM to parameter settings and object rotation; 2) the accuracy and efficiency of image segmentation using synthetic geometric shapes and the accuracy and robustness of the proposed method in segmenting brain tissues using: 3) simulated MRI data and 4) clinical MRI scans.

We first tested on synthetic geometric shapes with ground truth available similar to [38]. In these images, the objects were black (intensity value 0) on a white (intensity value 1) background. The size of each image was $100 \times 100 \times 100$ voxels. Gaussian noise was added with varying standard deviations σ . The shapes were generated using

$$\text{Torus : } \left(c - \sqrt{(x^2 + y^2)} \right)^2 + z^2 = a^2 \quad (6)$$

$$\text{Ellipsoid : } (x^2/a^2) + (y^2/b^2) + (z^2/c^2) = 1 \quad (7)$$

$$\text{Hyperboloid : } (x^2/a^2) + (y^2/b^2) - (z^2/c^2) = 1 \quad (8)$$

where (x, y, z) are the coordinates along the X -, Y -, and Z -axes, and a , b , and c are parameters governing the shape of each structure.

We also validated our method by applying the 3-D rbHMM to six simulated T1-weighted BrainWeb [48] MRI image volumes (with 0/1/3/5/7/9% noise, $181 \times 217 \times 181$ resolution, $1 \times 1 \times 1 \text{ mm}^3$ spacing), and 18 real high-resolution clinical T1-weighted MRI scans from the internet brain segmentation repository (IBSR) [49] (coronally acquired, $256 \times 128 \times 256$ resolution, $0.837 \times 0.837 \sim 1 \times 1 \text{ mm}^2$ in-plane spacing, 1.5-mm slice thickness). For both datasets, the ‘‘ground truth’’ is known for comparisons. For the BrainWeb dataset, the ground truth is the phantom atlas used to generate the simulated scans, whereas for the IBSR dataset the ground truth is the provided expert-guided manual segmentation label for each of the clinical scans.

TABLE I
MISCLASSIFICATION ERROR RATE OF SEGMENTATION RESULTS OBTAINED BY USING THE PROPOSED 3-D rbHMM METHOD ON 3-D TORII
WITH SHOWN PARAMETERS ($\sigma = 0.5$ AND 0.6)

<i>a</i>	<i>b</i>	<i>C</i>	$T_w=[0.65, 0.70, 0.75, 0.80, 0.85]$			
			$\sigma=0.5$		$\sigma=0.6$	
			Mean Misclass. Rate	Std. Dev.	Mean Misclass. Rate	Std. Dev.
20	30	40	0.60%	0.08%	0.66%	0.05%
40	30	40	0.78%	0.06%	0.93%	0.07%
40	20	40	0.77%	0.03%	0.86%	0.06%
40	10	40	0.74%	0.01%	0.90%	0.07%
30	30	40	0.81%	0.01%	0.92%	0.06%
30	30	30	0.84%	0.01%	0.93%	0.06%

a denotes the radius of the torus and *c* denotes the radius of the tube. We note the stable performance with regard to varying parameter T_w .

TABLE II
MISCLASSIFICATION ERROR RATE OF SEGMENTATION RESULTS OBTAINED BY USING THE PROPOSED 3D rbHMM METHOD ON 3-D ELLIPSOIDS
WITH SHOWN PARAMETERS ($\sigma = 0.5$ AND 0.6)

<i>a</i>	<i>b</i>	<i>C</i>	$T_w=[0.65, 0.70, 0.75, 0.80, 0.85]$			
			$\sigma=0.5$		$\sigma=0.6$	
			Mean Misclass. Rate	Std. Dev.	Mean Misclass. Rate	Std. Dev.
20	30	40	0.49%	0.04%	0.64%	0.08%
40	30	40	0.56%	0.00%	0.68%	0.01%
40	20	40	0.53%	0.01%	0.64%	0.01%
40	10	40	1.01%	0.25%	3.10%	1.02%
30	30	40	0.50%	0.01%	0.59%	0.01%
30	30	30	0.47%	0.02%	0.55%	0.04%

a, *b*, and *c* denote radii along each axis. We note the stable performance with regard to varying parameter T_w .

TABLE III
MISCLASSIFICATION ERROR RATE OF SEGMENTATION RESULTS OBTAINED BY USING THE PROPOSED 3D rbHMM METHOD ON 3-D HYPERBOLOIDS
WITH SHOWN PARAMETERS ($\sigma = 0.5$ AND 0.6)

<i>a</i>	<i>b</i>	<i>C</i>	$T_w=[0.65, 0.70, 0.75, 0.80, 0.85]$			
			$\sigma=0.5$		$\sigma=0.6$	
			Mean Misclass. Rate	Std. Dev.	Mean Misclass. Rate	Std. Dev.
20	30	40	0.58%	0.00%	0.74%	0.01%
40	30	40	0.57%	0.01%	0.70%	0.03%
40	20	40	0.70%	0.01%	0.83%	0.02%
40	10	40	0.63%	0.04%	0.82%	0.12%
30	30	40	0.69%	0.02%	0.80%	0.02%
30	30	30	0.65%	0.00%	0.84%	0.02%

a and *b* govern the skirt radii and *c* determines the sharpness. We note the stable performance with regard to varying parameter T_w .

The quantitative evaluation metrics in

$$\text{Misclassification rate : } \frac{F_+ + F_-}{T_+ + T_-} \times 100\% \quad (9)$$

$$\text{Dice similarity index : } \frac{2T_+}{2T_+ + F_+ + F_-} \times 100\% \quad (10)$$

were chosen to facilitate direct comparisons to other published results, where T_+ , T_- , F_+ , and F_- are the true positives, true negatives, false positives, and false negatives, respectively, between the known ground truth and the resulting segmentation results. The misclassification rate was used to assess the binary segmentation performance on the 3-D synthetic geometric shapes, whereas the Dice similarity index [50] was used for quantitative evaluation of the 3-D brain MRI scans. We compared our results with results presented by Joshi *et al.* [38] for

classification of 3-D geometric shapes and implemented their grid-based HMM for segmentation of brain MRIs.

A. Robustness to Parameter Settings and Object Rotation

1) *Parameter Sensitivity*: We first performed a two-class (object and background) segmentation on the 3-D geometric shapes. We examined the misclassification rates in the segmentation when varying the watershed threshold parameter T_w by $\pm 13.3\%$ around 0.75 (from 0.65 to 0.85). T_w denotes the percentage of significant edges present in the image, and, as its value increases, the sizes of the regions decrease. A general rule of thumb for selecting the value of T_w (and the maximum number of iterations) is to examine the granularity and performance versus efficiency tradeoffs for specific applications. By selecting $T_w = 0.75$, we achieved 90% reduction in the number of image regions used, i.e., $N_R \leq 100\,000$, as opposed to processing every image voxel ($W \times H \times D = 1\,000\,000$). This

TABLE IV
EFFECT OF ROTATION ON rbHMM SEGMENTATION RESULTS OF 3-D ELLIPSOIDS

Axis/Angle	0°	10°	20°	30°	40°
X	0.49%	0.50%	0.53%	0.53%	0.54%
Y	0.50%	0.51%	0.51%	0.52%	0.52%
Z	0.50%	0.51%	0.52%	0.53%	0.53%
Axis/Angle	50°	60°	70°	80°	90°
X	0.54%	0.54%	0.52%	0.52%	0.47%
Y	0.52%	0.55%	0.51%	0.51%	0.49%
Z	0.53%	0.52%	0.50%	0.50%	0.47%

We note the stable performance with varying angles of rotation with $\leq 0.07\%$ changes in the misclassification rate.

parameter sensitivity analysis was carried out first to ensure proper selection of the segmentation parameter for comparison to the best results presented by Joshi *et al.* [38] in Section III-B. Tables I–III show the results after three iterations of the parameter estimation algorithm with $\sigma = 0.5$ and $\sigma = 0.6$. Results indicate that our method is robust to changes in T_w . The worst case performance (ellipsoid, $a = 40, b = 10, c = 40, \sigma = 0.6$) was still significantly more stable and accurate than the results in [38], where misclassification rates of 5.76% \sim 22.30% were reported.

2) *Object Rotation*: To illustrate the robustness of rbHMM to object rotation, we demonstrated rotational invariance on a 3-D ellipsoid ($a = 40, b = 25, c = 30$) with noise ($\sigma = 0.5$). The image was progressively rotated between 0° to 90° along each axis. Our proposed parameter estimation algorithm was then applied for three iterations at a fixed $T_w = 0.75$. Table IV demonstrates our method's robustness to rotation. The misclassification error ranged from 0.47% to 0.55%, as compared with a range of 0.40% to 6.70% with the grid-based approach [38]. With only three iterations, we noted a very small difference (0.01%) at 0° rotation between each run of the proposed method due to the random initialization nature of the parameter estimation algorithm and such difference will diminish with increasing iterations. The proposed method is highly stable and repeatable. Additional repeatability studies on the same 3-D ellipsoid showed an error standard deviation due to random initializations of 0.0026% for twenty trials of the same dataset. Also, for 20 different random noise patterns applied at $\sigma = 0.5$, the repeatability studies showed an error standard deviation of 0.0089%. Qualitatively, compared with the segmentation results shown in Joshi *et al.* [38], Fig. 2 shows cleaner segmentation results based on the proposed method without significant observable artifacts such as horizontal streaks along the search direction.

B. Segmentation on Synthetic Geometric Shapes

1) *Accuracy*: Next, we compare the highest accuracies of the proposed 3-D rbHMM model from Section III-A1 above to the highest accuracy results reported in Joshi *et al.* [38] for the same shape parameter values tested at $\sigma = 0.5$ and $\sigma = 0.6$. Tables V–VII show that rbHMM outperformed grid-based HMM under various noise conditions in the vast majority of synthetic geometric shapes tested (32 out of 36). One result was equal to and three were only slightly lower (≤ 0.38 percentage point) than those of grid-based HMM. Overall, we achieved 66.24% ($\sigma = 0.5$) and 54.05% ($\sigma = 0.6$) reductions in the misclassification rates compared to those reported in [38]. Figs. 3–5 further illustrate the segmentation performance qualitatively.

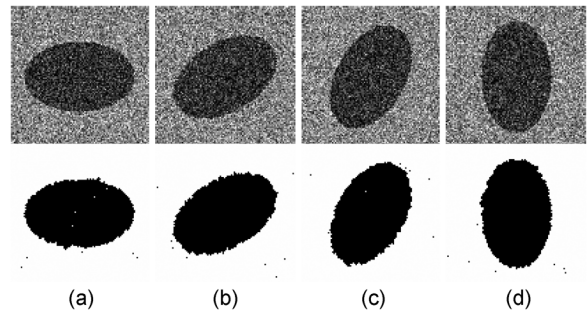


Fig. 2. Effect of varying rotational angles ($0^\circ \sim 90^\circ$) on the proposed rbHMM segmentation performance of 3-D ellipsoids. This figure shows greater invariance of the proposed method than the grid-based HMM [38], where striping artifacts are observable and classification performance is affected by the angles of rotation. Top row: original noisy images for $\sigma = 0.5$. Bottom row: segmentation results. (a) 0° . (b) 30° . (c) 60° . (d) 90° .

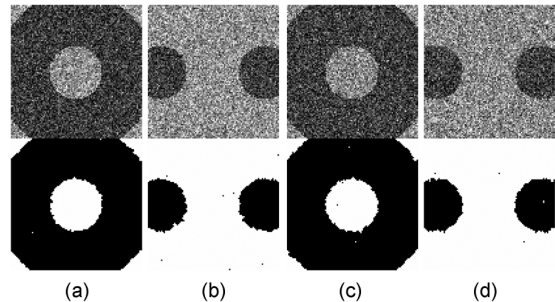


Fig. 3. rbHMM segmentation results on 3-D torus images with parameters $a = 20$ and $c = 40$. Top row: original noisy images for (a), (b) $\sigma = 0.5$ and (c), (d) $\sigma = 0.6$. Bottom row: segmentation outputs of corresponding slices.

2) *Efficiency*: Complexity-wise, applying 1-D Viterbi to 3-D grid-based HMM was shown to be $O(WHDN_s^2)$ [27], where $W, H,$ and D are the dimensions of the 3-D image. Our proposed 3-D rbHMM on the other hand can be processed in $O(N_N N_B N_s^2)$ time, where N_B is the number of tree branches and N_N is the average number of nodes per branch. If no branch overlaps exist, $N_N \times N_B = N_R$. The proposed algorithm is thus, generally, more computationally efficient, since, with proper watershed delineation, $N_N \times N_B \ll W \times H \times D$. The average running time of our 3-D rbHMM for the 18 tested 3-D geometric shapes above ($\sigma = 0.5, 100 \times 100 \times 100$ voxels) with three iterations at $T_w = 0.75$ was 170 s on a 3.0-GHz Xeon processor. This corresponds to roughly a 40% speed increase when compared with 280 s reported based on the grid-based 3-D HMM [38], adjusted for CPU clock speed.

TABLE V
 COMPARISON OF SEGMENTATION MISCLASSIFICATION ERROR RATE: GRID-BASED 3-D HMM VERSUS PROPOSED 3-D rbHMM ON 3-D TORI
 WITH SHOWN PARAMETERS

a	c	Grid HMM		Proposed 3D rbHMM	
		$\sigma=0.5$	$\sigma=0.6$	$\sigma=0.5$	$\sigma=0.6$
20	20	0.64%	1.70%	0.56%	0.56%
20	30	1.04%	1.49%	0.67%	0.81%
20	40	1.15%	1.64%	0.75%	0.76%
30	30	3.14%	4.33%	0.74%	0.77%
30	40	2.33%	2.83%	0.80%	0.82%
40	40	0.91%	0.45%	0.83%	0.83%
Mean		1.54%	2.07%	0.73%	0.76%
Std. Dev.		0.98%	1.34%	0.10%	0.10%

a denotes the radius of the torus and c denotes the radius of the tube. The best performance is highlighted in bold.

 TABLE VI
 COMPARISON OF SEGMENTATION MISCLASSIFICATION ERROR RATE: GRID-BASED 3-D HMM VERSUS PROPOSED 3-D rbHMM ON 3-D ELLIPSOIDS
 WITH SHOWN PARAMETERS

a	b	c	Grid HMM		Proposed 3D rbHMM	
			$\sigma=0.5$	$\sigma=0.6$	$\sigma=0.5$	$\sigma=0.6$
20	30	40	3.37%	4.61%	0.45%	0.56%
40	30	40	0.56%	0.89%	0.56%	0.67%
40	20	40	0.43%	0.64%	0.52%	0.63%
40	10	40	6.26%	5.76%	0.73%	1.86%
30	30	40	0.48%	0.86%	0.49%	0.58%
30	30	30	1.69%	4.41%	0.44%	0.51%
Mean			2.13%	2.86%	0.53%	0.80%
Std. Dev.			2.32%	2.31%	0.11%	0.52%

a , b and c denote radii along each axis. The best performance is highlighted in bold.

 TABLE VII
 COMPARISON OF SEGMENTATION MISCLASSIFICATION ERROR RATE: GRID-BASED 3-D HMM VERSUS PROPOSED 3-D rbHMM ON 3-D HYPERBOLOIDS
 WITH SHOWN PARAMETERS

a	b	c	Grid HMM		3D rbHMM	
			$\sigma=0.5$	$\sigma=0.6$	$\sigma=0.5$	$\sigma=0.6$
10	10	10	0.80%	1.12%	0.58%	0.74%
10	20	30	0.90%	2.24%	0.56%	0.67%
10	30	30	0.74%	3.69%	0.70%	0.81%
10	20	40	1.43%	1.99%	0.58%	0.68%
10	30	40	0.97%	1.24%	0.67%	0.79%
20	30	20	0.74%	1.19%	0.65%	0.82%
Mean			0.93%	1.91%	0.62%	0.75%
Std. Dev.			0.26%	0.99%	0.06%	0.07%

a and b govern the skirt radii and c determines the sharpness. The best performance is highlighted in bold.

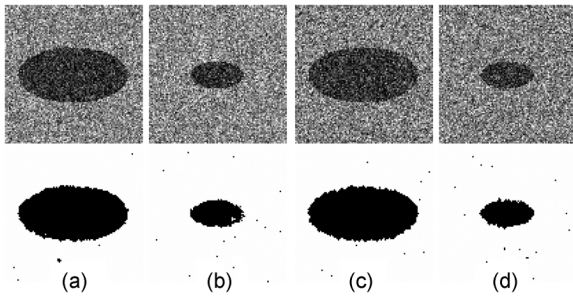


Fig. 4. rbHMM segmentation results on 3-D ellipsoid images with parameters $a = 40$, $b = 20$, and $c = 40$. Top row: original noisy images for (a), (b) $\sigma = 0.5$ and (c), (d) $\sigma = 0.6$. Bottom row: segmentation outputs of corresponding slices.

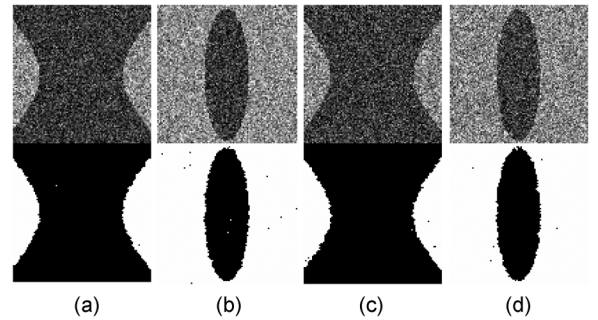


Fig. 5. rbHMM segmentation results on 3-D hyperboloid images with parameters $a = 10$, $b = 30$, and $c = 30$. Top row: original noisy images for (a), (b) $\sigma = 0.5$ and (c), (d) $\sigma = 0.6$. Bottom row: segmentation outputs of corresponding slices.

C. Segmentation of Simulated MRI Data

1) *Accuracy*: We performed four-class segmentation: background, white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF), on the simulated T1-weighted BrainWeb MR images [48] with varying noise levels ($n = 0\%$, 1% , 3% , 5% , 7% , and 9%). The proposed rbHMM was applied for ten

iterations at $T_w = 0.25$, which corresponded to approximately $96.7\% \sim 98.7\%$ reduction in the number of image regions used ($N_R = 89725 \sim 234501$) as opposed to processing every image voxel ($W \times H \times D = 7109137$). The Dice similarity index [50] was used to quantitatively evaluate WM and GM

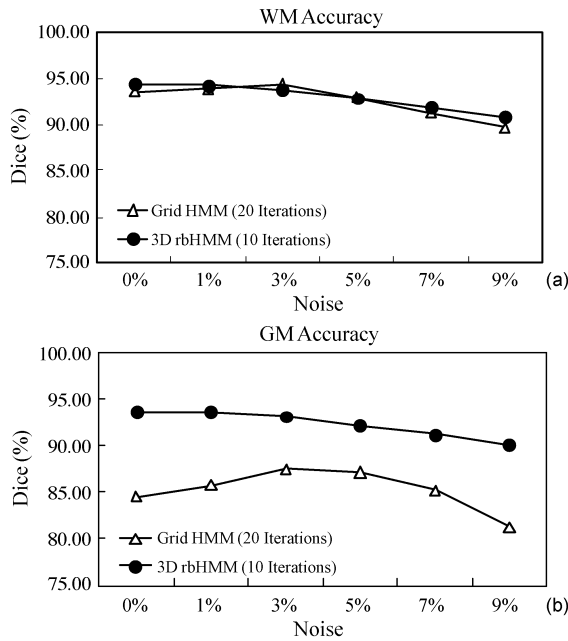


Fig. 6. Comparisons of segmentation accuracy of the proposed 3-D rbHMM and the grid-based HMM on simulated 3-D brain MRI images for (a) WM and (b) GM. Our method shows consistently high accuracies and stable results even when the noise level is varied from 0% to 9%.

segmentation accuracies by comparing with the provided phantom.

We compared our segmentation results against the grid-based HMM. The grid-based HMM implementation employed three states per class (K -means clustering was used to initialize each class, to which three states were then randomly assigned within each class) as described by Joshi *et al.* [38], full 3-D probability transitions, and 20 iterations due to its slower convergence as demonstrated by the results from the experiments of 3-D geometric shapes. Our proposed rbHMM achieved, on average over various noise levels, WM and GM similarity indices of 92.99% ($\sigma = 1.42\%$) and 92.27% ($\sigma = 1.44\%$), respectively. The grid-based HMM achieved 92.65% ($\sigma = 1.81\%$) and 85.22% ($\sigma = 2.23\%$) in WM and GM similarity indices. Fig. 6 illustrated that both our WM and GM segmentation results were consistently stable and highly accurate in the presence of noise. For the grid-based HMM approach, we observed consistent and comparable results for WM segmentation, but performance largely degrades for GM. This degradation in GM segmentation performance was also observable in the segmentation outputs illustrated in Fig. 7. Two main artifacts were observed in the grid-based HMM segmentation results. First, the structural details remained grainy in regions of high noise. A second horizontal striping artifact was also observed with the higher noise scans, resulting in increased jaggedness around class boundaries. This striping artifact was similarly observed with the segmentation results of the high noise 3-D synthetic geometric images [38]. We further examine this striping artifact in our subsequent experiment with object rotations. The average running time of the proposed method was 94 s per iteration on a 3.0-GHz Xeon processor.

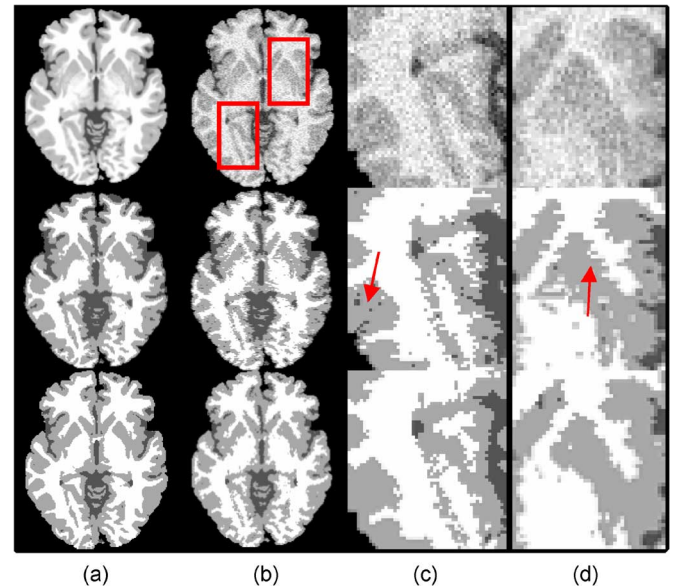


Fig. 7. 3-D simulated BrainWeb T1-weighted brain MR images showing effect of varying noise levels on segmentation performance. Top row: original noisy images. Middle row: grid-based HMM segmentation outputs. Bottom row: rbHMM segmentation outputs for noise levels of (a) 0% and (b) 9%. We demonstrated superior GM segmentation using the proposed rbHMM, whereas the grid-based HMM approach showed (c) grainy and (d) striping in the close-up results of the highlighted regions. Problematic regions are identified by arrows.

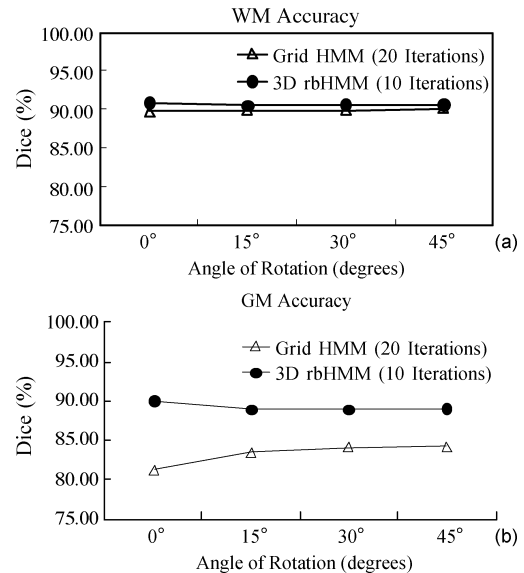


Fig. 8. Comparisons showing effect of varying rotational angles on segmentation accuracy of the proposed 3-D rbHMM and the grid-based HMM on simulated 3-D brain MRI images for (a) WM and (b) GM. Both methods show consistently stable results when the angle is varied from 0° to 45°.

2) *Object Rotation*: We next compared the rotational invariance capabilities of our proposed rbHMM to the grid-based HMM approach on a 3-D BrainWeb volume (9% noise) rotated for 15°, 30°, and 45° around the Z -axis. Rotated phantoms were generated using nearest neighbor interpolation for the quantitative comparison in Fig. 8 below. We observed that both methods appeared equally stable with respect to various rotational angles, but our results are consistently more accurate than the grid-based HMM, especially for gray matter segmentation. The small variations ($\sim 1\%$) in WM and GM similarity indices

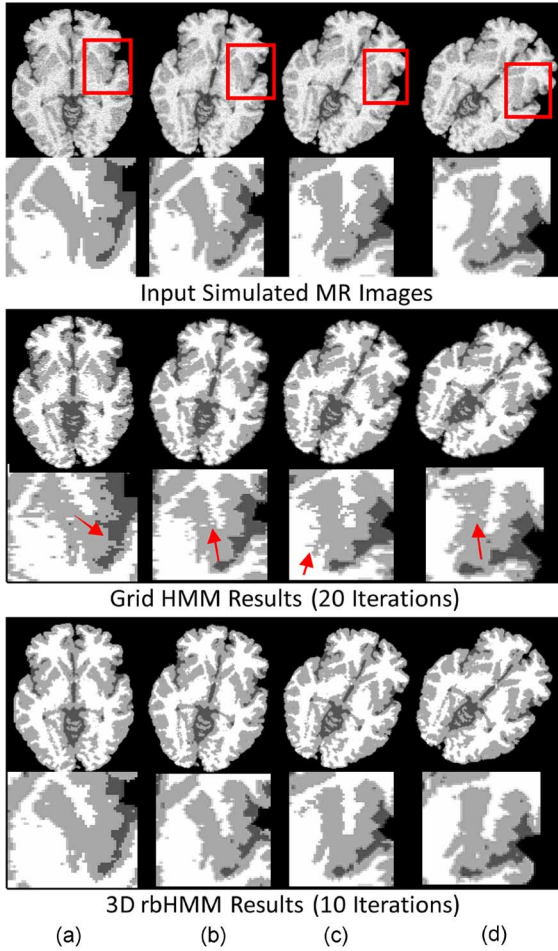


Fig. 9. 3-D simulated BrainWeb T1-weighted brain MR images showing effect of (a) 0°, (b) 15°, (c) 30°, and (d) 45° rotational angles on segmentation performance. Rows 1–2: original noisy simulated MR images (9%) and the corresponding phantom labels. Rows 3–4: grid-based HMM segmentation outputs and close-up results of the highlighted regions showing striping artifact in the horizontal decoding direction. Rows 5–6: rbHMM segmentation outputs and close-up results of the highlighted regions. Problematic regions are identified by arrows.

apparently due to rotation were largely caused by variations introduced with resampling of the discrete phantom ground truth.

Qualitatively, however, we were able to demonstrate that both methods, in fact, did not perform equally under object rotation. Fig. 9 illustrates qualitative segmentation results for both approaches. We confirmed no observable artifacts resulted from object rotations for the proposed rbHMM method. However, for the grid-based HMM method, as the object was rotated, the striping artifact remained in the direction of the horizontal grid-based decoding process, irrespective of the rotational angles. This was especially observable in the regions of caudate nucleus and putamen. A side effect of this artifact could also be observed in the quantitative results, where, at 45° rotation, the GM accuracy is increased by 3.04% in similarity index than that of 0° rotation. Since, at 45°, the subcortical GM structures appeared approximately horizontal/vertical in the simulated MRI scan, a reduced amount of errors were introduced, resulting in higher GM segmentation accuracy. Thus, although the quantitative results for the grid-based HMM method were stable, qualitative results showed that the proposed rbHMM method

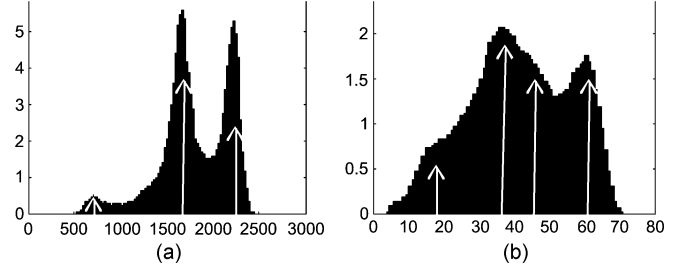


Fig. 10. Intensity histograms of MR images (excluding background) showing the differences between (a) simulated BrainWeb (3% noise) and (b) clinical IBSR V2.0 (#12) scans. The BrainWeb histogram demonstrated three distinct peaks, whereas the IBSR V2.0 histogram showed four mixtures.

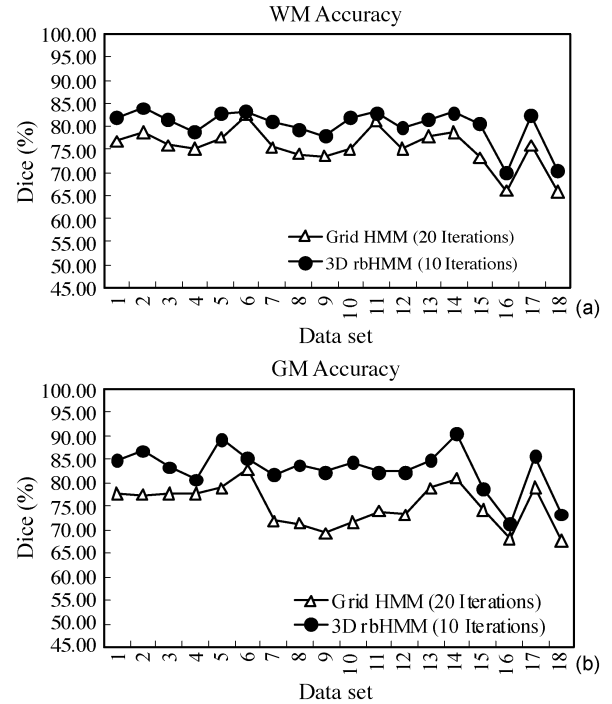


Fig. 11. Comparisons of segmentation accuracy of proposed 3-D rbHMM and grid-based HMM on clinical 3-D IBSR V2.0 brain MRI images for (a) WM and (b) GM. Our method shows consistently high accuracies and stable results.

demonstrated improved robustness to object rotations than the grid-based HMM approach.

D. Segmentation of Clinical MRI Data

We performed segmentation on 18 clinical T1-weighted IBSR V2.0 MR images [49]. 3-D rbHMM was applied for ten iterations at $T_w = 0.25$, which corresponded to approximately 98.8% ~ 99.2% reduction in the number of image regions used ($N_R = 64204 \sim 99008$) as opposed to processing every image voxel ($W \times H \times D = 8388608$). Evaluation of segmentation accuracies was performed by comparing against the provided expert-guided manual segmentation labels. A five-class: background, GM (two classes), WM, and CSF segmentation was performed rather than a four-class data representation because the intensity histogram of these clinical scans exhibited different mixture distribution properties than the simulated scans as shown in Fig. 10; thus, a different mixture model was necessary for the subsequent segmentation task.

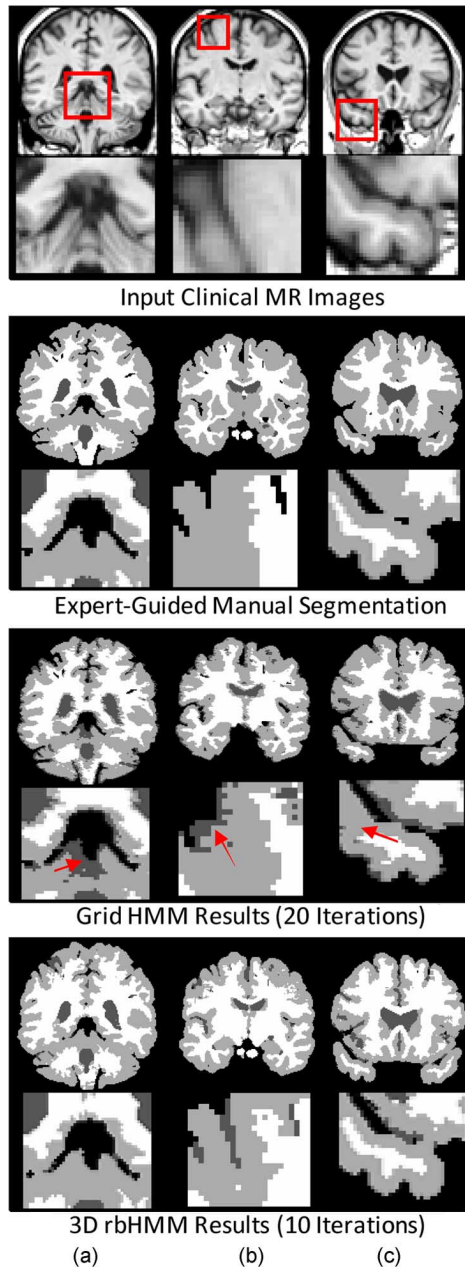


Fig. 12. 3-D clinical IBSR V2.0 T1-weighted brain MR images (normal adult scan #06) showing segmentation accuracy of the proposed 3-D rbHMM method. Rows 1–2: original image clinical MR images and close-up images of the highlighted regions. Rows 3–4: expert-guided manual segmentation labels. Rows 5–6: grid-based HMM segmentation outputs. Rows 7–8: 3-D rbHMM segmentation outputs of slice number (a) 45, (b) 65 and (c) 85. Problematic regions are identified by arrows.

Fig. 11 shows that the proposed rbHMM achieved, on average over scans, similarity indices of 80.08% ($\sigma = 3.95\%$) and 82.78% ($\sigma = 4.81\%$) with the expert-guided manual segmentation labels for WM and GM, respectively, whereas the grid-based HMM method achieved 75.48% ($\sigma = 4.21\%$) and 75.07% ($\sigma = 4.48\%$) for WM and GM overlaps. To statistically evaluate the differences of segmentation results between rbHMM and the grid-based HMM, we calculated the p values of WM and GM similarity indices ($p < 0.05$ indicates

significant difference). The proposed rbHMM method achieved significantly higher accuracies with increased similarity indices of 4.60% ($p = 0.0022$) in WM and 7.71% ($p < 0.0001$) in GM segmentation results. Further qualitative examples demonstrating the superior performance of the proposed segmentation method are shown in Fig. 12. We observed that our rbHMM segmentation results matched more closely to the manual segmentation labels than the results of using the grid-based HMM approach. However, both methods undersegmented subcortical GM structures (thalamus and putamen) when compared with the manual segmentation labels, as shown in Fig. 12(b). This was mainly due to the fact that the intensity features of these structures in this MR image were very similar to their surrounding WM. Without any other *a priori* cues such as expert knowledge, these subcortical structures were fairly indistinguishable. For the grid-based approach, the previously observed striping artifact was not as significant as before because the high quality of these clinical scans where noise was not a major factor. However, we still observed jaggedness between the WM and GM boundaries. The results generated by using the rbHMM approach, on the other hand, were much more robust, as identified by the arrows shown in Fig. 12. The average running time of the proposed method was 62 s per iteration on a 3.0-GHz Xeon processor.

IV. CONCLUSION

We proposed a novel 3-D rbHMM that employs a new unsupervised, computationally efficient parameter estimation algorithm and demonstrated its use for accurate and robust 3-D image segmentation. The main contributions of our work are; 1) a 3-D HMM framework based on *irregularly shaped* homogeneous regions, which reduces the overall model complexity thus resulting in a more efficient optimization process compared to current approaches which are based on rectangular lattices or grids and 2) a novel 3-D tree-structured algorithm that provides *rotationally invariant* estimates of the locally optimal MAP states of image regions by employing the Viterbi algorithm. We validated our technique first using synthetic geometric shapes where we demonstrated rbHMM's improved computational efficiency and increased invariance to 3-D object rotation compared with grid-based 3-D HMM frameworks using a similar optimization strategy. Furthermore, we successfully applied our proposed segmentation method to both simulated and real clinical brain MRI data and demonstrated improved accuracy and robustness of the rbHMM over the grid-based HMM approach. Extensive tests with varying model, shape, noise, and rotation parameters demonstrated the capability of the proposed 3-D rbHMM and parameter estimation algorithm to robustly handle 3-D image segmentation.

Plans for future work include incorporating texture cues to extend the model to segmentation of color images or natural scenes, incorporating multiple medical imaging modalities for the identification of anomalies such as tumors or lesions, and integrating iterative correction of medical imaging artifacts such as MR intensity inhomogeneity, a slow-varying nonuniformities in the magnetic field, which is a common difficulty for the segmentation task.

REFERENCES

- [1] J. K. Baker, "The dragon system—An overview," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, Feb. 1975, pp. 24–29.
- [2] L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [3] M. Diligenti, P. Frasconi, and M. Gori, "Hidden tree Markov models for document image classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 4, pp. 519–523, Apr. 2003.
- [4] J. Li, A. Najmi, and R. M. Gray, "Image classification by a two-dimensional hidden Markov model," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 517–533, Feb. 2000.
- [5] E. Levin and R. Pieraccini, "Dynamic planar warping for optical character recognition," in *Proc. Int. Conf. Acoust., Speech Signal Process.*, San Francisco, CA, Mar. 1992, vol. 3, pp. 149–152.
- [6] H. K. Lee and J. H. Kim, "An HMM-based threshold model approach for gesture recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 10, pp. 961–973, Oct. 1999.
- [7] S. Eickeler, S. Muller, and G. Rigoll, "Improved face recognition using pseudo 2-D hidden Markov models," in *Proc. Workshop Adv. Facial Image Anal. Recogn. Technol. (AFIART)*, Freiburg, Germany, Jun. 1998.
- [8] F. Hulsken, F. Wallhoff, and G. Rigoll, "Facial expression recognition with pseudo-3D hidden Markov models," in *Proc. DAGM-Symp.*, 2001, vol. 2191, Lecture Notes in Computer Science, pp. 291–297.
- [9] J. Boreczky and L. Wilcox, "A hidden Markov model framework for video segmentation using audio and image features," in *Proc. IEEE Conf. Acoust., Speech, Signal Process.*, 1998, vol. 6, pp. 3741–3744.
- [10] T. Vlachos and A. G. Constantinides, "A graph-theoretic approach to colour image segmentation and contour classification," in *Proc. Int. Conf. Image Process. Its Applications*, Maastricht, The Netherlands, Apr. 1992, pp. 298–302.
- [11] K. Barnard and D. A. Forsyth, "Learning the semantics of words and pictures," in *Proc. Int. Conf. Comput. Vis.*, 2001, pp. 408–415.
- [12] J. Jiten and B. Mérialdo, "Semantic image segmentation with a multidimensional hidden Markov model," *Lecture Notes in Comput. Sci.*, vol. 4351, pp. 616–624, 2006.
- [13] A. J. Frost, M. A. Thyer, R. Srikanthan, and G. Kuczera, "A general Bayesian framework for calibrating and evaluating stochastic models of annual multi-site hydrological data," *J. Hydrol.*, vol. 340, no. 3–4, pp. 129–148, Jul. 2007.
- [14] C.-C. Ke, J. G. Herrero, and J. Llinas, "Comparative analysis of alternative ground target tracking techniques," in *Proc. IEEE Inf. Fusion*, Jul. 2000, vol. 2, WEB5/3-WEB5/10.
- [15] R. Durbin, S. Eddy, A. Krogh, and G. Mitchison, *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge, U.K.: Cambridge Univ., 1998.
- [16] O. Cappe, E. Moutines, and T. Ryden, *Inference in Hidden Markov Models*. New York: Springer-Verlag, 2005.
- [17] O. Cappe, "Ten years of HMMs," Mar. 2001 [Online]. Available: <http://www.tsi.enst.fr/~cappe/docs/hmmbib.html>
- [18] L. E. Baum, "An inequality and associated maximization technique in statistical estimation for probabilistic functions of finite state Markov chains," *Inequalities*, vol. 3, pp. 1–8, 1972.
- [19] A. A. Markov, "An example of statistical investigation in the text of 'Eugene Onyegin' illustrating coupling of 'tests' in chains," *Proc. Acad. Sci.*, vol. 7, pp. 153–162, 1913.
- [20] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, Jul. 1948.
- [21] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
- [22] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *Ann. Math. Stat.*, vol. 41, no. 1, pp. 164–171, 1970.
- [23] A. J. Viterbi and J. K. Omura, "Trellis encoding of memoryless discrete-time sources with a fidelity criterion," *IEEE Trans. Inf. Theory*, vol. IT-20, pp. 325–332, May 1974.
- [24] G. D. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.
- [25] S. S. Kuo and O. E. Agazzi, "Machine vision for keyword spotting using pseudo 2-D hidden Markov models," *Proc. Int. Conf. Acoust., Speech Signal Process.*, vol. 5, pp. 81–84, 1993.
- [26] P. A. Devijver, "Probabilistic labeling in a hidden second order Markov mesh, pattern recognition," *Practice II*, pp. 113–123, 1985.
- [27] J. Li, D. Joshi, and J. Z. Wang, "Stochastic modeling of volume images with a 3-D hidden Markov model," in *Proc. IEEE Int. Conf. Image Process.*, 2004, pp. 2359–2362.
- [28] M. Ibrahim, N. John, M. Kabuka, and A. Younis, "Hidden Markov models-based 3-D MRI brain segmentation," *Image Vis. Comput.*, vol. 24, pp. 1065–1079, 2006.
- [29] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 11, pp. 721–741, Nov. 1984.
- [30] S. Z. Li, "Markov random field models in computer vision," in *Proc. Eur. Conf. Comput. Vis.*, Stockholm, Sweden, May 1994, pp. 361–370.
- [31] I. Patras, E. A. Hendriks, and R. L. Lagendijk, "Video segmentation by MAP labeling of watershed segments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 3, pp. 326–332, Mar. 2001.
- [32] Y. Tsaig and A. Averbuch, "A region-based MRF model for unsupervised segmentation of moving objects in image sequences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 2001, vol. 1, pp. 1899–1896.
- [33] A. Sarkar, M. K. Biswas, and K. M. S. Sharma, "A simple unsupervised MRF model based image segmentation approach," *IEEE Trans. Image Process.*, vol. 9, no. 5, pp. 801–812, May 2000.
- [34] H. Deng and D. A. Clausi, "Gaussian MRF rotation-invariant features for image classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 7, pp. 951–955, Jul. 2004.
- [35] 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. Imag.*, vol. 20, no. 1, pp. 45–47, Jan. 2001.
- [36] M. Maddah, K. H. Zou, W. M. Wells, R. Kikinis, and S. K. Warfield, "Automatic optimization of segmentation algorithms through simultaneous truth and performance level estimation (STAPLE)," in *Proc. Med. Image Computing Comput. Assisted Intervention*, 2004, pp. 274–282.
- [37] J. Besag, "On the statistical analysis of dirty pictures (with discussion)," *J. Royal Stat. Soc.*, ser. B, vol. 48, no. 3, pp. 259–302, 1986.
- [38] D. Joshi, J. Li, and J. Z. Wang, "A computationally efficient approach to the estimation of two- and three-dimensional hidden Markov models," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 1871–1886, Jul. 2006.
- [39] A. Huang, R. Abugharbieh, and R. Tam, "Image segmentation using an efficient rotationally invariant 3-D region-based hidden Markov model," in *Proc. IEEE Math. Methods Biomed. Image Anal.*, 2008, pp. 1–8.
- [40] S. Buecher, "Watersheds of functions and picture segmentation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Paris, France, May 1982, pp. 1928–1931.
- [41] J. Shi and J. Malik, "Normalized cuts and image segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 1997, pp. 731–737.
- [42] F. Li, Q. Dai, and W. Xu, "Region-based hidden Markov models for image categorization and retrieval," in *Proc. SPIE Electron. Imag.*, 2007, vol. 6508, p. 65081V.
- [43] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magnetic Resonance in Medicine*, vol. 34, no. 6, pp. 910–914, Dec. 1995.
- [44] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, 1974.
- [45] G. E. Schwarz, "Estimating the dimension of a model," *Ann. Stat.*, vol. 6, no. 2, pp. 461–464, 1978.
- [46] S. P. Lloyd, "Least square quantization in PCM," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 129–136, 1982.
- [47] G. J. McLachlan and K. E. Basford, *Mixture Models: Inference and Applications to Clustering*. New York: Marcel Dekker, 1988.
- [48] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, and A. C. Evans, "BrainWeb: Online interface to a 3-D MRI simulated brain database," *NeuroImage*, vol. 5, no. 4 Part 2/4, p. S425, 1997.
- [49] A. Worth, Nov. 2007, The Center for Morphometric Analysis at MGH [Online]. Available: <http://www.cma.mgh.harvard.edu/ibsr/>
- [50] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, pp. 297–302, 1945.



Albert Huang (S'05) received the B.A.Sc. degree in electrical and computer engineering in 2000 from the University of British Columbia (UBC), Vancouver, BC, Canada, in 2002, the S.M. degree in engineering science from Harvard University, Cambridge, MA, in 2002, and is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, UBC.

From 2000 to 2004, he was an Algorithm Developer with the Harvard Communications/Molecular and Cellular Biology Laboratories and Sony Taiwan

Limited. His current research interests include signal and image processing.



Rafeef Abugharbieh (M'03–SM'10) received the M.Sc. (with distinction), Technical Licentiate, and Ph.D. degrees from the School of Electrical and Computer Engineering, Chalmers University, Göteborg, Sweden.

She was a member of the Department of Signals and Systems, Chalmers University. She is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of British Columbia (UBC), Vancouver, BC, Canada. She is the cofounder and codirector of the Biomedical

Signal and Image Computing Lab (BiSICL), a multidisciplinary research laboratory focusing on visual computing in biomedical imaging.

Dr. Abugharbieh is a Registered Member of the Association of Professional Engineers and Geoscientists of British Columbia (P.Eng., APEGBC), a UBC Killam Faculty Research Fellow, a member of the IEEE Engineering in Medicine and Biology Society (EMBS), an associate founding member of the IEEE EMBS Vancouver Section, and a member of the Medical Image Computing and Computer Assisted Intervention (MICCAI) Society.



Roger Tam received the Ph.D. degree in computer science from the University of British Columbia (UBC), Vancouver, BC, Canada, in 2004.

He is currently an Assistant Professor with the Department of Radiology, UBC, and is also a member of the UBC MS/MRI Research Group, a pioneering research laboratory in the field of MRI analysis for the study of multiple sclerosis. His current research interests include medical image processing, computational shape modeling, and scientific visualization.