

# Enhanced Spatial Priors for Segmentation of Magnetic Resonance Imagery

Tina Kapur<sup>1\*\*</sup>, W. Eric L. Grimson<sup>1</sup>, Ron Kikinis<sup>2</sup>, and William M. Wells<sup>1,2</sup>

<sup>1</sup> MIT AI Laboratory, Cambridge MA, USA  
tkapur@ai.mit.edu

<http://www.ai.mit.edu/~tkapur>

<sup>2</sup> Brigham & Womens Hospital, Harvard Medical School, Boston MA, USA

**Abstract.** A Bayesian, model-based method for segmentation of Magnetic Resonance images is proposed. A discrete vector valued Markov Random Field model is used as a regularizing prior in a Bayesian classification algorithm to minimize the effect of salt-and-pepper noise common in clinical scans. The continuous Mean Field solution to the MRF is recovered using an Expectation-Maximization algorithm, and is a probabilistic segmentation of the image. A separate model is used to encode the relative geometry of structures, and as a spatially varying prior in the Bayesian classifier. Preliminary results are presented for the segmentation of white matter, gray matter, fluid, and fat in Gradient Echo MR images of the brain.

## 1 Introduction

The automatic segmentation of anatomical structures from medical images such as MRI or CT will likely benefit from the exploitation of four different kinds of knowledge: *intensity models* that describe the gray level appearance of individual structures (e.g. fluid appears bright in T2-weighted MRI), *relative geometric models* that describe the relative geometry of structures in a subject-specific reference frame (e.g. femoral cartilage is attached to the subject's femur), *shape models* that describe the shape of individual structures in a subject-independent reference frame (e.g. the brain-stem is tube-like), as well as *imaging models* that capture the relevant characteristics of the imaging process.

EM-Segmentation, a segmentation method for MRI images [1], employed Gaussian intensity models for the different tissue classes, and used an imaging model to account for some distortions of the signal that are unique to the MRI process. The work reported here continues along that theme, with two key parts. The first contribution is the addition of a regularizer to the *imaging model* used in the EM-Segmentation algorithm. Regularization combats salt-and-pepper noise

---

\*\* The authors would like to thank Martha Shenton for contributing data for this paper. W. Wells received support for this research in part from NIMH Research Scientist Development Awards K02 MH-01110 and R29 MH-50747 (Martha Shenton, PI) and from a Whitaker Foundation Biomedical Research Grant (W. Wells, PI). R. Kikinis received partial support from: NIH: RO1 CA 46627-08, PO1 CA67165-01A1, PO1 AG04953-14, NSF: BES 9631710 and Darpa: F41624-96-2-0001.

common in clinical scans. Previous implementations of EM-segmentation deal with this noise effectively by pre-processing the images with structure-preserving intensity smoothers, particularly gradient-limited diffusion methods [2, 3]. These methods are quite effective, but computationally costly, and not trivial to adjust. We leverage the Bayesian flavor of EM-Segmentation and regularize via a *prior* distribution on the labeling, without incurring undue additional computational cost. Specifically, we model the prior distribution as a Markov Random Field (MRF), and recover its Mean Field (MF) solution using the Expectation-Maximization algorithm. While MF approximations of MRFs have previously been used in computer vision, we believe that the reported work is novel in its use of this prior in conjunction with the EM-Segmentation algorithm.

In the second component, we propose an algorithm that leverages geometric relationships between structures for segmentation purposes. We observe that some structures can be directly segmented from medical images by using methods from low-level computer vision, e.g. skin surface is reproducibly segmented from head MRI using a combination of thresholding, connectivity, and morphological operations. Other structures, such as the brain tissue in head MRI, do not have as salient a combination of intensity and topology as the skin, and are harder to segment using low-level methods. We propose a “coarse to fine” strategy in feature (structure) space – a strategy in which the easily identifiable (“coarse”) structures are first segmented automatically and their geometry is then used to bootstrap the segmentation of other (“fine”) structures in the image. We present an implementation of this strategy in the form of a *relative geometric prior* (prior distribution on the geometry of “fine” structures, given the geometry of “coarse” structures), and integrate it into the EM-Segmentation algorithm along with the regularizing MRF prior summarized earlier.

Combining the two components, the contribution of this paper may be summarized as the enhancement of the EM-Segmentation algorithm using two priors: an MRF prior to encode piecewise-homogeneity of labels, and a spatial prior to encode the relative geometry of structures.

## 2 Background on EM Segmentation

**Expectation-Maximization (EM)** The EM algorithm is an iterative scheme for estimating the parameters of a model that maximize the likelihood of the observed signal. The key step in applying the EM algorithm is to identify a set of *hidden* variables, such that it becomes possible to directly compute the maximum-likelihood estimate of the model using the values of the observed variables and these hidden variables. Once the hidden variables are identified, and the model parameters initialized, the EM algorithm alternates between estimating the hidden variables (as the expected values of the hidden variables using the estimates of the model parameters; the E-step) and the model parameters (as the maximum-likelihood estimates of the model given the observed and hidden variables; the M-step). Each iteration improves the model estimate [4], and the EM algorithm converges to a local minimum of the likelihood function.

**EM-Segmentation** Segmentation of MRI images is a challenging problem due to the presence of a non-linear gain field attributable to inhomogeneities in the imaging equipment. The EM-Segmentation algorithm [1], approached the segmentation of MRI images as a *maximum likelihood* estimation problem and used the Expectation-Maximization algorithm [4] to simultaneously estimate the class label and gain at each voxel that maximize the likelihood of the observed signal.

The observed MRI signal was modeled as a product of the true signal generated by the underlying anatomy, and the non-linear gain artifact. Using this assumption, an iterative, supervised, Expectation-Maximization style segmentation algorithm was developed that treats the underlying label classes as hidden variables and alternates between estimating those classes (E-step) and the maximally probable gain field (M-step).

In this algorithm, intensity data is log-transformed, thus converting the multiplicative gain field to an additive bias field. Observed log intensity,  $Y_{ij}$ , at each pixel is modeled as a normal distribution, independent of all other pixels:

$$p(Y_{ij}|\Gamma_{ij} = k, \beta_{ij}) = N(Y_{ij} - \beta_{ij}; \mu_k, \sigma_k) \quad , \quad (1)$$

where  $N(x; \mu, \sigma)$  is the Gaussian distribution, with mean  $\mu$  and variance  $\sigma^2$ ;  $Y_{ij}$  is the observed log intensity at pixel location  $(i, j)$ ;  $\Gamma_{ij}$  is tissue class corresponding to intensity  $Y_{ij}$ ;  $\mu_k, \sigma_k$  are the mean and standard deviation in intensity for tissue class  $k$ ;  $\beta_{ij}$  is the bias field at pixel location  $(i, j)$ . The method used a spatially stationary prior probability on the tissue labels  $\Gamma$ :

$$P_{stat}(\Gamma) = \prod_{ij} p_{stat}(\Gamma_{ij}) \quad (2)$$

where  $p_{stat}(\Gamma_{ij})$  is the prior probability that a given voxel belongs to a particular tissue class. This prior probability is constant through the iterations. The bias field  $\beta$  is modeled as a multi-dimensional zero mean Gaussian random variable, to characterize its spatial smoothness.

The E-step computes the posterior tissue class probabilities,  $W_{ijk}$  (posterior probability of pixel  $ij$  belonging to tissue class  $k$ ), when the bias field is known:

$$W_{ijk} = \frac{p(Y_{ij}|\Gamma_{ij} = k; \beta_{ij})p_{stat}(\Gamma_{ij} = k)}{\sum_m p(Y_{ij}|\Gamma_{ij} = m; \beta_{ij})p_{stat}(\Gamma_{ij} = m)} \quad . \quad (3)$$

The M-Step computes the value of the bias field  $\beta$  that maximizes the average likelihood of observation, as  $\beta = FR$ , where  $R_{ij} = \sum_k \frac{W_{ijk}*(Y_{ij}-\mu_k)}{\sigma_k^2}$ , and  $F$  is a linear operator that can be approximated by smoothing filters. This step is equivalent to a MAP estimator of the bias field when the tissue probabilities  $W$  are known. Detailed derivations of these steps can be found in [1].

In the next two sections, we preserve this EM framework for iterating between tissue classification and bias field estimation, and present two different methods for computing spatially varying priors on tissue class to enhance the spatially stationary prior shown in Equation 2 and used in Equation 3 of the E-step.

### 3 Addition of Markov Prior

As noted, EM-Segmentation [1] uses a spatially stationary prior on tissue class, i.e. at each iteration, the prior probability that a voxel belongs to a particular tissue class remains constant, and is independent of the labels of voxels in its neighborhood. In this work, we incorporate a Markov prior on tissue class, under which the prior probabilities at a voxel are influenced by the labels in its immediate neighborhood. This prior model acts as a regularizer and biases the solution towards piecewise-homogeneous labelings. Such a regularizing prior is useful in segmenting scans corrupted by salt and pepper noise.

MRF priors have been used in computer vision to model smoothness as well as different textures (e.g., [5]). Typical solvers for MRFs include Gibbs sampling [6], the Metropolis algorithm [7], Iterated Conditional Modes (ICM) [8], and Mean-Field (MF) methods [9]. ICM solvers have been used for the segmentation of medical images [10, 11].

**MRF Formulation:** We describe the reformulation of the prior distribution on the tissue labels  $\Gamma$  (from Equation 2) as a Markov Random Field to represent the piecewise homogeneity and the local compatibility of different tissues. The MRF parameters are obtained from manually labeled training data, and its Mean Field solution is recovered using the EM framework of the previous section.

Some notation:  $S = \{S_{ij} | 1 \leq i \leq m, 1 \leq j \leq n\}$  is the lattice on which the MRF is defined, and each site of this lattice – referred to as either  $S_{ij}$  or simply  $ij$  – corresponds to the pixel location  $(i, j)$  in the image.  $N = \{N_{ij} | 1 \leq i \leq m, 1 \leq j \leq n\}$  defines the neighborhood system for the MRF, where  $N_{ij}$  refers to the four neighbors of pixel  $ij$  that share an edge with it, i.e.  $N_{ij} = \{S_{i,j-1}, S_{i,j+1}, S_{i-1,j}, S_{i+1,j}\}$ .

The tissue labels  $\Gamma = \{\Gamma_{ij} | S_{ij} \in S\}$  are modeled as an MRF with the neighborhood system  $N$  on the lattice  $S$ .  $\Gamma_{ij}$  is a discrete-valued random-vector drawn from the set  $\{[100 \dots 0]^T, [010 \dots 0]^T, \dots [000 \dots 1]^T\}$ , and assigning the value  $[0 \dots 1 \dots 0]^T$  (with a 1 in the  $k^{th}$  position) to  $\Gamma_{ij}$  is equivalent to assigning the  $k^{th}$  tissue class to the pixel  $ij$  in the image.  $\Gamma$  satisfies the Markov condition, given by:  $P(\Gamma_{ij} | S \setminus S_{ij}) = P(\Gamma_{ij} | N_{ij}), \forall ij$ , which means that the value of each random variable  $\Gamma_{ij}$  in the field depends only on its four neighbors.

The Hammersley-Clifford theorem established the Markov-Gibbs equivalence and states that the probability of a particular configuration  $\gamma$  of any MRF  $\Gamma$  can be computed using the Gibbs probability of its clique potentials (a clique is a subset of nodes of  $S$  that are each others neighbors) over  $N$ :

$$P(\Gamma = \gamma) = \frac{1}{Z} e^{-\sum_c V_c(\gamma)} \quad (4)$$

Here  $V_c$  is a clique potential which describes the prior probability of a particular realization of the elements of the clique  $c$ .

The spatially stationary prior on the labeled image  $\Gamma$ , given by Equation 2, may be interpreted as an MRF on the image lattice with zeroth order cliques, i.e. there is no interaction between lattice sites. We impose local spatial coherence

on the labeling  $\Gamma$  by using first-order cliques. Clique potentials are computed using an Ising-like model derived from training data, i.e., the prior probability of tissue classes  $k$  and  $l$  occurring adjacent to each other is computed from manually labeled images. Thus, the prior probability on the labeling,  $P_{mrf}(\Gamma)$ , is not spatially stationary; it interacts with the labels in its neighborhood system.

Computing the field configuration with the maximum Gibbs probability is computationally intractable, so we use a Mean Field approximation to the general Markov model. We approximate the values of the field  $\Gamma$  at neighboring sites by their statistical means, and rewrite  $P_{mf}(\Gamma)$ , a Mean Field approximation to  $P_{mrf}(\Gamma)$ , as a product of single site probabilities  $p_{mf}(\Gamma_{ij}, N_{ij})$ :<sup>1</sup>

$$P_{mf}(\Gamma) = \prod_{ij} p_{mf}(\Gamma_{ij}, N_{ij}) , \quad (5)$$

where the single site probability  $p_{mf}(\Gamma_{ij}, N_{i,j})$  is written as a product of the single-site prior  $p(\Gamma_{ij})$  and the probability of each clique involving pixel  $ij$ :

$$p_{mf}(\Gamma_{ij}, N_{ij}) = \frac{1}{Z} p(\Gamma_{ij}) \cdot P_{h-}(\bar{\Gamma}_{i,j-1}) \cdot P_{h+}(\bar{\Gamma}_{i,j+1}) \cdot P_{v-}(\bar{\Gamma}_{i-1,j}) \cdot P_{v+}(\bar{\Gamma}_{i+1,j}) \quad (6)$$

where  $\bar{\Gamma}_{ij}$  is a continuous MF approximation to  $\Gamma_{ij}$ ,  $\cdot$  is component-wise vector multiplication, and  $Z$  is a normalizing constant. Equations 5 and 6 describe a general model for a discrete first-order pseudo-Ising MRF, using a particular mean-field approximation. To apply this model we must choose specific representations for  $p$  and each of the four neighborhood terms  $P_{h-}$ ,  $P_{h+}$ ,  $P_{v-}$ , and  $P_{v+}$ . In addition we need to supply values for the mean field  $\bar{\Gamma}$  in Equation 6.

**Using MRF Prior on Tissue Classes in EM Framework:** We incorporate the model into the EM framework in the following way: For the single site probability  $p$  we use the independent stationary prior,  $P_{stat}$ , that was formerly used in the work described in Section 2. For each of  $P_{h-}$ ,  $P_{h+}$ ,  $P_{v-}$ , and  $P_{v+}$ , we use a model based on the empirical joint probability of neighboring pixels. For  $\bar{\Gamma}$  we use the estimates on per-pixel tissue probability produced by the previous iteration of the EM segmentation algorithm as described in [1].

Since the basic EM segmentation algorithm is already computing such tissue probabilities, the net effect is a simple modification to the E-step that relates tissue co-occurrence statistics to the tissue probabilities that were computed on neighboring pixels in the previous iteration.

Since  $\Gamma$  is a random variable drawn from unit vectors, we may write the probability distribution modeled by  $P_{h-}$  as a linear form  $P_{h-}(g) \equiv \Lambda_{h-}g$ , where  $\Lambda_{h-}$  is a  $k \times k$  matrix, where  $k$  is the number of tissue classes in the training data, and its  $mn^{th}$  element  $\lambda_{h-,mn}$  gives the prior probability of tissue class  $m$  and  $n$  occurring as a horizontal pair in the image, with the left pixel in the pair being tissue class  $n$  and the right one being  $m$ . The distributions  $P_{h+}$ ,  $P_{v-}$ , and

<sup>1</sup> A different tractable MRF model for brain segmentation, using an Iterated-Conditional-Modes solver has been explored in [10].

$P_{v+}$ , and the corresponding  $A$ 's are defined in a similar fashion. We use this representation for the neighborhood probabilities in Equation 6 since it may be directly evaluated on the more general vectors  $\bar{I}$ .

## 4 Addition of Relative Geometric Prior

In this section we describe the second component of our work: a method that leverages geometric relationships between structures for segmentation purposes. The motivating observation is that while some structures are easily segmented using low-level computer vision methods (*primary* structures), there are structures whose segmentation is facilitated by knowledge of their spatial layout (geometry) relative to other structures (*secondary* structures).

**Summary of Method:** In order to use the relative geometry information for segmentation of a secondary structure in a given image, we first *identify a set of primitives* in terms of which to define its local geometric relationship to one or more primary structures. For example, the *distance* between points on the outside surface of structure  $P$  and the closest points on the inside surface of structure  $B$  is a primitive that describes local relative geometry of the two surfaces. Next, we *construct the relative geometric model* from training (segmented) images. In order to do this, a random variable is defined for each primitive, and segmented images are used to construct an empirical joint probability distribution over these random variables. This probability distribution serves as a model of the relative geometric relationship between the primary and secondary structures in question. For example, if one primitive is the distance between the *outer* surface of  $P$  and the outer surface of  $S$ , and another is the distance between the *inner* surface of  $P$  and the outer surface of  $S$ , then two random variables  $d_1$  and  $d_2$  are defined, one for each primitive relationship, and an empirical joint probability distribution for  $d_1$  and  $d_2$  is constructed from the segmented images. This joint serves as the relative geometric model for structures  $P$  and  $S$ . Following the construction of the model, we *segment primary structures* in the given image using appropriate algorithms. Finally, we use the geometric model as a *prior on the spatial layout* of the secondary structure, conditioned on the geometry of the segmented primary structures and used the EM-Segmentation algorithm to segment the secondary structure in question.

Note that this method is profitably used for the segmentation of a pair of structures in which one is a primary structure and the geometric relationship between the pair is informative (in an information theoretic sense). If either constraint is violated (neither of the structures is primary, or the relationship is uninformative), this method does not help the resulting segmentation.

**Previous Work:** A similar relative geometric prior was used in a traditional Bayesian classifier to segment femoral cartilage from Knee MRI images [12]. Also, this work is similar in spirit to landmark based segmentation [13], and different in its detection of a *dense* set of features as the landmarks.

### Example Usage of Method to Segment Brain Tissue from MRI Images

We observe that the skin surface and the ventricles are easily segmented in head MRI images, and use those as primary structures for segmentation of brain tissue (white matter and gray matter); the relationship between brain tissue and these primary structures is well described using two primitives:  $ds$ , the distance to the inside skin surface, and  $dv$ , the distance to the outside ventricle surface.

Next, we detail the algorithm for constructing the aforementioned empirical geometric model, and its usage with the EM-Segmentation algorithm for segmentation of white matter.

**Empirical Joint Density Estimation:** Example images in which the skin, the ventricles, and white matter have been manually labeled by experts are used to construct a non-parametric estimate for this joint density function. In particular, chamfer distance transforms[14] are computed for the inside skin surface and for the outside ventricle surface. These chamfer maps are used to find  $ds_i$  and  $dv_i$ , the distance to skin and ventricle surfaces for all pixels  $i$  that are labeled white matter, and the values are histogrammed jointly. The histogram is normalized to obtain an empirical estimate of the joint density of  $ds$  and  $dv$  for white matter. Note that instead of histogramming the values of the random variables, methods such as Parzen Windowing [15] could be used effectively for density estimation.

**Usage with EM-Segmentation:** The class conditional density is thus:

$$P(ds_i, dv_i | x_i \in WM) \quad (7)$$

where  $x_i$  are the spatial coordinates of the  $i$ th data pixel;  $WM$  is the set of all pixels belonging to white matter;  $S$  is the set of all pixels belonging to the skin;  $V$  is the set of all pixels belonging to the ventricles;  $ds_i$  is short for  $ds_i(S)$ , which is the distance from  $x_i$  to the inside surface of the skin;  $dv_i$  is short for  $dv_i(V)$ , which is the distance from  $x_i$  to the outside surface of the ventricles.

Bayes rule allows us to express the posterior probability that a pixel should be classified as white matter based on observations of its intensity and spatial relation to the skin and the ventricles ( $P(x_i \in WM | ds_i(S), dv_i(V), I_i)$ ) as a product of the prior probability that a given pixel belongs to white matter ( $P(x_i \in WM)$ ) and the class conditional density  $P(ds_i(S), dv_i(V), I_i | x_i \in WM)$  as follows:

$$P(x_i \in WM | ds_i, dv_i, I_i) = \frac{P(ds_i, dv_i, I_i | x_i \in WM)P(x_i \in WM)}{P(ds_i, dv_i, I_i)} \quad (8)$$

where  $I_i$  is the intensity at  $x_i$ , and the other terms are as in Equation 7. This expression may be rewritten assuming independence between the intensity at a pixel and its spatial relationship to skin and ventricles as:

$$P(x_i \in WM | ds_i, dv_i, I_i) = \frac{P(ds_i, dv_i | x_i \in WM)P(I_i | x_i \in WM)P(x_i \in WM)}{P(ds_i, dv_i, I_i)} \quad (9)$$

The first term in the numerator is the class conditional density for the model parameters, and is estimated using the method described above. The second term is a Gaussian intensity model for tissue class, obtained from samples of white

matter intensity. The third term is the prior probability that a pixel belongs to white matter, computed as a ratio of white matter volume to total head volume in a segmented scan. The denominator is a normalization factor.

This spatial probability distribution (Equation 9) can be used either in conjunction with the Mean-Field prior of Section 3, or by itself, instead of the spatially stationary prior in the E-step of the EM-Segmentation algorithm.

The above method is repeated to obtain a segmentation of gray matter.

## 5 Results

We have used the work presented in this paper to classify several images from different Gradient Echo brain MRI scans. Two examples are described here.

**Gradient Echo Brain MRI with MF:** Figure 1 shows the results of EM Segmentation using a Mean-Field prior on a sagittal slice of a Gradient Echo brain MRI. In the left column of the figure, the top image is the gray scale slice with additive white noise. The second image, provided as a baseline, is its classification (gray – gray matter, white – white matter, black – air/csf, red – skin/fat) that was obtained using a standard MAP classifier. The third image in the first column is the classification obtained using EM Segmentation with a spatially stationary prior, and the fourth image is the classification obtained using EM Segmentation with a Mean Field prior. Notice that the segmentation that uses the Mean-Field prior is much less fragmented compared to the segmentation that uses only the spatially stationary prior. Since each of these segmentations is obtained by thresholding the respective weights ( $W_{ijk}$  from Equation 3) associated with each tissue class, the middle and the right column of the figure show the weights for each tissue class (gray matter, white matter, csf/air, skin/fat) when the spatially stationary prior and Mean-Field prior are used, respectively. Again, the point to note is the lack of fragmentation when the MF prior is used.

**Gradient Echo Brain MRI with MF and Conditional-Spatial Priors:** Figure 2 shows the results of EM Segmentation using a Spatial-Conditional prior in conjunction with a Mean-Field prior on a coronal slice of a Gradient Echo brain MRI. In the left column of the figure, the top image is the grayscale slice. The second image, provided as a baseline, is its classification (gray – gray matter, white – white matter, black – air/csf, red – skin/fat) that was obtained using a standard MAP classifier. The third image in the first column is the classification obtained using EM Segmentation with a spatially stationary prior, and the fourth image is the classification obtained using EM Segmentation with Spatial-Conditional and Mean Field priors. Notice that the segmentation that uses the relative spatial priors is much less fragmented, and shows improved distinction between skin and brain tissue, as well as in the segmentation of white matter in the brain stem, compared to the segmentation that uses only the spatially stationary prior. Since each segmentation is obtained by thresholding the respective weights ( $W_{ijk}$  from Equation 3) associated with each tissue class, the middle and the right column of the figure show the weights for each tissue class (gray matter, white matter, csf/air, skin/fat) when the spatially stationary



prior and Mean-Field prior are used, respectively. Again, the point to note is the lack of fragmentation due to the MF prior, and the improved distinction between brain tissue and skin as well as improved segmentation of white matter.

## 6 Discussion

**Tradeoff between Prior and Observation:** A characteristic of Bayesian methods is the delicate balance that needs to be maintained between the influence of the prior term and fidelity to the observed data. If the degree of faith in the prior term is high (i.e. it models the underlying phenomenon accurately) and the observation is noisy, then conflicts between the prior and the observations are resolved in favor of the prior. In contrast, if there is negligible noise in the observations, then the prior can be discarded altogether, giving rise to a prior-less or maximum-likelihood solution. Unfortunately, it is often the case that the prior term is somewhat accurate, and the data is somewhat noisy i.e. it is not as clear how the two terms should be traded off in Bayes rule. The art of maintaining this balance is colloquially referred to as “tweaking the Bayesian fudge factor” and is arguably crucial to the success of the resulting algorithm.

Empirically speaking, in our case, the relative importance of the regularizing (Markov) prior is inversely proportional to the signal to noise ratio (SNR) in the MRI scan. Since SNR in MR scans is directly proportional to imaging parameters such as the strength of the ambient magnetic field, we weigh the prior by these parameters. For example, a scan acquired with a 0.5 Tesla magnet is segmented using a higher weight on the MRF prior, as compared with a scan acquired using a 1.5 Tesla magnet.

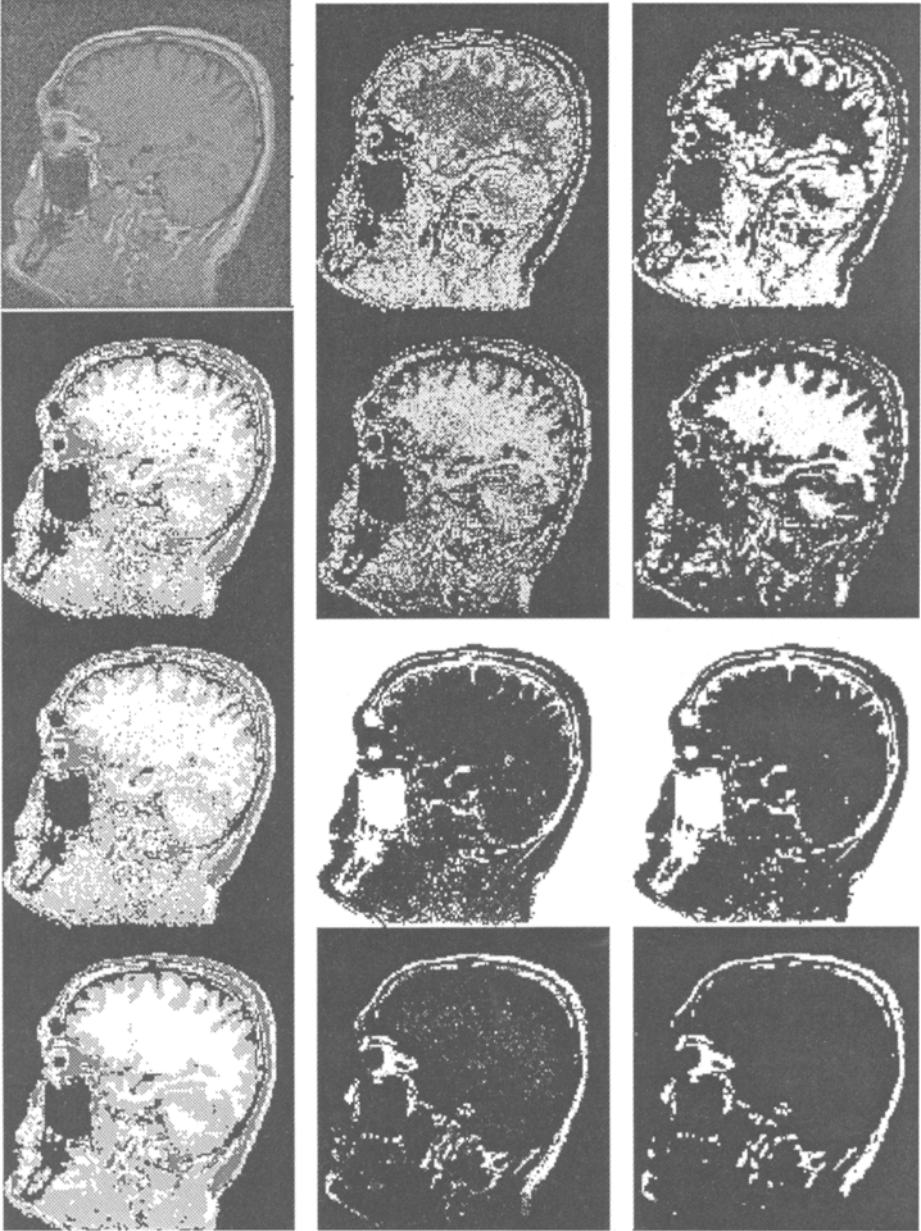
How best to characterize the weighing scheme for the geometric prior is less obvious. It would not be unreasonable, however, to measure the variation of the geometric prior model across individuals and assign a relative importance that is inversely proportional to that measure of variance. As a side-effect of this process, the variance in the relative geometric model could be used to characterize which structures this method is best suited for, and analyze its failure modes. Unless better schemes become apparent, this is the approach we plan to take for characterizing the importance of the geometric prior in classification.

**Moving onto 3D:** Since the images we are dealing with are inherently 3D volumes, the natural next step is to extend the reported priors by a dimension. While the 3D extension of the regularizing prior is simple to conceptualize and implement, the extension of the geometric prior to 3D will require a convention for normalizing images from different subjects, so that the prior usefully encodes information across a population. The Talairach coordinate system is a popular normalization method and a possible choice for us.

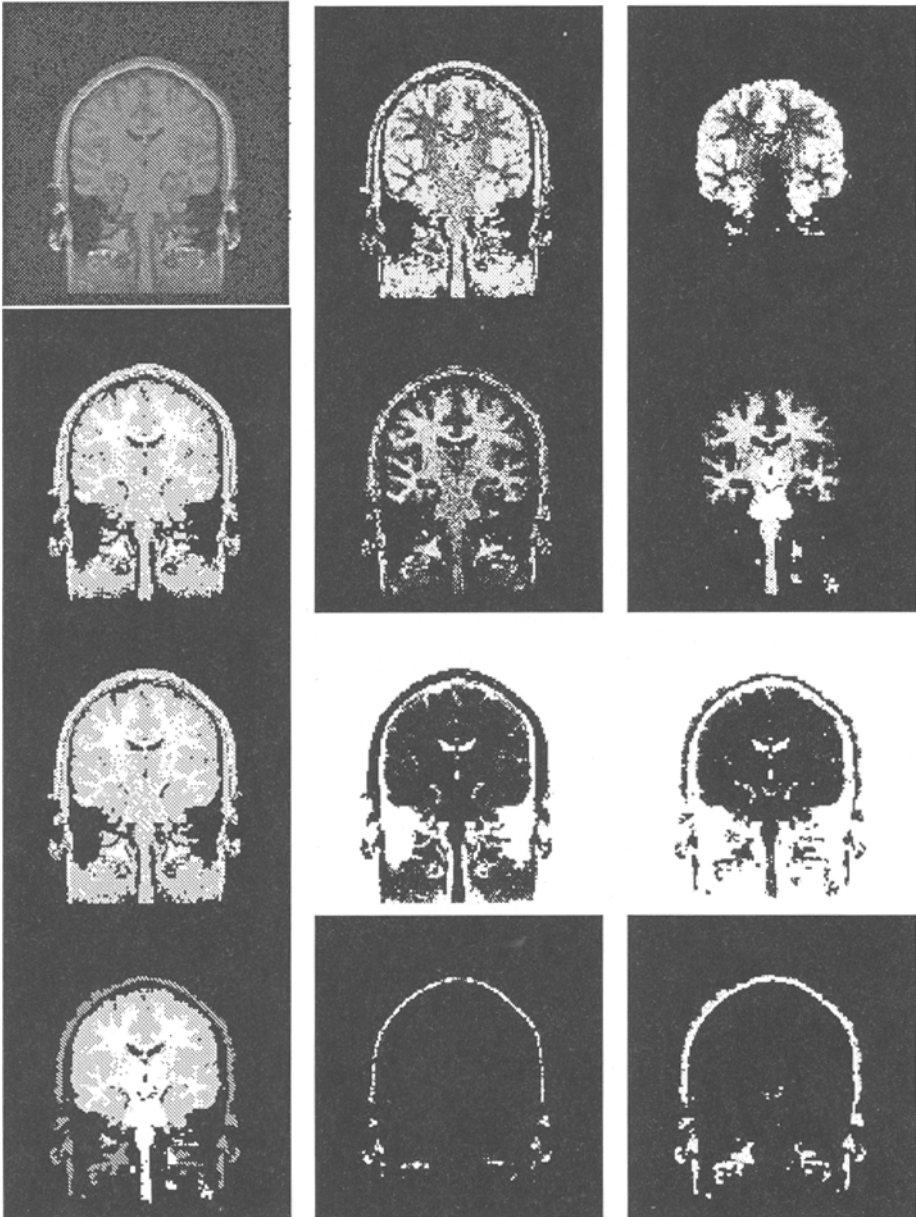
## References

1. W.M. Wells III, R. Kikinis, W.E.L. Grimson, and F. Jolesz. Adaptive segmentation of mri data. *IEEE Transactions on Medical Imaging*, 1996.

2. P. Perona, T. Shiotan, and J. Malik. Anisotropic diffusion. *Geometry-Driven Diffusion in Computer Vision*, pages 73–92, 1994.
3. R. Whitaker and G. Gerig. Vector-valued diffusion. In B.M. ter Haar Romeny, editor, *Geometry-Driven Diffusion in Computer Vision*, pages 93–133, 1994.
4. A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximal likelihood form incomplete data via the em algorithm. *RoyalStat*, B 39:1–38, 1977.
5. S.Z. Li. *Markov Random Field Modeling in Computer Vision*. Springer-Verlag, 1995.
6. S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *PAMI*, 6(6):721–741, November 1984.
7. N. Metropolis, A.W. Rosenbluth, A. N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.
8. J. Besag. On the statistical analysis of dirty pictures. *RoyalStat*, B-48(3):259–302, 1986.
9. D. Geiger and F. Girosi. Parallel and deterministic algorithms from mrfs: Surface reconstruction. *PAMI*, 13(5):401–412, May 1991.
10. K. Held, E. Rota Kopps, B. Krause, W. Wells, R. Kikinis, and H. Muller-Gartner. Markov random field segmentation of brain mr images. *IEEE Transactions on Medical Imaging*, 16:878–887, 1998.
11. D. Vandermeulen, X. Descombes, P. Suetens, and G. Marchal. Unsupervised regularized classification of multi-spectral MRI. In *Proceedings of the Fifth Conference on Visualization in Biomedical Computing*. SPIE, 1996.
12. T. Kapur, P.A. Beardsley, S.F. Gibson, W.E.L. Grimson, and W.M. Wells III. Model based segmentation of clinical knee mri. In *IEEE International Workshop on Model Based 3D Image Analysis*, 1998.
13. D. Collins, T. Peters, W. Dai, and A. Evans. Model based segmentation of individual brain structures from mri data. In R. Robb, editor, *Proceedings of the First Conference on Visualization in Biomedical Computing*, pages 10–23. SPIE, 1992.
14. G. Borgefors. Distance transformations in digital images. *Computer Vision, Graphics, and Image Processing*, 34:344–371, 1986.
15. R.O Duda and P.E. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, 1973.



**Fig. 1.** The results of segmentation using an MRF Prior within EM-Segmentation. The top left image is the input image, and the bottom image in the first column is its segmentation. The middle column shows the weights  $W_{ijk}$  (for gray matter, white matter, air/csf, skin respectively) upon convergence of the EM-Segmentation algorithm with spatially stationary priors. The right column shows the weights when the MRF prior is used with EM-Segmentation. See the text for a discussion of the results.



**Fig. 2.** The results of segmentation using a Relative Geometric Prior within EM-Segmentation. The top left image is the input image, and the bottom image in the first column is its segmentation. The middle column shows the weights  $W_{ijk}$  (for gray matter, white matter, air/csf, skin respectively) upon convergence of the EM-Segmentation algorithm with spatially stationary priors. The right column shows the weights when the geometric prior is used with EM-Segmentation. See the text for a discussion of the results.