

A Fuzzy Region-Based Hidden Markov Model for Partial-Volume Classification in Brain MRI

Albert Huang¹, Rafeef Abugharbieh¹, and Roger Tam²

¹ Department of Electrical and Computer Engineering

{alberth, rafeef}@ece.ubc.ca

² Department of Radiology, The University of British Columbia, Vancouver, B.C., Canada

roger@msmri.medicine.ubc.ca

Abstract. We present a novel fuzzy region-based hidden Markov model (*frbHMM*) for unsupervised partial-volume classification in brain magnetic resonance images (MRIs). The primary contribution is an efficient graphical representation of 3D image data in which irregularly-shaped image regions have memberships to a number of classes rather than one discrete class. Our model groups voxels into regions for efficient processing, but also refines the region boundaries to the voxel level for optimal accuracy. This strategy is most effective in data where partial-volume effects due to resolution-limited image acquisition result in intensity ambiguities. Our *frbHMM* employs a forward-backward scheme for parameter estimation through iterative computation of region class likelihoods. We validate our proposed method on simulated and clinical brain MRIs of both normal and multiple sclerosis subjects. Quantitative results demonstrate the advantages of our fuzzy model over the discrete approach with significant improvements in classification accuracy (30% reduction in mean square error).

1 Introduction

Graphical models have long been successfully used in various signal processing and analysis applications such as speech recognition, computer vision, error correction coding, and genome analysis. Such models provide a graphical representation of probabilistic distributions by expressing complex computations for inference and learning as simpler graphical structures [1]. There exist two major classes of graphical models - directed and undirected. The directed graphical models, or Bayesian Networks (BN), specify a particular directionality on the links of the graphs, which are useful for conveying causal relationships between the underlying random variables. An example of BN is the hidden Markov model (HMM), which is commonly used to represent stochastic processes. On the other hand, undirected graphical models or Markov random fields (MRF) do not carry any directional implications but rather specify some constraints between the random variables. Both directed and undirected graphical models have been applied successfully in the context of brain tissue segmentation from magnetic resonance imaging (MRI) data, such that each of the image voxels represents an underlying hidden random variable of tissue class label, which cannot be observed directly but can be indirectly estimated given observations by inference.

Performing exact inference on a fully connected high dimensional graphical model is a daunting task. Regardless of whether the model is directed or undirected, the amount of computations is expensive if not intractable. For the directed representation, 2D/3D HMMs (causal MRF) have been proposed on a regular rectangular pixel lattice [2], but estimating the exact states requires exponential time thus a block-based approach [3] and an iterative approximation method [4] have been proposed to improve efficiency. However, modeling an image in such a way is odd as voxels do not typically bear causal orders. Recently, a more data-driven model [5] was proposed as a tree-structure extension of the iterative approximation method [4]. For the undirected representation, performing a maximum a posteriori (MAP) estimation on a hidden MRF (HMRF) is a computationally difficult problem [6]. In such cases, the optimal solutions were usually computed using some optimization techniques through local optimization or optimization on a relaxed problem [7-9].

For image segmentation tasks, the above mentioned estimation methods commonly provide a single discrete label to each image voxel. An alternative model subdivides the content of one voxel into numerous classes simultaneously, allowing for a more accurate modeling of a common physical imaging limitation, namely, partial volume effects. Discrete segmentations thus appear as a special case by assuming that no partial volume voxels exist, where all classes are null except for the hard estimate. One can intuitively infer the partial volumes from the class distributions of a discrete labeling process using techniques such as the classical forward-backward algorithm [10]. However, a more accurate model should simultaneously model all pixel likelihoods without assuming one single, particular true label class. Furthermore, classifications based on partial-volume models have shown to achieve improved accuracies in tissue volume measurements [11, 12]. A number of interesting works in this aspect have been done using graphical models. Bricq *et al.* [13] converted 3D scan data into a 1D chain. Such vectorization is not truly 3D as only a single fixed-ordering neighboring pixel is considered in the estimation process, and the scan order is predefined irrespective of the data. Others [12, 14] utilized fuzzy hidden MRFs to incorporate information from immediate neighbors. However, parameter estimation in such an undirected model is known to be difficult and time consuming. In contrast, estimation in a directed model is comparatively easy with the results achieved being generally similar to those of an undirected model [15]. Therefore, starting with a more natural, data-driven region-based HMM (rbHMM) that was proposed in [5], where the advantages of rotational invariance and increased efficiency were shown, we propose, in this paper, a novel partial-volume extension for brain tissue segmentation – henceforth referred to as fuzzy rbHMM or *frbHMM*. We integrate the classical forward-backward scheme in the tree-structured parameter estimation algorithm and refine region boundaries to the voxel level resulting in a more accurate classification model for partial-volume effects. We present quantitative validation results demonstrating the advantages of modeling each voxel in brain MRI scans as mixtures of multiple classes as opposed to one single label in simulated as well as real clinical MR data of both normal and Multiple Sclerosis (MS) patient subjects.

2 Methods

We first briefly describe the discrete rbHMM of [5] and present our proposed extension for modeling partial volumes within a new fuzzy framework (hereafter referred

to by *frbHMM*). We then describe how the forward-backward algorithm is employed for estimating the 3D *frbHMM* parameters and region class likelihoods, and how the classification resolution is improved by further refining image regions.

2.1 Region-Based Hidden Markov Model Overview

In [5], Huang *et al.* proposed a method where an image U is first divided into a set R of contiguous homogeneous local regions r_i , each of size $N_i = \|r_i\|$, where $1 \leq i \leq N_R$ and $N_R = \|R\|$ is the total number of regions. Each voxel with coordinates (x, y, z) belongs to a region r_i if the pixel is labeled as $L(x, y, z) = i$. The assumption was that each region would exhibit similar properties such as intensity or texture, and that by grouping them together using, e.g., a watershed transform or a normalized cut, the complexity and computational cost can be largely reduced. For grayscale images, e.g. MRI, such regional features f_i can be simply defined by the mean observed voxel intensities of r_i . The observed f_i represented noisy regions with true underlying discrete states s_i , which can then be optimally estimated in a MAP sense based on the three statistical model assumptions:

Assumption 1: The observations for the underlying model states $l \in \{1, 2, \dots, N_s\}$, where N_s is the total number of given underlying states, follow Gaussian distributions with mean μ_l and variance σ_l^2 , which are estimated from the observed samples..

Assumption 2: If s_i is known, then f_i is conditionally independent of other regions.

Assumption 3: The true underlying state s_i is governed by an irregular Markov mesh such that each region r_i has a set of spatially neighboring regions R_i' . The transitional probabilities are defined as $P(s_i = l | n_i) = \alpha_{n_i, l}$, where $n_i = \{(s_i'', f_i'') : r_i'' \in R_i'\}$ is a set of states and features of some preceding neighbors R_i'' , which is a subset of R_i' where states s_i'' are already known.

The primary advantages of this region-based modeling and its tree-structured parameter estimation scheme are increased efficiency over pixel-based methods and invariance to rotations that are commonly observed in medical images.

2.2 Proposed Fuzzy Region-Based Hidden Markov Model

Our main contribution here is the extension of *rbHMM* by introducing fuzzy states that allow each region to belong to multiple classes simultaneously. Rather than considering one single ‘true’ underlying state s_i , we now consider an underlying state vector $S_i = (t_{i,1}, t_{i,2}, \dots, t_{i,N_s})$ such that $\sum_{l=1}^{N_s} t_{i,l} = 1$, $t_{i,l} \in [0,1]$ represents the proportion of the l^{th} class in region r_i for $l = 1, \dots, N_s$. The term $t_{i,l}$ can also be seen as the probability of labeling region r_i with label l given the model and observations. For discrete or crisp segmentation, $t_{i,l} = 0$ for all l in S_i except for the single element where $s_i = l$, then $t_{i,l} = 1$.

To compute the model parameters, K-means clustering is used as a simple way to initialize the state means and variances from the samples. The state probabilities $S_i = (t_{i,1}, t_{i,2}, \dots, t_{i,N_s})$ for $1 \leq i \leq N_R$ are then calculated based on a Gaussian mixture assumption. The transition probabilities $\alpha_{n_i, l}$ are determined based on probability-weighted empirical frequencies given the constraint $\sum_{l=1}^{N_s} \alpha_{n_i, l} = 1$ for $1 \leq i \leq N_s$. Once these model

parameters are initialized, the likelihoods, $t_{i,l}$, are computed using the forward-backward algorithm, which has been shown to be efficient in 1D [10]. Thus, similar to rbHMM, we iteratively construct a 3D Markov tree from a randomly selected region outwards by traversing all regions [5].

For each region, we substitute the global region index i with (b,k) such that r_i is the k^{th} region on branch b ; thus, $t_{i,l}$ can be represented as $t_{(b,k),l}$. For each of the tree branches formed, the forward-backward algorithm is applied to estimate the likelihood of class l by defining a forward term, $\zeta_{(b,k)}(l)$, and a backward term, $\eta_{(b,k)}(l)$:

$$\zeta_{(b,k)}(l) = P(\bar{f}_{(b,k)}, t_{(b,k),l} = 1 | \mu_l, \sigma_l^2) \quad 1 \leq l \leq N_s \quad (1)$$

$$\eta_{(b,k)}(l) = P(\bar{f}_{(b,k)} | t_{(b,k),l} = 1, \mu_l, \sigma_l^2) \quad 1 \leq l \leq N_s \quad (2)$$

where $\bar{f}_{(b,k)}$ is the set of the observed mean intensity features from the start of the tree branch to region $r_{(b,k)}$, and $\bar{f}_{(b,k)}$ is the set of the observed mean intensity features from region $r_{(b,k+1)}$ to the end of the branch. We can solve both terms at each region $i = (b,k)$ inductively by:

$$\zeta_{(b,k+1)}(l) = [\sum_{m=1}^{N_s} \zeta_{(b,k)}(m) \alpha_{n_i,l}] P(f_{(b,k+1)} | \mu_l, \sigma_l^2) \quad k=1, \dots, \text{length}-1 \quad 1 \leq l \leq N_s \quad (3)$$

$$\eta_{(b,k)}(l) = \sum_{m=1}^{N_s} [\eta_{(b,k+1)}(m) \alpha_{n_i,l} P(f_{(b,k+1)} | \mu_m, \sigma_m^2)] \quad k=\text{length}-1, \dots, 1 \quad 1 \leq l \leq N_s \quad (4)$$

where *length* is the number of regions in the branch. The forward term is initialized as $\zeta_{(b,1)}(l) = P(f_{(b,0)} | \mu_l, \sigma_l^2)$ and the backward term is initialized as $\eta_{(b,\text{length})}(l) = 1/N_s$. The likelihood, $t_{i,l}$, is thus calculated based on both the forward and the backward terms as:

$$t_{i,l} = t_{(b,k),l} = [\zeta_{(b,k)}(l) \eta_{(b,k)}(l)] / \sum_{m=1}^{N_s} [\zeta_{(b,k)}(m) \eta_{(b,k)}(m)] \quad (5)$$

Once all region likelihoods are found, the algorithm then re-evaluates the model parameters $[\mu_l, \sigma_l^2, \alpha_{n_i,l}]$ for each state $l \in \{ 1, 2, \dots, N_s \}$ using probability weighting, by assuming a linear model such that the estimated region class probabilities represent the underlying partial volumes. A new Markov tree is then constructed from a random region and the estimation process repeats. The convergence criterion is defined as a minimum mean absolute change or a maximum number of iterations reached.

While using watershed regions to divide the image into contiguous homogeneous regions as proposed by Huang *et al.* [5] works well for discrete classification, the fuzzy approach can benefit from further subdivision along the region boundaries. The boundary between two watershed regions represents an optimal sharp division based on intensity gradients, and in the discrete rbHMM framework provides an adequate level of detail because each region only receives one class label (Fig. 1a). However, in our proposed fuzzy rbHMM framework, a finer resolution around the region boundaries would allow for superior capturing of gradient changes due to partial volume effects (Fig. 1b). Thus, we first pre-segment an image using the watershed transform. Once the watershed regions are established, we refine them by assigning voxels around all region boundaries as individual regions so as to provide increased boundary resolution for the likelihood estimation procedure (Fig. 1c).

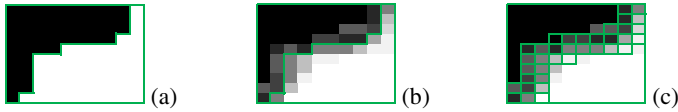


Fig. 1. Image (grayscale) and watershed subdivision (green) for (a) ideal and (b) partial-volume edges with two regions. By assigning boundary voxels as individual regions (c), we increase the resolution for subsequent estimations.

3 Results and Discussion

We tested our fuzzy partial-volume classification technique, *frbHMM*, on 3D MR brain images and compared the performance to that of the discrete approach, *rbHMM* [5]. We first validated on simulated T1-weighted BrainWeb scans ($181 \times 217 \times 181$ dimensions, $1\text{mm} \times 1\text{mm} \times 1\text{mm}$ spacing, 0% noise) of normal and MS anatomical models to quantify our accuracy as the ground truths are available [16]. We then applied our method to real clinical 3DT1 MR scans ($256 \times 256 \times 120$ dimension, $0.837\text{mm} \times 0.837\text{mm} \times 1.10\text{mm}$ spacing) of 18 relapse-remitting MS (RRMS) patients and 14 healthy controls from the UBC MRI Research Centre to demonstrate the robustness of the proposed method in maintaining control/subject measurements at reduced resolution, and therefore, increased partial-volume effects. We performed a 4-class segmentation - background, white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF). For MS data, lesions cause errors in CSF and/or GM classification due to intensity overlap, but *frbHMM* remains robust for WM and more importantly, does not show unexpected stability problem when model assumptions are violated. *frbHMM* required approximately twice the runtime as *rbHMM*. Results of classification improve iteratively, but for the purpose of fair comparisons, both discrete and fuzzy estimations were run for 10 iterations.

3.1 Simulated Images

We examined the segmentation accuracies of both *rbHMM* and *frbHMM*. Fig. 2 shows qualitative results obtained by both methods. For MS data, both methods classified periventricular lesions as GM; however, *frbHMM* performed better than *rbHMM* in capturing the partial WM details of ambiguous regions (Fig. 2). Quantitatively, Table 1 shows that mean square errors (MSE) of the segmentation results. At the original resolution, *frbHMM* achieved approximately 20% lower MSE for both WM and GM results. Analyzing the MSE gains of partial volume regions ($0 < \text{probability} < 1$) yielded similar performance. To simulate progressively reduced resolution, hence increased partial volumes, we performed smoothing by using a Gaussian kernel ($7 \times 7 \times 7$ dimension) with varying standard deviations to incorporate intensities within $\pm 3\text{mm}$. Again, *frbHMM* was superior (about 40% MSE reduction).

3.2 Clinical Scans

Next, we tested both *rbHMM* and *frbHMM* classification accuracy on high resolution 3DT1 scans of real clinical control (C) and subject (S) groups. With no ground truths available, in order to evaluate the robustness of the proposed *frbHMM* approach with increased partial-volumes, we examined the WM volume fraction measurements

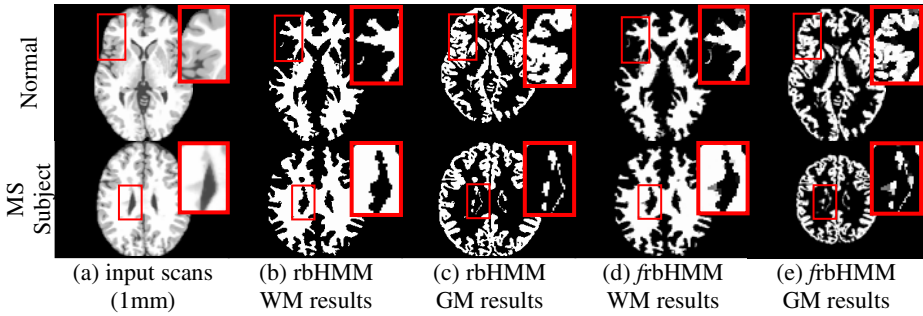


Fig. 2. Experimental results on simulated BrainWeb MRI scans. (a) Input T1-weighted scans at 1mm slice thickness, (b-c) results of discrete rbHMM, and (d-e) results of proposed frbHMM.

Table 1. Quantitative MSE of simulated BrainWeb MRIs of normal and MS subject anatomical models. Classifications were performed using discrete and fuzzy rbHMMs, and results were compared to the discrete and fuzzy phantoms, respectively. The better performance for each comparison is highlighted in bold. Gain is defined as (rbHMM-frbHMM)/rbHMM.

Data	Tissue	Original Scan			Gaussian kernel std. dev. = 1.0 voxel			Gaussian kernel std. dev. = 2.0 voxel		
		rbHMM	frbHMM	% gain	rbHMM	frbHMM	% gain	rbHMM	frbHMM	% gain
Normal	WM	0.0089	0.0069	22.47	0.0217	0.0133	38.71	0.0352	0.0213	39.49
	GM	0.0149	0.0119	20.13	0.0364	0.0213	41.48	0.0572	0.0345	39.69
MS Subject	WM	0.0090	0.0069	23.33	0.0222	0.0136	38.74	0.0348	0.0213	38.79
	GM	0.0151	0.0121	19.87	0.0364	0.0213	41.48	0.0565	0.0345	38.94

($WMVF = V_{WM} / V_{BRAIN} \times 100\%$, where V_{WM} and V_{BRAIN} are the WM and intradural volumes respectively). We targeted the WM as it is the largest tissue by volume and the pathological regions in patient scans are the most distinct in WM and thus can be readily excluded. We first demonstrate that both methods (rbHMM and frbHMM) performed comparably in terms of the WMVF measure when using the high resolution scans. However, we then demonstrate how the proposed frbHMM is far more robust when the image resolution is reduced.

Table 2 shows the average WMVF obtained by using the original (high resolution) scans. Both methods showed that the average WMVF in the controls were significantly ($p < 0.05$) higher (+3.53% for discrete, +3.11% for fuzzy) than those of the

Table 2. Quantitative average WMVF measurement on high resolution clinical MRI data. Note that both the rbHMM and frbHMM segmentation methods show consistent and significant differences between the control and the subject groups, while no significant differences were observed between the two methods for either group, which is expected in high resolution data.

	rbHMM	frbHMM	Δ (methods)	p -Value
Controls (C)	42.57 ($\sigma=1.66$)	41.98 ($\sigma=1.43$)	0.58	0.33
Subjects (S)	39.03 ($\sigma=2.82$)	38.88 ($\sigma=2.59$)	0.15	0.87
Δ (groups)	3.53	3.11		
p -Value	<0.05	<0.05		

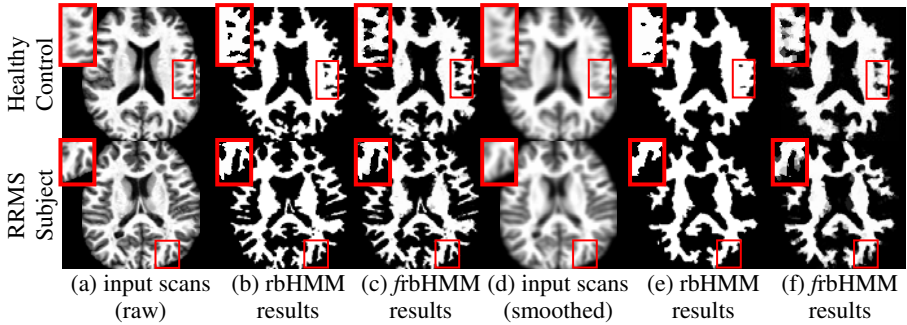


Fig. 3. Example results on real clinical MRI scans. (a-c) Input 3DT1 scans at original resolution and results of discrete and fuzzy rbHMMs, (d-f) smoothed 3DT1 scans (kernel standard deviation=2.0 voxel) and results of discrete and fuzzy rbHMMs.

Table 3. Quantitative segmentation comparisons on real clinical data with escalated partial volume effects. Classifications were performed using discrete and fuzzy rbHMMs, and results were compared to the segmentations obtained based on the original high resolution scans. The superior performance is highlighted in bold. Gain is defined as (rbHMM-frbHMM)/rbHMM.

Metric	Gaussian kernel std. dev. = 1.0 voxel			Gaussian kernel std. dev. = 2.0 voxel		
	rbHMM	frbHMM	% gain	rbHMM	frbHMM	% gain
MSE (C)	0.0159 ($\sigma=0.0018$)	0.0102 ($\sigma=0.0010$)	35.89 ($p<0.05$)	0.0250 ($\sigma=0.0027$)	0.0179 ($\sigma=0.0022$)	28.40 ($p<0.05$)
MSE (S)	0.0155 ($\sigma=0.0015$)	0.0100 ($\sigma=0.0008$)	35.48 ($p<0.05$)	0.0244 ($\sigma=0.0022$)	0.0173 ($\sigma=0.0016$)	29.10 ($p<0.05$)
MSD (C)	11.37 ($\sigma=6.33$)	2.29 ($\sigma=3.10$)	79.86 ($p<0.05$)	35.48 ($\sigma=11.05$)	18.06 ($\sigma=10.30$)	49.10 ($p<0.05$)
MSD (S)	6.72 ($\sigma=5.97$)	1.43 ($\sigma=1.75$)	78.72 ($p<0.05$)	22.71 ($\sigma=13.74$)	8.55 ($\sigma=8.24$)	62.35 ($p<0.05$)

RRMS patients due to the presence of WM lesions and enlarged ventricles in MS patients. Comparing the two methods within groups, no significant evidence ($p \geq 0.05$) of differences in WMVF was observed demonstrating that both methods achieved similar performances on high resolution data.

Similar to the simulated scans case, we then progressively reduced the image resolutions by Gaussian smoothing. Fig. 3 shows the effects of such smoothing on the classifications of control and subject scans. Again, the proposed frbHMM classified ambiguous intensity regions (zoomed-in regions of Fig. 3) with greater accuracy than the discrete model. Quantitatively, Table 3 shows the average MSE (image-to-image comparison) and the mean square difference or MSD (WMVF comparison) of the control and subject groups. The proposed frbHMM consistently achieved significantly ($p < 0.05$) lower MSE (approximately 30% improvement) and MSD for both groups.

4 Conclusions

We introduced a novel fuzzy 3D region-based hidden Markov model for modeling and estimation of partial volumes in image voxels within an unsupervised framework.

The paper's main contribution is a new *fuzzy* 3D HMM framework based on irregularly-shaped homogeneous regions with further spatial refinement at the region boundaries, where the contents of each region are assigned to multiple underlying classes simultaneously rather than assuming a single true discrete label. To compute the region class likelihoods, we employ a classical iterative forward-backward scheme. We evaluated the classification accuracy and robustness of our method under increased partial volume effects using both simulated and real clinical brain MRI data of healthy controls and MS subjects. Results showed the proposed *frbHMM* approach to be consistently superior to the discrete *rbHMM* in labeling intensity-ambiguous regions, such as white matter with reduced signal due to pathology. Future work includes investigating the utility of our *frbHMM* approach for quantifying the extent of white matter disease such as diffusely-abnormal white matter (DAWM).

References

1. Bishop, C.M.: Pattern Recognition and Machine Learning, pp. 359–422. Springer, Heidelberg (2006)
2. Li, J., Najmi, A., Gray, R.M.: Image classification by a two dimensional hidden Markov model. *IEEE Trans. Signal Proc.* 48(2), 517–533 (2000)
3. Ibrahim, M., John, N., Kabuka, M., Younis, A.: Hidden Markov models-based 3D MRI brain segmentation. *Image and Vision Computing* 24, 1065–1079 (2006)
4. Joshi, D., Li, J., Wang, J.Z.: A comp. efficient approach to the estimation of two- and three-dimensional hidden Markov models. *IEEE Trans. Image Proc.* 15(7), 1871–1886 (2006)
5. Huang, A., Abugharbieh, R., Tam, R.: Image segmentation using an efficient rotationally invariant 3D region-based hidden Markov model. In: *IEEE CVPR, Workshop on Mathematical Methods in Biomed. Image Analysis (MMBIA)*, Anchorage-Alaska, pp. 107–111 (2008)
6. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imag.* 20(1), 45–57 (2001)
7. Besag, J.: On the statistical analysis of dirty pictures (with discussion). *J. of Royal Statist. Soc., Ser.B* 48(3), 259–302 (1986)
8. Komodakis, N., Tziritas, G., Paragios, N.: Fast, approximately optimal solutions for single and dynamic MRFs. In: *IEEE CVPR, Minneapolis-Minnesota* (2007)
9. Kolmogorov, V.: Convergent tree-reweighted message passing for energy minimization. *IEEE Trans. Pattern Anal. Mach. Intell.* 28(10), 1568–1583 (2006)
10. Rabiner, L.R.: An intro. to hidden Markov model. In: *IEEE ASSP Magazine*, pp. 4–16 (1986)
11. Santago, P., Gage, H.D.: Quantification of MR brain images by mixture density and partial volume modeling. *IEEE Trans. Med. Imag.* 12(3), 566–574 (1993)
12. Van Leemput, K.: A unifying framework for partial volume segmentation of brain MR images. *IEEE Trans. Med. Imag.* 22(1), 105–119 (2003)
13. Bricq, S., Collet, C., Armspach, J.P.: Unifying framework for multimodal brain MRI segmentation based on hidden Markov chains. *Medical Image Analysis* 12, 639–652 (2008)
14. Choi, H.S., Haynor, D.R., Kim, Y.: Partial volume tissue classification of multichannel magnetic resonance images – a mixel model. *IEEE TMI* 10(3), 395–407 (1991)
15. Domke, J., Karapurkar, A., Aloimonos, Y.: Who killed the directed model? In: *CVPR* (2008)
16. BrainWeb: Simulated Brain Database,
<http://www.bic.mni.mcgill.ca/brainweb/>