Generalizing Diffusion Tensor Model Using Probabilistic Inference in Markov Random Fields

Çağatay Demiralp and David H. Laidlaw

Brown University
Providence, RI, USA

Abstract. We give a proof of concept for efficiently modeling tensor configuration distributions with Markov random fields (MRFs) and inferring the most likely tensor configurations with maximum a posteriori (MAP) estimations. We demonstrate the plausibility of our method by resolving fiber crossings in a synthetic dataset, experimenting with three different MAP estimation methods on a grid MRF model. The power of the MAP-MRF framework comes from its mathematical convenience in modeling prior distributions and the fact that it yields a global optimization driven by local neighborhood interactions.

1 Introduction

It is well known that the standard single second-order tensor model of diffusion falls short in quantifying fiber geometries such as crossing, kissing, and branching. This limitation has led to the development of different acquisition techniques along with more complex models of diffusion. One strategy for recovering the underlying complex fiber architecture is to nonparametrically estimate the displacement probability distribution function, which characterizes the diffusion locally, using the Fourier relationship between diffusion and the signal echo attenuation in q-space [1, 2]. These q-space methods, however, require a large number of scans, and hence long acquisition times, making them impractical in a clinical setting. Therefore, various regularization techniques have been proposed to sample the q-space efficiently and thereby reduce the number of scans required (e.g., [3–5]).

Some of earlier work has also proposed multi-tensor as well as simple, restricted geometric models to address the issue of mixed tract geometries (e.g., [6, 7]). In this context, the work presented here can be viewed as a regularized multi-tensor model. Given signal values (i.e., DWI sequences) and a set of prior constraints, we would like to know the most likely tensor configuration that "explains" the signal. This is a typical Bayesian inference problem. Therefore, we model the probability distribution of configurations with a Markov random field (MRF) and obtain the most probable configuration using the maximum a priori (MAP) estimate [8]. Note that, while it is applied to extend the tensor model here, as a general spatial probabilistic model selection approach, the MAP-MRF framework can be applied to other diffusion models easily.

2 Background and Related Work

2.1 Diffusion Imaging

The current standard method for measuring diffusion in tissues with magnetic resonance imaging (MRI) is based on an experiment using spin echoes in the presence of a time-dependent field gradient (the Stejskal-Tanner echo-sequence experiment [9]). This method samples from the Fourier transform of molecular displacement that is characterized by a probability density function $p(\mathbf{r}|\tau)$, where $\mathbf{r}$ is the displacement vector and τ is the diffusion time. In this context, the relation between the MRI signal $S(\mathbf{q}, \tau)$ measured in the direction of a diffusion gradient $\mathbf{G}$ and $p(\mathbf{r}|\tau)$ in a Stejskal-Tanner spin-echo sequence experiment is given by

$$S(\mathbf{q}, \tau) = S_0 \int_{R^3} p(\mathbf{r}|\tau) e^{-2\pi i \mathbf{q}^T \mathbf{r}} d\mathbf{r} \tag{1}$$

$$= S_0 \int_{R^3} p(\mathbf{r}|\tau) e^{-i\gamma\delta|\mathbf{G}|\mathbf{g}^T \mathbf{r}} d\mathbf{r} \tag{2}$$

where $\mathbf{q} = \gamma\delta\mathbf{G}/2\pi$, S_0 is a reference signal obtained without using a diffusion gradient, γ is the gyromagnetic ratio of water protons, δ is the diffusion gradient duration, and $\mathbf{g}$ is the unit vector in the direction of the diffusion gradient $\mathbf{G}$ (i.e., $\mathbf{g} = \mathbf{G}/|\mathbf{G}|$).

2.2 Diffusion Models

What we want to recover from equation (1) is the probability density function $p(\mathbf{r}|\tau)$ of molecular displacement. In this sense, modeling diffusion locally is a probability density estimation problem. There are essentially two basic approaches, parametric and nonparametric.

If we take the probability density function $p(\mathbf{g}|\tau)$ to be Gaussian, then the original equation (1) reduces to

$$S(\mathbf{g}, \tau) = S_0 e^{-b\mathbf{g}^T \mathbf{D} \mathbf{g}} \tag{3}$$

where $b = \tau|\mathbf{q}|^2 = \tau(\gamma\delta|\mathbf{G}|/2\pi)^2$ and $\mathbf{D}$ is a three-dimensional, second-order symmetric tensor. In this case, $p(\mathbf{r}|\tau)$ has the following form:

$$p(\mathbf{r}|\tau) = \frac{e^{-\frac{\mathbf{r}^T \mathbf{D}^{-1} \mathbf{r}}{4\tau}}}{\sqrt{(4\pi\tau)^3 |\mathbf{D}|}} \tag{4}$$

Since the mean of this three-dimensional normal distribution is assumed to be $(0, 0, 0)$, the covariance matrix (or the second-order symmetric tensor) $\mathbf{D}$ determines the distribution.

The single second-order tensor model of diffusion is the most widely used model quantifying anisotropic diffusion with MRI [10]. As remarked above, one of the limitations of this model is that it cannot handle mixtures of different tissue orientations. To address the issue, multi-tensor models have been proposed that turn the single Gaussian density estimation problem into the Gaussian mixture density estimation. (e.g., [11, 6]). An n-tensor model of diffusion is given by

$$S_i = S_0 \sum_{j=1}^{n} \alpha_j e^{-b\mathbf{g_i^T D_j g_i}} \tag{5}$$

where α is the mixing coefficient (or volume fraction) with $\sum_j \alpha_j = 1$.

Our method can be considered a multi-tensor model in which the number of tensors, and possibly tensor parameters, is estimated using Bayesian model selection. In the demonstration here, we use a two-tensor model with $\alpha_1 = \alpha_2 = 0.5$.

It is possible to estimate $p(\mathbf{r}|\tau)$ without assuming a particular shape for the distribution. All existing nonparametric approaches try to reconstruct a probability density function defined on the sphere. One approach is to use the Fourier relationship between the diffusion function and the signal echo attenuation in q-space. Diffusion spectrum imaging (DSI) samples diffusion on a volumetric grid [1], while qball imaging samples the density function on a sphere [2].

2.3 Markov Random Fields

A Markov random field is a conditional probability distribution with a Markov property over a set of random variables described by an undirected graphical model $G(V, E)$ with a set V of vertices and a set E of edges [8]. The vertex set V of the graph corresponds to the random variables and the edge set E determines the conditional dependencies (i.e., Markov properties). One of the advantages of MRFs is that it is easy to model local dependencies (or interactions) with arbitrary energy functions while ensuring that the probability distribution remains a proper probability distribution. Note that the Hammersley-Clifford theorem indicates that all MRFs have Gibbs representations, given $P(x) > 0$ [12].

MRFs have been used in the past to regularize diffusion tensor fields [13, 14]. The main difference between the two earlier works using MRFs and the present effort is that while we estimate the most likely configurations that can include multiple tensors for some voxels (crossing, branching, etc.), they estimate only the most likely single-tensor configurations. In other words, the MRF-based regularization method in [13, 14] does not resolve mixed tissue orientations. Also, at the technical level, we exploit more recent developments in MAP estimation algorithms: In addition to the greedy algorithm iterated conditional mode (ICM) used in [13, 14], we also experiment with belief propagation (BP) and tree-reweighted message passing (TRW) algorithms. Our results suggest that, even in the simple setting used, BP and TRW are more accurate and faster than ICM.

3 Methods

We express the problem of finding the most likely tensor configuration for a given sequence of diffusion images as a Bayesian inference problem. If x is a random variable representing the hidden tensor configuration and y is another random variable corresponding to the diffusion signal, then Bayes' rule suggests that

$$P(x|y) \propto P(y|x)P(x)$$
$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

where we model the prior distribution $P(x)$ using an MRF. In this context, the most likely tensor configuration is an assignment of x that maximizes the posterior probability $P(x|y)$.

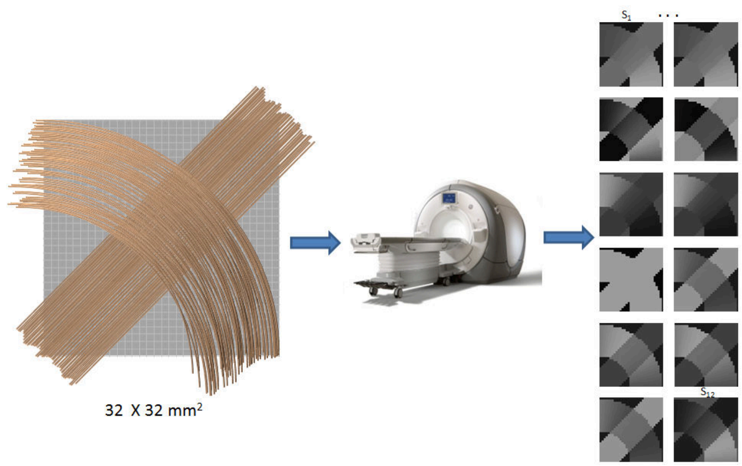


Fig. 1: Synthetic diffusion MRI data of resolution 32×32 were generated using $b = 1500$ s/mm^2 and 12 diffusion gradient directions.

4 MAP Estimation and Equivalent Energy Formulation

The maximum a posteriori (MAP) estimate is a choice of x that maximizes $P(x|y)$; MAP estimation in MRFs has an equivalent energy formulation [15] that we have used in our experiments. Let $V = \{v_1, v_2, \ldots\}$ be a set of random variables (e.g., corresponding to the voxels in a DWI sequence) and L be a finite set of labels (e.g., number of tensors to fit). Let x_i denote a labeling of v_i; then

$$E(\mathbf{x}) = \underbrace{\sum_{v_i \in V} D(x_i)}_{\text{data cost}} + \underbrace{\sum_{(i,j) \in N} W(x_i, x_j)}_{\text{smoothness cost}} \tag{6}$$

where $D(x_i)$ is the cost of assigning label x_i to v_i and $W(x_i, x_j)$ is the cost of assigning label x_i and x_j to two neighboring v_i and v_j. Minimizing $E(\mathbf{x})$ means finding a labeling assignment vector $\mathbf{x}$ that minimizes the cost.

5 Experiment

We demonstrate the use of the MAP-MRF framework on a synthetic fiber-crossing dataset.

5.1 Synthetic Data and MRF Model

We generate a synthetic fiber crossing diffusion MRI sequence using 12 gradient directions with $b = 1500$ s mm^{-2} [16]. We use a two-dimensional 4-connected grid MRF.

For our experiment, we use a simple smoothness function $W(x_i, x_j)$, where x_i and x_j are the random variables corresponding to the two neighboring nodes (voxels) i and j in the graphical model (MRF). Note that $x_i, x_j \in L = \{1, 2\}$ and, hence, $\mathbf{x} \in \{1, 2\}^{32\times 32}$.

We first define a "distance" between two tensors $\mathbf{D}_i$ and $\mathbf{D}_j$ using their eigenvectors:

$$\rho(\mathbf{D}_i, \mathbf{D}_j) = \sum_{k=1}^{3} \theta(\mathbf{e}_i^k, \mathbf{e}_j^k) \tag{7}$$

where $\mathbf{e}_l^k$ is the kth eigenvector of the tensor $\mathbf{D}_l$, and $\theta(\mathbf{u}, \mathbf{v}) = \arccos(|\mathbf{u}^T\mathbf{v}|)$ (the smaller angle between the two vectors). Therefore, ρ is basically the sum of the angles between corresponding eigenvectors of the two tensors.

Our smoothness function tries to capture how well the principal axes of tensors associated with one node are aligned with those of the other node:

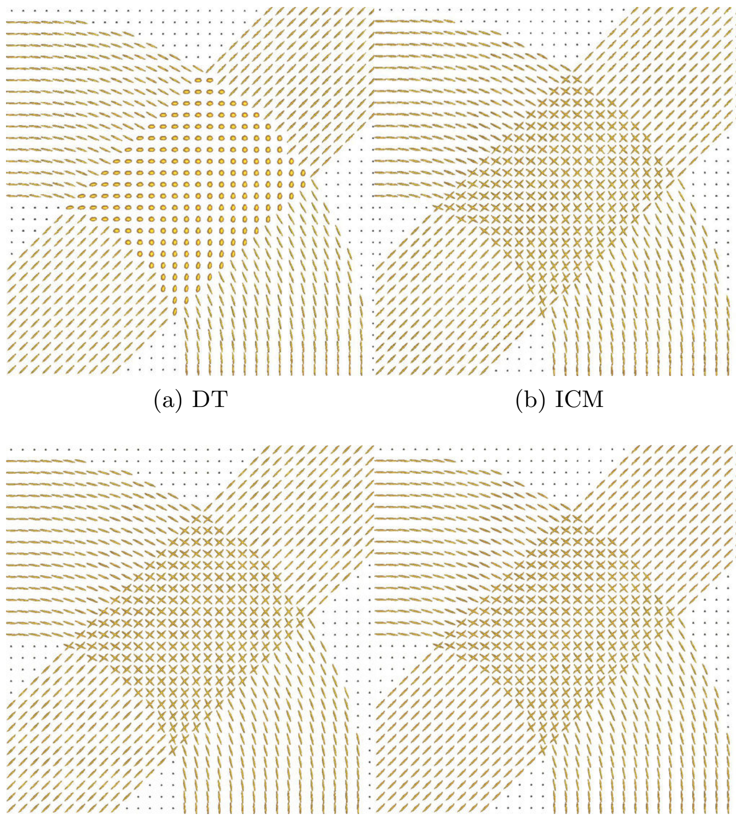
$$W(x_i, x_j) = \begin{cases} \rho(\mathbf{D}_i^1, \mathbf{D}_j^1) & \text{if } x_i = 1 \text{ and } x_j = 1 \\ \min(\rho(\mathbf{D}_i^1, \mathbf{D}_j^1), \rho(\mathbf{D}_i^1, \mathbf{D}_j^2)) & \text{if } x_i = 1 \text{ and } x_j = 2 \\ W(x_j, x_i) & \text{if } x_i = 2 \text{ and } x_j = 1 \\ \rho(\mathbf{D}_i^l, \mathbf{D}_j^m) + \rho(\mathbf{D}_i^{\{1,2\}-l}, \mathbf{D}_j^{\{1,2\}-m}) & \text{if } x_i = 2 \text{ and } x_j = 2 \end{cases}$$

where $(l, m) = \arg\min_{(l,m)} \rho(\mathbf{D}_i^l, \mathbf{D}_j^m)$.

The last case, where $x_i = 2$ and $x_j = 2$, needs some explanation. Here we have exactly two tensors $\{\mathbf{D}_i^1, \mathbf{D}_i^2\}$ and $\{\mathbf{D}_j^1, \mathbf{D}_j^2\}$ associated with x_i and x_j, respectively. First, we find the pair $(\mathbf{D}_i^l, \mathbf{D}_j^m)$ for which ρ is the minimum. Then we define W to be the sum of this minimum and the distance between the remaining pair.

The data cost term used in this case is also very simple:

$$D(x_i) = \begin{cases} 1 & \text{if } x_i = 1 \\ 2 & \text{otherwise} \end{cases}$$



(a) DT (b) ICM

(c) BP (d) TRW

Fig. 2: Results for a 2D synthetic fiber-crossing data set. While the single-tensor fitting (DT) cannot resolve the crossing, all the MAP estimation algorithms used (ICM, BP, and TRW) "select" the underlying orientation configuration correctly.

6 Results

We use three different MAP estimation algorithms: iterated conditional mode (ICM), belief propagation (BP), and tree-reweighted message passing (TRW) [15]. All the three algorithms, even one as simple as ICM, provide good results, effectively selecting the configuration that represents the fiber crossing accurately (see Figure 2).

7 Discussion and Conclusions

While diffusion is a local phenomenon, connectivity is global. Therefore, a model fitting for diffusion should consider goodness of fit in both aspects. We have

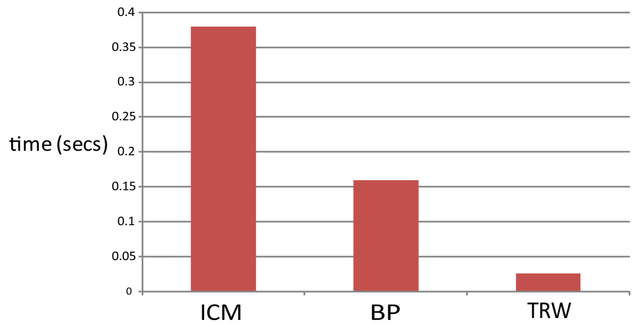


Fig. 3: Running times for the MAP estimation algorithms.

provided a proof of concept for efficiently modeling tensor configuration distributions with MRFs and their practical MAP estimations. We believe that the MAP-MRF framework can address the two main limitations of multi-compartment models: nonlinear optimization and model selection. Note that it is easy to incorporate a noise (degradation) factor into our model. Having shown the plausibility of the approach, we plan to test it on synthetic and real data sets with different noise levels.

References

1. Wedeen, V.J., Hagmann, P., Tseng, W.Y.I., Reese, T.G., Weisskoff, R.M.: Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magnetic Resonance in Medicine **54**(6) (2005) 1377–1386
2. Tuch, D.S.: Q-ball imaging. Magnetic Resonance in Medicine **52**(6) (2004) 1358–1372
3. Deriche, R., Calder, J., Descoteaux, M.: Optimal real-time q-ball imaging using regularized Kalman filtering with incremental orientation sets. Medical Image Analysis **13**(4) (2009) 564 – 579
4. Merlet, S., Deriche, R.: Compressed sensing for accelerated EAP recovery in diffusion MRI. In: MICCAI (1). (2010)
5. Michailovich, O.V., Rathi, Y.: Fast and accurate reconstruction of HARDI data using compressed sensing. In: MICCAI (1). (2010) 607–614
6. Alexander, D.C.: Multiple-fibre reconstruction algorithms for diffusion MRI. Annals of the New York Academy of Sciences **1064** (December 2005) 113–133
7. Peled, S., Friman, O., Jolesz, F., Westin, C.F.: Geometrically constrained two-tensor model for crossing tracts in DWI. Magnetic Resonance Imaging **24** (2006) 1263–1270
8. Geman, S., Graffigne, C.: Markov random field image models and their applications to computer vision. In: Proceedings of the International Congress of Mathematicians (1986)
9. Stejskal, E.O.: Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow. The Journal of Chemical Physics **43**(10) (1965) 3597–3603

10. Basser, P.J., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. J Magn Reson B **103**(3) (March 1994) 247–254
11. Alexander, A.L., Hasan, K.M., Lazar, M., Tsuruda, J.S., Parker, D.L.: Analysis of partial volume effects in diffusion-tensor MRI. Magnetic Resonance in Medicine **45**(5) (2001) 770–780
12. Hammersley, J.M., Clifford, P.: Markov fields of finite graphs and lattices. Technical report, University of California-Berkeley, preprint (1968)
13. Poupon, C., Clark, C.A., Frouin, V., Rgis, J., Bloch, I., Bihan, D.L., Mangin, J.F.: Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. NeuroImage **12**(2) (2000) 184–195
14. Poupon, C., Mangin, J.F., Clark, C., Frouin, V., Rgis, J., Bihan, D.L., Bloch, I.: Towards inference of human brain connectivity from MR diffusion tensor data. Medical Image Analysis **5**(1) (2001) 1–15
15. Szeliski, R., Zabih, R., Scharstein, D., Veksler, O., Kolmogorov, V., Agarwala, A., Tappen, M., Rother, C.: A comparative study of energy minimization methods for Markov random fields with smoothness-based priors. IEEE Trans. Pattern Anal. Mach. Intell. **30** (June 2008) 1068–1080
16. Barmpoutis, A., Jian, B., Vemuri, B.C.: Adaptive kernels for multi-fiber reconstrucion. In: Proceedings of the 21st International Conference on Information Processing in Medical Imaging. IPMI '09, Berlin, Heidelberg, Springer-Verlag (2009) 338–349