INTERPOLATING MULTI-FIBER MODELS BY GAUSSIAN MIXTURE SIMPLIFICATION

Maxime Taquet, Benoît Scherrer, Christopher Benjamin, Sanjay Prabhu, Benoît Macq, Simon K. Warfield

Computational Radiology Laboratory, Harvard Medical School, Boston, USA
ICTEAM Institute, Université catholique de Louvain, Belgium

ABSTRACT

Multi-fiber models have been introduced to leverage the accuracy of the diffusion representation in crossing fiber areas. The improved accuracy may, however, be impaired by poor processing of the multi-fiber models. In particular, interpolating multi-fiber models proves challenging, while it is a pervasive and recurrent task in many processes. The error accumulated from iterating a poor interpolation may yield significantly corrupted global results. In this paper, we propose an interpolation scheme based on gaussian mixture simplification and demonstrate its benefits over a heuristic approach in terms of spatial normalization and tractography results.

Index Terms— Multi-Fiber Models, Interpolation, Tractography, Spatial Normalization, Diffusion Tensor Imaging

1. INTRODUCTION

Brain diffusion tensor imaging (DTI) enables the visualization and characterization of fiber tracts in the white matter. A classical limitation of DTI is its incapacity to represent complex structures such as crossing fibers [1]. To overcome this limitation, novel model-based and model-free methods to analyze the diffusion signals have emerged (see [1], chapter 2 for a detailed review). Multi-fiber models (MFM) are of particular attractiveness since they allow the computation of diffusion parameters for each fiber bundle independently. This property is of central interest for tractography and fiber integrity assessment [2]. MFM represent the diffusion as a gaussian mixture model:

$$S(b, \boldsymbol{g}) = S_0 \left(f_0 e^{-b \boldsymbol{D}_{\text{iso}}} + \sum_{i=1}^{N} f_i e^{-b \boldsymbol{g}^T \boldsymbol{D}_i \boldsymbol{g}} \right), \quad (1)$$

where b is the b-value at which the signal is acquired, $\boldsymbol{g}$ is the gradient direction, D_{iso} is the diffusivity of free water, $\boldsymbol{D}_i$ are the anisotropic diffusion tensors and f_i are the relative volumetric occupancy. In this model, water molecules are assumed to be in one of the compartment with some probability f_i. These models can then be used in various applications, including tractography [3, 4].

However, the structure of MFM makes it challenging to perform basic tasks such as interpolation. The difficulty mainly comes from the arbitrary assignment of a compartment to the tensors (*e.g.* for a two-tensor model, the parameterization (f_1, D_1, f_2, D_2) is equivalent to (f_2, D_2, f_1, D_1)). A simple generalization of single tensor interpolations is therefore illicit. For this reason, authors have performed the interpolation on the raw diffusion-weighted sequences [4] or they have defined some heuristics to cluster the tensors into groups on which single tensor interpolation can be performed [3]. One such heuristics would be to define the clusters based on the principal eigenvector of each component. This heuristic method may lead to ill-posed selection and arbitrary choices when different voxels contain different number of tensors.

In this paper, we present an interpolation method which fully accounts for the structure of multi-fiber models, based on recent developments in gaussian mixture simplification. The remaining of this paper is organized as follows. Section 2 introduces the theory of gaussian mixture simplification and applies it to multi-fiber models. Section 3 then presents the results in terms of tractography and spatial normalization. Finally, Section 4 concludes the paper.

2. GAUSSIAN MIXTURE SIMPLIFICATION

The central idea of the proposed method is to define a complete gaussian mixture containing the information of all the initial mixtures and to subsequently reduce this model to a mixture of the same order as the initial data. This section introduces the theory of gaussian mixture simplification and then applies it to multi-fiber model simplification.

2.1. General Theory

Gaussian mixture simplification (GMS) has been developed in distribution-based soft clustering where a mixture models needs be learnt from a set of distributions. A seminal paper in this field is [5], which has been extended in [6] for the particular case of gaussian mixtures.

Let $G_i(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_i, \Sigma_i)$ be a multivariate gaussian and

$$p_N(\boldsymbol{x}|\boldsymbol{G}) = \sum_{i=1}^{N} \alpha_i G_i(\boldsymbol{x})$$

be a gaussian mixture model of N components. Simplifying this gaussian mixture consists in defining a new gaussian mixture of order $K \leq N$:

$$p_K(\boldsymbol{x}|\boldsymbol{R}) = \sum_{j=1}^{K} f_j R_j(\boldsymbol{x}),$$

where $R_j(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{m}_j, \boldsymbol{T}_j)$ is another set of multivariate gaussians. This simplification is based on the minimization of some energy function:

$$p_K^*(\boldsymbol{x}|\boldsymbol{R}) = \underset{p_K(\boldsymbol{x}|\boldsymbol{R})}{\arg\min}\, D(p_N(\boldsymbol{x}|\boldsymbol{G}), p_K(\boldsymbol{x}|\boldsymbol{R})).$$

In information theoretic approaches, the energy function is related to the amount of information lost when approximating p_N by p_K. In [6], the cumulative differential relative entropy is used, that is the weighted sum of the differential relative entropies between the gaussian components in p_N and their best representative component in p_K. In other words,

$$D(p_N(\boldsymbol{x}|\boldsymbol{G}), p_K(\boldsymbol{x}|\boldsymbol{R})) = \sum_{j=1}^{K} \sum_{i:\pi_i=j} \alpha_i D(G_i||R_j), \quad (2)$$

where the π_i's are latent variables indicating which component R_j best represents G_i. An expectation-maximization scheme can efficiently minimize (2). In a first step, the N components are assigned to one of K clusters and, in a second step, the parameters of the representatives R_j are optimized in each cluster.

Davis and Dhillon [6] showed that the optimization of the parameters of the R_j's can be carried out in closed form once the cluster assignments are known. We have:

$$\boldsymbol{m}_j = \frac{\sum_{i:\pi_i=j} \alpha_i \mu_i}{\sum_{i:\pi_i=j} \alpha_i}, \quad (3)$$

$$T_j = \frac{\sum_{i:\pi_i=j} \alpha_i \left(\Sigma_i + (\boldsymbol{\mu}_i - \boldsymbol{m}_j)(\boldsymbol{\mu}_i - \boldsymbol{m}_j)^T\right)}{\sum_{i:\pi_i=j} \alpha_i}. \quad (4)$$

In the clustering step, G_i is assigned the cluster j such that R_j is the closest to G_i in terms of differential entropy. Alternating these two steps converges to a gaussian mixture model of K elements that minimizes (locally) the cumulative differential entropy. The mixture weights f_j are naturally equal to the sum of the initial weights α_i in the clusters: $f_i = \sum_{i:\pi_i=j} \alpha_i$.

2.2. Multi-Fiber Model Simplification

The problem of interpolating between MFM can be interpreted as a particular case of GMS by defining and simplifying a complete mixture comprising all the neighboring components. Let M be the number of neighbors from which we want to interpolate the mixture at a floating location $\boldsymbol{x}$ (typically, $M = 8$) and let $\mathcal{M}_m$ be the gaussian mixture at the neighbor voxel m ($1 \leq m \leq M$):

$$\mathcal{M}_m = \sum_{i=1}^{N_m} \alpha_{m,i} \mathcal{N}(\boldsymbol{g}|\boldsymbol{0}, D_{m,i}^{-1}).$$

Each component of this mixture has a zero mean and a covariance matrix equal to the inverse of the diffusion tensor $D_{m,i}$. The number of components N_k may vary from one voxel to the other. The *complete* mixture model is then:

$$p_N(\boldsymbol{g}|\mathcal{M}) = \sum_{m=1}^{M} w_m \sum_{i=1}^{N_m} \alpha_{m,i} \mathcal{N}(\boldsymbol{g}|\boldsymbol{0}, D_{m,i}^{-1}).$$

with weights w_m defined by classical scalar interpolation (*e.g.* trilinear). The number of components in the final mixture at location $\boldsymbol{x}$ depends on our choice of model. In this paper, we choose $K = \max_m N_m$. GMS is used to reduce the complete model with $\sum_{m=1}^{M} N_m$ components to a model with K components. Since tensors in DTI have zero mean, the update rules (3) and (4) are simplified. The former can be ignored and the latter amounts to computing the weighted average of the initial covariances in each cluster. As for the cluster assignment, it is carried out by minimizing the Burg matrix divergence between the covariance matrices, *i.e.*

$$\pi_i = \underset{j}{\arg\min}\, D(G_i||R_j) = \underset{j}{\arg\min}\, B(D_i^{-1}, T_j)$$

where the Burg matrix divergence for 3×3 matrices is defined as $B(A, B) = \mathrm{Tr}(AB^{-1}) - \log|AB^{-1}| - 3$.

One may be concerned about the swelling effect stemming from averaging the covariance matrices in (4). Although this effect reflects the intrinsic lack of knowledge about the diffusion signal at the considered location, it can be undesired, depending on the application [7]. This entails us to define a log-euclidean version of the GMS described above as it was defined for single-tensor interpolation [8]. This is achieved by replacing the covariance matrices by their matrix logarithm prior to performing GMS. The update of the covariance matrices now reads:

$$\log T_j = \frac{\sum_{i:\pi_i=j} \alpha_i \log D_i^{-1}}{\sum_{i:\pi_i=j} \alpha_i}.$$

Interestingly, since $\log A^{-1} = -\log A$, the multi-fiber interpolation in the log-domain, reduces to the single tensor interpolation in areas where a single component is present in each neighboring voxel ($N_k = 1$) and $K = 1$. This is not the case in the direct domain, since GMS averages the covariance matrices rather than the tensors.

The EM optimization in GMS converges to a local minimum of the cumulative differential entropy. An initialization step is thus required. In this paper, we initialize the clusters by spectral clustering [9] with the cosine similarity between the primary eigenvectors of each tensor as a similarity matrix.

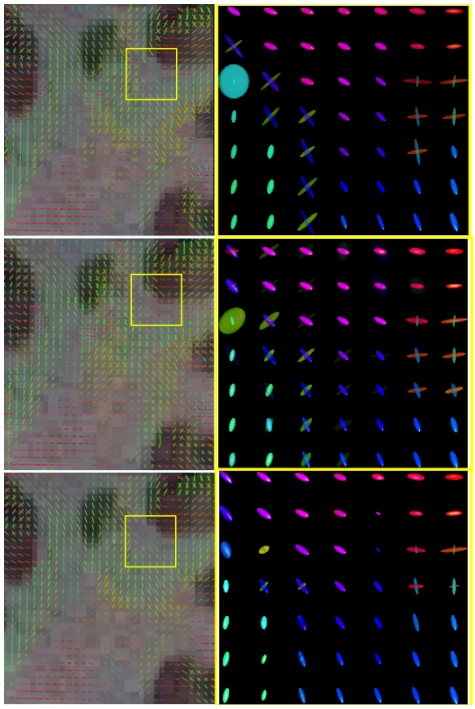

Fig. 1. GMS interpolation better preserves the information contained in an original multi-fiber model. The results here compares the original image (top) with the result projected backward using GMS (middle) and heuristic interpolation (bottom). A zoom of the yellow box shows how GMS better preserves areas with multiple fibers. In the zoomed version, fractions are encoded as the transparency of the ellipsoids.

3. APPLICATIONS

In this section, we present the results of the GMS interpolation compared to the heuristic interpolation presented in the introduction. The interpolation was applied in two contexts: spatial normalization and two-tensors tractography. Ten multi-fiber DTI (resolution: $1.8\times1.8\times2.4\text{mm}^3$) of different subjects with two anisotropic and one isotropic compartments were estimated using the method described in [2]. The log-euclidean version of GMS was used and compared with the log-euclidean heuristic interpolation.

3.1. Spatial Normalization

To assess the quality of the interpolation in terms of spatial normalization, we sequentially apply a spatial transform and its inverse to the data and compare the final image with the initial one. This tells us the amount of information lost in the process. The transform used here is obtained by registering a T1-weighted image of each subject to an atlas. The results show that GMS interpolation better preserves the information contained in the original image (Fig. 1).

By selecting the tensor beforehand and subsequently interpolating single tensor fields, the heuristic interpolation tends to lose information in areas where both single tensor and multiple-tensors are present, as shown by the yellow box in Fig. 1. In contrast, by evaluating a global model at once, GMS preserves as much information as possible.

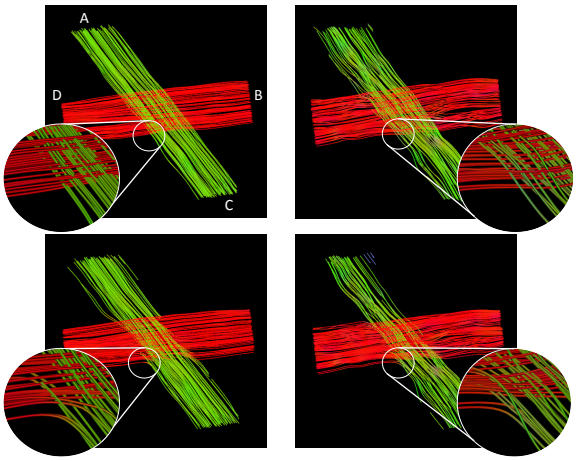

Fig. 2. GMS (top) prevents confounding the fiber bundles in tractography, as compared with heuristic interpolation (bottom). Results are shown for 5% (left) and 15% (right) gaussian noise.

To assess the gain in accuracy obtained by GMS interpolation, a voxel-wise distance between the original data, (f_0, F_0, f_1, F_1) and the backward projected result, (g_0, G_0, g_1, G_1) was computed for all ten multi-fiber brain DTI as:

$$\min\Big(\frac{f_0+g_0}{2}||F_0-G_0||^2+\frac{f_1+g_1}{2}||F_1-G_1||^2,$$
$$\frac{f_0+g_1}{2}||F_0-G_1||^2+\frac{f_1+g_0}{2}||F_1-G_0||^2\Big), \quad (5)$$

The mean distance was 27.2%($\pm$.05%) of the mean tensor norm for GMS and 29.4%($\pm$.1%) for the heuristic interpolation. When restricted to areas with multiple tensors, the difference is larger with a mean distance of 22.3%($\pm$.2%) for GMS and 38.0%($\pm$.4%) for the heuristic interpolation. A one-tail t-test performed on the voxel-wise difference between the GMS interpolation error and the heuristic interpolation error showed that, on average, GMS performs significantly better ($p < 10^{-6}$ for each subject individually).

3.2. Tractography

In the context of tractography, interpolation is required to estimate the value of the MFM at the floating end of the tract being constructed. We applied and compared both interpolation schemes in a probabilistic tractography algorithm [3].

3.2.1. Synthetic Data

Tractography was performed on a digital phantom made of two crossing fiber bundles (Fig. 2) under the influence of

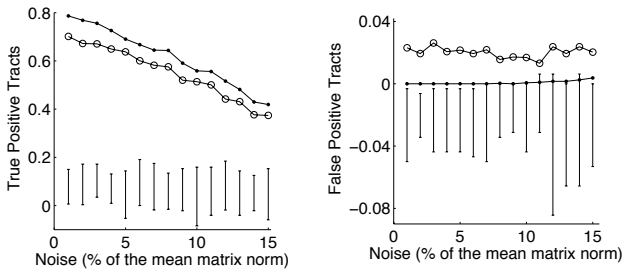

Fig. 3. GMS (•) significantly increases the number of tracts correctly identified (true positive rate, left) and reduces the proportion of confounded tracts (false positive rates, right) as compared to heuristic interpolation (○). The plot represents the mean rate for each noise level. A 95% confidence interval of the difference in these statistics between GMS and heuristic interpolation is depicted by vertical bars.

noise. Symmetric matrices of gaussian noise (20 repetitions of each noise level between 1%-15% of the mean Frobenius tensor norm) were added independently to the MFM compartments in the log-domain. The results reveal more confounds in the fiber bundles when heuristic interpolation is used than when GMS is used (Fig. 2). In terms of the areas they connect (A,B,C,D, see Fig. 2), the true positive tract rates (number of correctly identified tracts over number of true tracts in the phantom) and the false positive rates (number of confounded tracts over number of true tracts) were computed. On average, for any noise level, GMS significantly increases the true positive rates ($p < .03$) and decreases the false positive rates ($p < .001$), as seen on Fig. 3.

3.2.2. Clinical Data

The performance of the interpolation schemes were compared in the context of a clinical application: the visualization of the optic radiation, a set of axons carrying visual stimuli to the visual cortex. These neural pathways present areas of crossing fibers whose disentanglement is critical to visualize particular structures such as the Meyer's loop. Performing probabilistic tractography with both interpolation schemes demonstrates that GMS is better at unravelling tracts (Fig. 4).

4. CONCLUSION

This paper has introduced a novel approach to the interpolation of multi-fiber models. Experiments on synthetic and real world data demonstrate the benefits of this approach over a more heuristic method. In particular, spatial normalization presents a lower information loss and tractography reveals more subtle structures and yields fewer spurious tracts. We believe that gaussian mixture simplification (GMS) should be used every time interpolating between multi-fiber models is required, since small interpolation errors tend to accumulate to corrupt the global results in practical contexts.

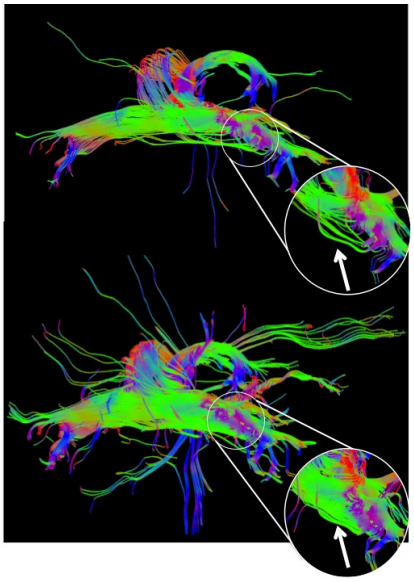

Fig. 4. GMS (top) better unravels neural pathways in the presence of crossing fibers than heuristic interpolation (bottom). The zoomed-in area is the Meyer's loop. The loop is not clearly visible at the bottom, due to the poor interpolation in crossing fiber areas. Note also the increased number of spurious tracts on the bottom image.

5. REFERENCES

[1] H. Johansen-Berg and T.E.J. Behrens, *Diffusion MRI: from quantitative measurement to in-vivo neuroanatomy*, Academic Press, 2009.

[2] B. Scherrer and S.K. Warfield, "Toward an accurate multi-fiber assessment strategy for clinical practice," in *IEEE ISBI 2011*. IEEE, 2011, pp. 2140–2143.

[3] TEJ Behrens, H.J. Berg, S. Jbabdi, MFS Rushworth, and MW Woolrich, "Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?," *Neuroimage*, vol. 34, no. 1, pp. 144–155, 2007.

[4] O. Bergmann, G. Kindlmann, S. Peled, and C.F. Westin, "Two-tensor fiber tractography," in *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*. IEEE, 2007, pp. 796–799.

[5] A. Banerjee, S. Merugu, I.S. Dhillon, and J. Ghosh, "Clustering with bregman divergences," *The Journal of Machine Learning Research*, vol. 6, pp. 1705–1749, 2005.

[6] J.V.D.I. Dhillon, "Differential entropic clustering of multivariate gaussians," in *Advances in Neural Information Processing Systems*. The MIT Press, 2007, vol. 19, p. 337.

[7] O. Pasternak, N. Sochen, and P.J. Basser, "The effect of metric selection on the analysis of diffusion tensor mri data," *NeuroImage*, vol. 49, no. 3, pp. 2190–2204, 2010.

[8] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-euclidean metrics for fast and simple calculus on diffusion tensors," *MRM*, vol. 56, no. 2, pp. 411–421, 2006.

[9] J. Shi and J. Malik, "Normalized cuts and image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 8, pp. 888–905, 2000.