UNSUPERVISED AUTOMATIC WHITE MATTER FIBER CLUSTERING USING A GAUSSIAN MIXTURE MODEL

Meizhu Liu[1], *Baba C. Vemuri*[1] *and Rachid Deriche*[2]

[1]Department of CISE, University of Florida, Gainesville, FL, 32611, USA
[2]Project-Team ATHENA, INRIA, Sophia Antipolis - Méditerranée, 06902, France

ABSTRACT

Fiber tracking from diffusion tensor images is an essential step in numerous clinical applications. There is a growing demand for an accurate and efficient framework to perform quantitative analysis of white matter fiber bundles. In this paper, we propose a robust framework for fiber clustering. This framework is composed of two parts: accessible fiber representation, and a statistically robust divergence measure for comparing fibers. Each fiber is represented using a Gaussian mixture model (GMM), which is the linear combination of Gaussian distributions. The dissimilarity between two fibers is measured using the total square loss function between their corresponding GMMs (which is statistically robust). Finally, we perform the hierarchical total Bregman soft clustering algorithm on the GMMs, yielding clustered fiber bundles. Further, our method is able to determine the number of clusters automatically. We present experimental results depicting favorable performance of our method on both synthetic and real data examples.

Index Terms— Fiber clustering, Gaussian mixture model, Gaussian process, total Bregman divergence, total square loss.

1. INTRODUCTION

Diffusion magnetic resonance imaging (dMRI) is a very popular technique that has been widely studied and reported in literature. It uses diffusion sensitizing gradients to non-invasively capture the anisotropic properties of tissue. Diffusion tensor imaging (DTI) approximates the diffusivity function by a symmetric positive definite tensor of order two [1]. DTI allows one to capture the diffusion direction of water molecules in the brain, associated with the direction of fiber tracts in the white matter. When this movement is hindered by membranes and macromolecules, water diffusion becomes anisotropic. Therefore, in highly structured tissues such as nerve fibers, this anisotropy can be used to characterize local tissue structure. Fiber tractography and fiber clustering are promising for many applications such as analyzing brain tumors, migraines, eclampsia and other disorders. All of these latter tasks will benefit from an accurate and robust fiber analysis technique.

However, accurate and efficient visualization of this large and complicated 3D fiber tract data set to gain clinical insight is extremely difficult. To tackle this problem, many fiber clustering methods have been proposed recently, e.g., the tracts have been successfully clustered into bundles using spectral clustering [2, 3], normalized cuts [4], agglomerative hierarchical clustering [5], fuzzy c-means [6], k-nearest neighbors [7]. For fiber representation, typical methods represent a fiber using a set of points [8, 9, 7], shapes [10, 11, 12], splines [13], geometric characteristics of fibers (e.g. length, center of mass, and second order moment) [14] and so on. For fiber comparison, the difference among fibers is quantified using different metrics including closest point distance, chamfer distance, Hausdorff distance, and their symmetric versions [15, 11, 9, 13, 16, 8].

In this paper, we propose a novel method for fiber clustering. The novelty lies in three parts. First, we present an easy to access fiber representation method. Second, we present a robust and efficient divergence measure to measure the dissimilarity between fibers. Finally, we present a hierarchical clustering framework which is not only robust, but also can automatically determine the number of clusters. To be specific, we use Gaussian mixture model (GMM) to represent each fiber. This GMM is calculated based on the points lying on the fiber. Note that earlier work in literature [17] reported a Gaussian process representation but not a GMM representation of the fibers. After getting the GMM for each fiber, the dissimilarity between two fibers is computed using the total square loss (tSL) between their corresponding GMMs. tSL belongs to the class of total Bregman divergences (tBD), which was proposed recently [18]. tBD has been proved to be a statistically robust divergence measure. This divergence measure is based on the orthogonal distance between the convex generating function of the divergence and its tangent approximation at the second argument of the divergence. For more details, we refer readers to [18]. Finally, the tBD hierar-

This research was in part funded by the NIH grant NS066340 to Baba C. Vemuri, the University of Florida Alumni Fellowship and the INRIA Internship to Meizhu Liu, and the Association France Parkinson and the French National Research Agency (ANR, Neurodegenerative and Psychiatric Diseases) to Rachid Deriche.

chical soft clustering algorithm is performed on GMMs. The clustering yields a grouping of the fiber bundles. The obvious benefit of using this clustering algorithm is that it is very efficient because the cluster center for a set of GMMs is in closed form. Furthermore, the number of clusters can be automatically detected. The entire framework is explained at length in Section 2.

The rest of the paper is organized as follows. In Section 2, we introduce our proposed method, followed by the empirical validation in Section 3. Finally we conclude in Section 4.

2. PROPOSED METHOD

In this part, we will first introduce the GMM representation method for fibers, and then we will present the tBD soft clustering algorithm for fiber clustering. Finally, we describe how our method can automatically determine the number of clusters.

2.1. Fiber representation

Suppose the set of points on a fiber $\mathcal{F}$ are denoted by $\{\mathbf{f}_i\}_{i=1}^{|\mathcal{F}|}$, where $|\mathcal{F}|$ denotes the number of points on the fiber $\mathcal{F}$. A fiber $\mathcal{F}$ is characterized by a blurred indicator function $I_{\mathcal{F}}(\mathbf{f})$, which is modeled by a Gaussian process as in [17], and

$$I_{\mathcal{F}}(\mathbf{f}) \sim \mathcal{N}(I^*_{\mathcal{F}}(\mathbf{f}), \sigma_{\mathcal{F}}(\mathbf{f})). \tag{1}$$

The maximal level-set of this indicator function corresponds to the fiber $\mathcal{F}$ [17]. In this paper, we represent the fiber $\mathcal{F}$ as a GMM, which is consistent with the above definitions,

$$p_{\mathcal{F}} = \frac{1}{|\mathcal{F}|}\sum_{i=1}^{|\mathcal{F}|}\mathcal{N}(I^*_{\mathcal{F}}(\mathbf{f}_i), \sigma_{\mathcal{F}}(\mathbf{f}_i)). \tag{2}$$

The dissimilarity between two fibers $\mathcal{F}_1$ and $\mathcal{F}_2$ is computed using the total square loss (tSL) defined in [18] and reproduced below for convenience. Given two GMMs p_1 and p_2,

$$p_1(x) = \sum_{i=1}^{m} a_i^{(1)}\mathcal{N}(x; \mu_i^{(1)}, \Sigma_i^{(1)}), \tag{3}$$

$$p_2(x) = \sum_{i=1}^{m} a_i^{(2)}\mathcal{N}(x; \mu_i^{(2)}, \Sigma_i^{(2)}), \tag{4}$$

where $x \in \mathbf{R}^n$, m is the number of Gaussian components, and a_i is the coefficient for the ith component. The tSL between p_1 and p_2 is

$$d(p_1, p_2) = \frac{\int (p_1 - p_2)^2 dx}{\sqrt{1 + \int 4p_2^2 dx}} = \frac{d_1 + d_2 - d_{1,2}}{\sqrt{1 + 4d_2}}, \tag{5}$$

where

$$d_1 = \sum_{i,j=1}^{m} a_i^{(1)} a_j^{(1)} \mathcal{N}(0; \mu_i^{(1)} - \mu_j^{(1)}, \Sigma_i^{(1)} + \Sigma_j^{(1)}), \tag{6}$$

$$d_2 = \sum_{i,j=1}^{m} a_i^{(2)} a_j^{(2)} \mathcal{N}(0; \mu_i^{(2)} - \mu_j^{(2)}, \Sigma_i^{(2)} + \Sigma_j^{(2)}), \tag{7}$$

$$d_{1,2} = 2\sum_{i,j=1}^{m} a_i^{(1)} a_j^{(2)} \mathcal{N}(0; \mu_i^{(1)} - \mu_j^{(2)}, \Sigma_i^{(1)} + \Sigma_j^{(2)}), \tag{8}$$

$$\mathcal{N}(0; \mu, \Sigma) = \frac{1}{(\sqrt{2\pi})^n \sqrt{\det(\Sigma)}} \exp\left(-\frac{\mu' \Sigma^{-1} \mu}{2}\right). \tag{9}$$

The tSL between two GMMs only depends on the means and covariance matrices of the Gaussian components, and is very efficient to compute.

Given a set of fibers and their corresponding GMMs, their center (average or representative) is defined through the following minimization,

$$\bar{p} = \arg\min_{p} \sum_{l=1}^{n} d(p, p_l). \tag{10}$$

The solution to this minimization can be written in closed form as,

$$\bar{p} = \frac{\sum_{l=1}^{n} w_l p_l}{\sum_{l=1}^{n} w_l}, \quad w_l = (1 + 4d_l)^{-1/2}. \tag{11}$$

This closed form solution leads to fast computations in our clustering algorithm described below.

2.2. Hierarchical soft clustering

A hierarchical divisive clustering process is performed using the tBD soft clustering algorithm proposed in [19], which is presented here in the Algorithm block 1 for readers' convenience.

Let l denote the level of clustering. The original (entire) dataset is at level $l = 0$. We first split the whole dataset into two clusters, which are at level $l = 1$. Then from the level $l = 1$ clusters, we choose the cluster with largest maximum divergence from the cluster center, and split this cluster into two clusters. These two clusters along with the level $l = 1$ clusters, except the split cluster, will form the clusters at the $l = l + 1$ level. The above process is repeated on the new l level clusters.

2.3. Choosing the Number of Clusters

We define the effort "e" of splitting a cluster $\mathbf{C}$ into $\mathbf{A}$ and $\mathbf{B}$ as

$$e(\mathbf{C}) = \frac{2}{|\mathbf{C}|}\sum_{i\in\mathbf{C}} d(\bar{p}_{\mathbf{C}}, p_i) - \frac{1}{|\mathbf{A}|}\sum_{i\in\mathbf{A}} d(\bar{p}_{\mathbf{A}}, p_i) - \frac{1}{|\mathbf{B}|}\sum_{i\in\mathbf{B}} d(\bar{p}_{\mathbf{B}}, p_i), \tag{12}$$

Algorithm 1 Total Bregman Soft Clustering Algorithm

Input: The data to be clustered $X = \{x_i\}_{i=1}^N$, and the number of clusters c.
Output: $M = \{m_j\}_{j=1}$ and $Q = \{q_j\}_{j=1}$, m_j is the cluster center for the j^{th} cluster with probability q_j.
Initialization: Randomly choose c elements from X as M and set Q to the uniform probability.
repeat
  {assign x_i to clusters}
  for $i = 1$ to N **do**
    for $j = 1$ to c **do**
      $q(j \mid x_i) \leftarrow \frac{q_j \exp(-d(m_j, x_i))}{\sum_{j'=1}^{c} q_{j'} \exp(-d(m_{j'}, x_i))}$
    end for
  end for
  {update cluster centers}
  for $j = 1$ to c **do**
    $q_j \leftarrow \frac{1}{N} \sum_{i=1}^{N} q(j \mid x_i)$
    $m_j \leftarrow t$-center for cluster j
  end for
until The change of the results between two consecutive iterations is below some threshold.

where $\bar{p}_{\mathbf{C}}$, $\bar{p}_{\mathbf{A}}$ and $\bar{p}_{\mathbf{B}}$ are the centers of the sets $\mathbf{C}$, $\mathbf{A}$ and $\mathbf{B}$ as defined in (11). $|\mathbf{A}|$ represents the number of fibers in $\mathbf{A}$. Note that the higher the value of this effort, the better the splitting, since the effort is the amount of reduction in the average dissimilarity between the center and the members of the set in the original cluster $\mathbf{C}$.

The number of clusters is determined in the following way. First, we perform hierarchical clustering for a known number of times, K. Compute the effort $\{e_i\}_{i=1}^K$. Then for each $k \in \{2, \cdots, K\}$, we use total least squares fitting to fit a linear function to the computed efforts $\{(i, e_i)\}_{i=1}^k$ and $\{(i, e_i)\}_{i=k+1}^K$ respectively. Let's denote the total fitting error as ϵ_k. Then the number of optimal clusters will be $k^* = \arg\min_k \epsilon_k, k = 1, \cdots, K$. This is illustrated in Fig. 1. In this paper, we stop the splitting process when the number of clusters reaches $\sqrt{N}$, i.e., $K = N$. This number was chosen heuristically.

3. EXPERIMENTAL RESULTS

We evaluated our method on both synthetic and real datasets. We first generate a $160 \times 160 \times 160$ diffusion tensor field and the principle eigenvectors of the DTI are shown in Fig. 2(a). This data was generated such that the eigenvectors of the diffusion tensors group into 5 different fiber clusters which provides the ground truth for our clustering method. The tractography was performed using Slicer, and the streamline tractography was implemented by seeding in sub-voxel resolution by taking every voxel with fractional anisotropy higher than 0.1, with the stopping track curvature 0.8, and integration step length 0.5mm. The minimum length is 10mm. In this way, we got 400 fibers which are shown in Fig. 2(b). There are about 50 points on each fiber. The clustering results are shown in Fig. 2(d) and the final cluster centers are shown in Fig. 2(c) in bold. The clustering results yield precisely 5 clusters almost the same as the ground truth, thus validating the accuracy of our proposed clustering algorithm.

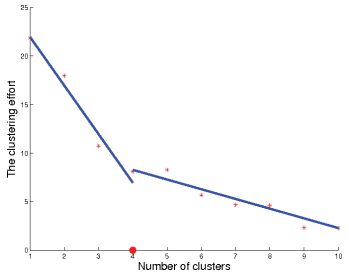


Fig. 1. Determine the number of clusters according to total least squares fitting. The x-axis represents the number of clusters, while the y-axis represents the clustering effort "e". The two blues lines are fitting lines. The red circle represents the determined optimal number of clusters.

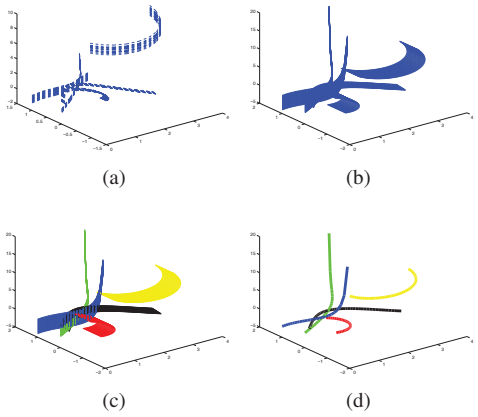


Fig. 2. Clustering on synthetic data. The principle eigenvectors are shown in (a), and the fibers are shown in (b). The final clustering fiber bundles along with the final centers are shown in Fig. 2(c)(d).

To evaluate our method on the real dataset, we used the human brain DWI dataset ($93 \times 116 \times 93$) provided by Alfred Anwander of the Max Planck Institute for Human Neuroscience [20]. The DWIs were acquired with a whole-body 3 Tesla Magnetom TRIO operating at 3T (Siemens Medical Solution) equipped with an 8-channel head array coil. The twice-refocused spin-echo EPI sequence (TR = 12 s, TE = 100 ms) consists of 60 diffusion gradients with a b-value of $1000 s/mm^2$. The DTI image was computed using non-

local means variational framework [21]. The tractography was implemented by seeding in sub-voxel resolution by taking every voxel with fractional anisotropy higher than 0.1, with the stopping track curvature 0.8, integration step length 0.5mm. The minimum length is 10mm, and maximum length is 800mm. In total, 611 fiber tracts were obtained. The fibers are shown in Fig. 3(a). The initial centers are shown in Fig. 3(b). The final clustering fiber bundles along with the final centers are shown in Fig. 3(c)(d).

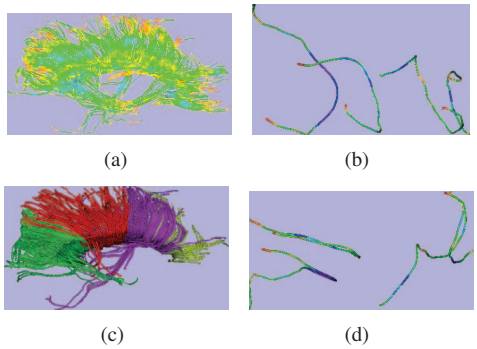


Fig. 3. The fibers from the tractography process are shown in Fig. 3(a). The initial centers are shown in Fig. 3(b). The final clustering fiber bundles along with the final centers are shown in Fig. 3(c)(d).

4. CONCLUSIONS

In this paper, we proposed an unsupervised fiber clustering framework that can automatically determine the number of clusters. We use a Gaussian mixture model (GMM) to represent a fiber, and the dissimilarity between two fibers is defined using the total square loss between the corresponding GMMs. We use total Bregman soft clustering algorithm to cluster the fibers into different clusters, and each cluster corresponds to a fiber bundle. The number of clusters is determined according to the total least squares fitting error in the clustering effort defined earlier in (12). The proposed method has been evaluated favorably on both synthetic and real datasets.

References

[1] P.J. Basser, J. Mattiello, and D. LeBihan, "Estimation of the effective self-diffusion tensor from the NMR spin echo," *J. Magn. Res.*, vol. 103, no. 3, pp. 247–254, 1994.

[2] L. Jonasson, P. Hagmann, J.P. Thiran, and V.J. Wedeen, "Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering," *Proc. ISMRM*, p. 1310, 2005.

[3] L. O'Donnell and C.F. Westin, "White matter tract clustering and correspondence in populations," *MICCAI*, pp. 3570–3577, 2005.

[4] A. Brun, H. Knutsson, H.J. Park, M.E. Shenton, and C.F. Westin, "Clustering fiber traces using normalized cuts," *MICCAI*, pp. 368–375, 2004.

[5] S. Zhang and D.H. Laidlaw, "DTI fiber clustering in the whole brain," *IEEE Visualization*, p. 28p, 2004.

[6] J.S. Shimony and A.Z. Snyder abd N. Lori and T.E. Conturo, "Automated fuzzy clustering of neur-sonal pathways in diffusion tensor tracking," *Proc. ISMRM*, p. 10, 2002.

[7] Z. Ding, J.C. Gore, and A.W. Anderson, "Classification and quantification of neuronal fiber pathways using diffusion tensor MRI," *MRM*, vol. 49, pp. 716–721, 2003.

[8] M. Maddah, W. E. L. Grimson, S. K. Warfield, and W. M. Wells, "A unified framework for clustering and quantitative analysis of white matter fiber tracts," *MedIA*, vol. 12, pp. 191–202, 2008.

[9] L. O'Donnell and C.F. Westin, "High-dimensional white matter atlas generation and group analysis," *MICCAI*, vol. 4191, pp. 243–251, 2006.

[10] P.G. Calamante F. Tournier J.-D. Atkinson D. Hill D.L.G. Connelly A. Batchelor, "Quantification of the shape of fiber tracts," *MRM*, vol. 55, no. 4, pp. 894–903, 2006.

[11] I. Corouge, P.T. Fletcherb, S. Joshic, S. Gouttardd, and G. Geriga, "Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis," *MedIA*, pp. 786–798, 2005.

[12] A. JunLeemans, J. Sijbers, S. D. Backer, E. Vandervliet, and P. Parizel, "Multiscale white matter fiber tract coregistration: a new feature-based approach to align diffusion tensor data," *MRM*, vol. 55, no. 6, pp. 1414–1423, 2006.

[13] J. Glaunes, A. Qiu, M. Miller, and L. Younes, "Large deformation diffeomorphic metric curve mapping," *International Journal of Computer Vision*, vol. 80, no. 3, pp. 317–336, 2008.

[14] I. Corouge, S. Gouttard, and G. Gerig, "Towards a shape model of white matter fiber bundles using diffusion tensor mri," *Int. Symp. Biomed. Imag.*, pp. 344–347, 2004.

[15] A. Brun, H. Park, H. Knutsson, and C.F. Westin, "Coloring of DT-MRI fiber traces using laplacian eigenmaps," *Computer Aided Systems Theory - EUROCAST*, vol. 2809, pp. 518–529, 2003.

[16] A. Tsai, C.F. Westin, A.O. Hero, and A.S. Willsky, "Fiber Tract Clustering on Manifolds With Dual Rooted-Graphs," *IEEE CVPR*, vol. 51, pp. 1–6, 2007.

[17] D. Wassermann, L. Bloy, E. Kanterakis, R. Verma, and R. Deriche, "Unsupervised white matter fiber clustering and tract probability map generation: Applications of a Gaussian process framework for white matter fibers," *NeuroImage*, vol. 51, pp. 228–241, 2010.

[18] B.C. Vemuri, M. Liu, S. Amari, and F. Nielsen, "Total Bregman Divergence and its Applications to DTI Analysis," *IEEE TMI*, vol. 30, no. 2, pp. 475–483, 2011.

[19] M. Liu, B.C. Vemuri, S. Amari, and F. Nielsen, "Shape Retrieval Using Hierarchical Total Bregman Soft Clustering," *TPAMI, under revision*, 2011.

[20] M. Makuuchi, J. Bahlmann, A. Anwander, and A. D. Friederici, "Segregating the core computational faculty of human language from working memory," *Proc. National Acad. Sciences*, vol. 106, no. 20, pp. 8362–8367, 2009.

[21] M. Liu, B.C. Vemuri, and R. Deriche, "Simultaneous Smoothing & Estimation of DTI via Robust Variational Non-local Means," *MICCAI CDMRI*, 2011.