Clustering of the Human Skeletal Muscle Fibers Using Linear Programming and Angular Hilbertian Metrics

Radhouène Neji, Ahmed Besbes, Nikos Komodakis, Mezri Maatoouk, Jean-François Deux, Alain Rahmouni, Guillaume Bassez, Gilles Fleury, Nikolaos Paragios

Clustering of the Human Skeletal Muscle Fibers Using Linear Programming and Angular Hilbertian Metrics

Radhouène Neji[1,2,3], Ahmed Besbes[1,2], Nikos Komodakis[4], Jean-François Deux[5], Mezri Maatouk[5], Alain Rahmouni[5], Guillaume Bassez[5], Gilles Fleury[3], and Nikos Paragios[1,2]

[1] Laboratoire MAS, Ecole Centrale Paris, Châtenay-Malabry, France
[2] Equipe GALEN, INRIA Saclay - Île-de-France, Orsay, France
[3] Département SSE, Ecole Supérieure d'Electricité, Gif-sur-Yvette, France
[4] Department of Computer Science, University of Crete, Crete, Greece
[5] Centre Hospitalier Universitaire Henri Mondor, Créteil, France ⋆

Abstract. In this paper, we present a manifold clustering method for the classification of fibers obtained from diffusion tensor images (DTI) of the human skeletal muscle. Using a linear programming formulation of prototype-based clustering, we propose a novel fiber classification algorithm over manifolds that circumvents the necessity to embed the data in low dimensional spaces and determines automatically the number of clusters. Furthermore, we propose the use of angular Hilbertian metrics between multivariate normal distributions to define a family of distances between tensors that we generalize to fibers. These metrics are used to approximate the geodesic distances over the fiber manifold. We also discuss the case where only geodesic distances to a reduced set of landmark fibers are available. The experimental validation of the method is done using a manually annotated significant dataset of DTI of the calf muscle for healthy and diseased subjects.

1 Introduction

Diffusion Tensor Imaging (DTI) has started to become more ubiquitous in other fields than brain white matter study [1]. Indeed, this modality has been used for other anatomical regions such as the tongue [2] and the human skeletal muscles [3]. The latter are of particular interest because they present an architecture of elongated myofibers with well known anatomy. Furthermore, the study of the effects of myopathies (neuromuscular diseases) on water diffusion in muscle tissues is essential to assess the possibility of the use of DTI in a diagnosis procedure and early detection of diseases. Since myopathies result in an atrophy and weakness of the muscle, we expect an alteration of the diffusion properties

⋆ This work was partially supported by Association Française contre les Myopathies (AFM: `http://www.afm-france.org`) under the DTI-MUSCLE project.

among diseased subjects. It is therefore important to cluster fiber tracts for local statistical analysis of diffusion information.

DTI previous studies of the human skeletal muscle [3, 4] provided a comparative study between subjects and different muscle regions of scalar values derived from tensors like trace, fractional anisotropy, etc. They also evaluated experimentally the physiological cross-sectional area (PCSA), which is an important measure of muscle architecture since it is related to the maximum muscle force. However little emphasis was put on muscle segmentation in comparison with brain white matter, where several approaches were proposed. The use of graph theory and manifold learning has been extensively explored in the previous literature. For instance, in [5] the distribution of points along each fiber tract is considered to be Gaussian, which allows to derive a Euclidean distance between each pair of fibers. Fiber bundling is done using a normalized cut. In [6], the affinity between fibers is based on the symmetrized Hausdorff distance and spectral clustering is achieved using an eigenanalysis of the affinity matrix and k-means in the embedding space. The method presented in [7] relies on Laplacian Eigenmaps and similarity between fibers is determined using their end points. In [8], the authors construct a graph-based distance between fiber tracts where both local and global dissimilarities are taken into account. The considered distance is then incorporated in a Locally Linear Embedding framework and clustering is done using k-means. Curve modeling has attracted attention and was handled in [9] by defining a spatial similarity measure between curves and using the Expectation-Maximization algorithm for clustering. The method proposed in [10] considers the simultaneous use of medoid-shift clustering and Isomap manifold learning and proposed to include prior knowledge in the segmentation process using a white matter fiber atlas. Mean-shift was also used in [11] where each fiber is first embedded in a high dimensional space using its sequence of points, and kernels with variable bandwidths are considered in the mean-shift algorithm. More recently, fibers were represented in [12] using their differential geometry and frame transportation and a consistency measure was used for clustering. Another class of methods suggested to circumvent the limitation of unsupervised clustering where the obtained segmentation may not correspond to anatomical knowledge. They opt for supervised algorithms that try to achieve a clustering consistent with a predefined atlas. Expert manual labeling of the fibers for one subject provides an atlas in [13]. This is followed by the registration of B0 images and a hierarchical classification of fibers where the B-spline coefficients of the curves are considered to measure curve similarity. In [6], a Nystrom approximation of the out-of-sample extension of the spectral embedding is considered to build an atlas of fibers.

We can note that the existing literature puts a lot of emphasis on manifold embeddings. They are considered crucial to reflect faithfully the diffusion process modeled by tensors and fibers, and proved to be useful for a more accurate analysis of DTI information [14]. However, the use of embeddings and common clustering techniques like k-means requires to choose the dimension of the embedding and the number of clusters. It would be preferable to obtain the

number of clusters as a result of the clustering algorithm, especially when the inter-subject variability (which is rather important for skeletal muscles) may require the use of different numbers of clusters across patients. Moreover, selecting the embedding dimension is an issue since a too low dimension will result in information loss and a too high dimension will include an important dispersion in the data. Furthermore, clustering on the manifold directly is a tricky issue since one has to compute intrinsic means on submanifolds where an explicit expression of geodesic distances is not necessarily available. Another issue is the sensitivity of methods like k-means to initialization and the possible failure of the medoid-shift technique to determine correctly the modes of a density[15]. Besides, when dealing with fiber similarities, the prior art seems to discard the information provided by the tensor field when considering metrics between fibers. In [16], we proposed a kernel between tensors primarily, generalized it to fiber tracts and used k-means clustering after kernel PCA and Isomap embedding. In this paper, we propose a method that performs manifold clustering of fibers without resorting to manifold embeddings or computations of intrinsic means. It is based on a linear programming (LP) technique [17] and uses the geodesic distances from the fibers to a reduced set of landmark fibers to perform the clustering. Unlike k-means, the algorithm provides automatically the number of clusters, is not sensitive to initialization and the class centers are chosen as examplars from the dataset. As far as fiber similarity is concerned, we develop the viewpoint that we proposed in [16] and build Hilbertian angular metrics between fibers. These are derived from their counterparts between tensors, providing a more general and much simpler formulation than [16]. The metrics are incorporated afterwards in the Dijkstra algorithm to approximate the geodesic distances along the manifold of fibers.

The remainder of the paper is organized as follows: in section 2, we present the clustering method and develop the landmark-based geodesic clustering costs. In section 3, we discuss and derive the family of Hilbertian angular metrics between tensors and propose their extension to fiber tracts. Section 4 is dedicated to the experimental results and we discuss the perspectives of this work in section 5.

2 Manifold Clustering via Linear Programming

Clustering refers to the process of organizing a set of objects into groups such that the members of each group are as similar to each other as possible. A common way of tackling this problem is to formulate it as the following optimization task: given a set of objects $\mathcal{V} = \{p_1, \ldots, p_n\}$, endowed with a distance function $d(\cdot, \cdot)$ that measures dissimilarity between objects, the goal of clustering is to choose K objects from $\mathcal{V}$, say, $\{q_1, \ldots, q_K\}$ (these will be referred to as cluster centers hereafter) such that the obtained sum of distances between each object and its nearest center is minimized, or:

$$\min_{q_1,\ldots,q_K \in \mathcal{V}} \sum_{p \in \mathcal{V}} \min_i d(p, q_i) \quad . \tag{1}$$

An important drawback of the above formulation is that it requires the number of clusters K to be provided beforehand, which is problematic as this number is very often not known in advance. Note that a wrong value for K may have a very negative effect on the final outcome. One would thus prefer K to be automatically estimated by the algorithm as a byproduct of the optimization process. To address this issue, we will let K be a variable here, and, instead of (1), we will use the following modified objective function, which additionally assigns a penalty $g(q_i)$ to each one of the chosen cluster centers q_i:

$$\min_{K} \min_{q_1,\ldots,q_K \in \mathcal{V}} \left(\sum_{p \in \mathcal{V}} \min_{i} d(p, q_i) + \sum_{i=1}^{K} g(q_i) \right) \quad . \tag{2}$$

But, even if K is known, another serious drawback of many of the existing optimization-based techniques for clustering is that they are particularly sensitive to initialization and thus may get easily trapped in bad local minima. For instance, K-means (one of the most commonly used clustering methods) is doomed to fail if its initial cluster centers happen not to be near the actual cluster centers. To deal with that, here we will rely on a recently proposed clustering algorithm [17], which has been shown to yield approximately optimal solutions to the NP-hard problem (2). This algorithm relies on reformulating (2) as an equivalent integer program, whose LP-relaxation (denoted as PRIMAL hereafter) has the following form:

$$\text{PRIMAL} \equiv \min_{\mathbf{x}} \sum_{p,q \in \mathcal{V}, p \neq q} d(p, q) x_{pq} + \sum_{q \in \mathcal{V}} g(q) x_{qq} \tag{3}$$

$$\text{s.t.} \sum\nolimits_{q \in \mathcal{V}} x_{pq} = 1,\ x_{pq} \leq x_{qq},\ x_{pq} \geq 0 \tag{4}$$

If constraints $x_{pq} \geq 0$ are replaced with $x_{pq} \in \{0, 1\}$, then the resulting integer program is equivalent to clustering problem (2). In this case, each binary variable x_{pq} with $p \neq q$ indicates whether object p has been assigned to cluster center q or not, while binary variable x_{qq} indicates whether object q has been chosen as a cluster center or not. Constraints $\sum_{q \in \mathcal{V}} x_{pq} = 1$ simply express the fact that each object must be assigned to exactly one center, while constraints $x_{pq} \leq x_{qq}$ require that if p has been assigned to q then object q must obviously be chosen as a center. The most crucial issue for tackling this integer LP is setting the variables x_{qq} correctly, *i.e*, deciding which objects will be chosen as centers. To this end, the so-called *stability* of an object has been introduced in [17]. This is a measure which tries to quantitatively answer the following question: *how much does one need to further penalize an object to ensure that it will never be selected as an optimal cluster center?* Intuitively, the greater the stability of an object, the more appropriate that object is to become a cluster center. For having a practical algorithm based on object stabilities, an efficient way of estimating them is required. It turns out that this can indeed be done very fast by moving to the dual domain and appropriately updating a solution of a dual relaxation to PRIMAL. Since each dual cost provides a lower bound to the cost of the optimal

clustering, an additional advantage of working in the dual domain is the ability to provide online optimality guarantees and to avoid bad local minima. We refer the reader to [17] for more details.

We now discuss the case where the objects lie on a manifold. This implies the use of the geodesic distance as a similarity measure. Ideally this distance should correspond to the pairwise cost $d(p,q)$ for $p \neq q$ in the linear programming formulation proposed in (3). A first possible choice is to compute the geodesic distances between all the pairs of points using the Dijkstra algorithm in an Isomap-like fashion, as suggested in [10]. The shortest path is found using a local approximation of the geodesic distance, for example a Euclidean distance. The pairwise cost $d(p,q)$ is set to $d(p,q) = d_g(p,q)$ where d_g is the corresponding geodesic distance. However, inspired by the landmark Isomap algorithm [18], we can compute the geodesic distances from all the data points to a reduced set of randomly selected landmarks. This will reduce the computational load that a full computation of the geodesic distances between every pair of data points would entail. Let $(l_m)_{m=1\dots n_l}$ be a set of such chosen n_l landmarks. We would like to replace $d_g(p,q)$ by a reasonable approximation. Given that the geodesic distance between two points is the length of the shortest path linking these points, we note the following $\forall m \in [1 \dots n_l]$, $|d_g(p,l_m) - d_g(q,l_m)| \leq d_g(p,q) \leq d_g(p,l_m) + d_g(q,l_m)$, which implies

$$\sup_m |d_g(p,l_m) - d_g(q,l_m)| \leq d_g(p,q) \leq \inf_m (d_g(p,l_m) + d_g(q,l_m)) \tag{5}$$

This provides a lower bound and an upper bound to the cost $d_g(p,q)$ in the case where only the geodesic distances to some landmarks are computed. Note that in the particular case where p and q are landmarks $d_g(p,q) = \sup_m |d_g(p,l_m) - d_g(q,l_m)| = \inf_m (d_g(p,l_m) + d_g(q,l_m))$. On the other hand we can also note that

$$\inf_m (d_g(p,l_m) + d_g(q,l_m)) - 2\eta \leq \quad d_g(p,q) \tag{6}$$

$$d_g(p,q) \leq \quad \sup_m |d_g(p,l_m) - d_g(q,l_m)| + 2\eta \tag{7}$$

where $\eta = \inf_m \min(d_g(p,l_m), d_g(q,l_m))$. Therefore it makes sense to replace the cost $d_g(p,q)$ whether by its upper bound or its lower bound, since both approximate the cost up to 2η. A byproduct of inequalities (6) and (7) is that both approximations are exact if p *or* q are landmarks, since in that case we have $\eta = 0$.

It is interesting to note in this setting that the lower bound is the L^∞ norm between the distance-to-landmarks representation of p and q. Indeed, let $\mathbf{u_p}$ (resp. $\mathbf{u}_q$) be the n_l-dimensional vector of geodesic distances of p (resp. q) to the landmarks

$$\mathbf{u}_p = [d_g(p,l_1), \dots, d_g(p,l_{n_l})]^t \ , \ \mathbf{u}_q = [d_g(q,l_1), \dots, d_g(q,l_{n_l})]^t \tag{8}$$

By definition, $\sup_m |d_g(p,l_m) - d_g(q,l_m)| = ||\mathbf{u}_p - \mathbf{u}_q||_\infty$. Thus the lower bound approximation has the advantage of defining a metric cost. Intuitively, for a number of landmarks sufficiently larger than the intrinsic dimension of the manifold,

the distance vector representation will provide a good characterization of the points on the manifold.

In order to apply the clustering framework to fiber tracts, we will define a Euclidean structure over the fiber domain. For this purpose, we show in the next section how to map the fibers to a Hilbert space and derive corresponding metrics that will provide a local approximation of the geodesic distance.

3 From Metrics on Tensors to Metrics on Fibers

In this section, we build a family of Hilbertian metrics between fibers that will be incorporated in the Dijkstra algorithm to find the shortest path (and thus d_g) between two elements of the fiber set. The starting point is to consider angular distances between diffusion tensors based on Gaussian probability densities and generalize these distances to the fiber domain.

3.1 Multivariate Normals: a Subset of the Exponential Distributions Family

The structure of the set of multivariate normal distributions $\mathcal{M}$ as a statistical manifold endowed with the Fisher information geometry was discussed in [19], where a closed-form solution of the geodesic distance over this manifold is available for the particular case of Gaussian distributions with common mean. Here we view the multivariate normal distributions as a subset of the exponential distributions family. Let us consider a normal probability density p. In this context, given the exponential decay of the distribution, it is interesting to notice that not only p is an element of the Hilbert space L^2 of square integrable functions but any power p^α, with α a strictly positive real number is also square integrable. This motivates the use of normalized probability product kernels [20] to define a family of angular similarities between multivariate normal distributions. Indeed, considering two elements p_1 and p_2 of $\mathcal{M}$ and $\alpha \in \mathbb{R}_+^*$, we can define the following similarity $C_\alpha(p_1, p_2)$ between p_1 and p_2 as follows:

$$C_\alpha(p_1, p_2) = \frac{\int p_1(\mathbf{x})^\alpha p_2(\mathbf{x})^\alpha d\mathbf{x}}{\sqrt{\int p_1(\mathbf{x})^{2\alpha} d\mathbf{x}} \sqrt{\int p_2(\mathbf{x})^{2\alpha} d\mathbf{x}}} \tag{9}$$

C_α is simply the normalized L^2 inner product between p_1^α and p_2^α. It is therefore the cosine of the angle between p_1^α and p_2^α. It defines a Mercer kernel over the space of multivariate normal distributions, i.e. for any subset $(p_i)_{i=1\ldots N}$ of $\mathcal{M}$, the Gram matrix G of C_α with entries $G_{ij} = C_\alpha(p_i, p_j)$ is semi-definite positive. The Mercer property allows the construction of a mapping ϕ_α associated with the kernel C_α that provides an embedding of $\mathcal{M}$ in the Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}_\alpha$ such that $C_\alpha(p_1, p_2) = < \phi_\alpha(p_1), \phi_\alpha(p_2) >_{\mathcal{H}_\alpha}$, where $< .,. >_{\mathcal{H}_\alpha}$ is the inner product of $\mathcal{H}_\alpha$. Given that C_α is a normalized scalar product, i.e. $C_\alpha(p, p) = 1$, we can define the following Hilbertian metric $d_{\alpha|\mathcal{H}_\alpha}$:

$$d_{\alpha|\mathcal{H}_\alpha}(p_1, p_2) = \sqrt{C_\alpha(p_1, p_1) - 2C_\alpha(p_1, p_2) + C_\alpha(p_2, p_2)} = \sqrt{2 - 2C_\alpha(p_1, p_2)} \tag{10}$$

In the following subsection, we derive the closed-form expression of C_α and $d_{\alpha|\mathcal{H}_\alpha}$ for normal distributions that model a local diffusion process.

3.2 Explicit Derivation of the Angular Distances

Let us consider the Gaussian distribution p that models the motion distribution of water protons at a location $\mathbf{x}$ with a tensor $\mathbf{D}$. Given a diffusion time t, the probability of displacement from the position $\mathbf{x}$ to the position $\mathbf{y}$ is provided by the following equation:

$$p(\mathbf{y}|\mathbf{x}, t, \mathbf{D}) = \frac{1}{\sqrt{\det(\mathbf{D})(4\pi t)^3}} \exp(-\frac{(\mathbf{y}-\mathbf{x})^t \mathbf{D}^{-1}(\mathbf{y}-\mathbf{x})}{4t}) \tag{11}$$

We now consider two normal distributions p_1 and p_2 with parameters $(x_1, \mathbf{D}_1)$ and $(x_2, \mathbf{D}_2)$ respectively. Based on [20] and (11), we can see that C_α is the product of two terms:

$$C_\alpha(p_1, p_2) = C_\alpha^{tensor}(\mathbf{D}_1, \mathbf{D}_2) C_\alpha^{spatial}(p_1, p_2) \tag{12}$$

where

$$\begin{aligned} C_\alpha^{tensor}(\mathbf{D}_1, \mathbf{D}_2) &= 2\sqrt{2}\, \frac{\det(\mathbf{D}_1)^{\frac{1}{4}} \det(\mathbf{D}_2)^{\frac{1}{4}}}{\sqrt{\det(\mathbf{D}_1 + \mathbf{D}_2)}} \\ C_\alpha^{spatial}(p_1, p_2) &= \exp\left(-\frac{\alpha}{4t}(\mathbf{x}_1^t \mathbf{D}_1^{-1} \mathbf{x}_1 + \mathbf{x}_2^t \mathbf{D}_2^{-1} \mathbf{x}_2)\right) \times \\ &\quad \exp\left(\frac{\alpha}{4t}(\mathbf{D}_1^{-1}\mathbf{x}_1 + \mathbf{D}_2^{-1}\mathbf{x}_2)^t (\mathbf{D}_1^{-1} + \mathbf{D}_2^{-1})^{-1} (\mathbf{D}_1^{-1}\mathbf{x}_1 + \mathbf{D}_2^{-1}\mathbf{x}_2)\right) \end{aligned} \tag{13}$$

We notice that $C_\alpha^{spatial}$ has a much simpler expression. Indeed, using the following inversion properties

$$(\mathbf{D}_1 + \mathbf{D}_2)^{-1} = \mathbf{D}_1^{-1} - \mathbf{D}_1^{-1}(\mathbf{D}_1^{-1} + \mathbf{D}_2^{-1})^{-1}\mathbf{D}_1^{-1} \tag{14}$$

$$(\mathbf{D}_1 + \mathbf{D}_2)^{-1} = \mathbf{D}_2^{-1} - \mathbf{D}_2^{-1}(\mathbf{D}_1^{-1} + \mathbf{D}_2^{-1})^{-1}\mathbf{D}_2^{-1} \tag{15}$$

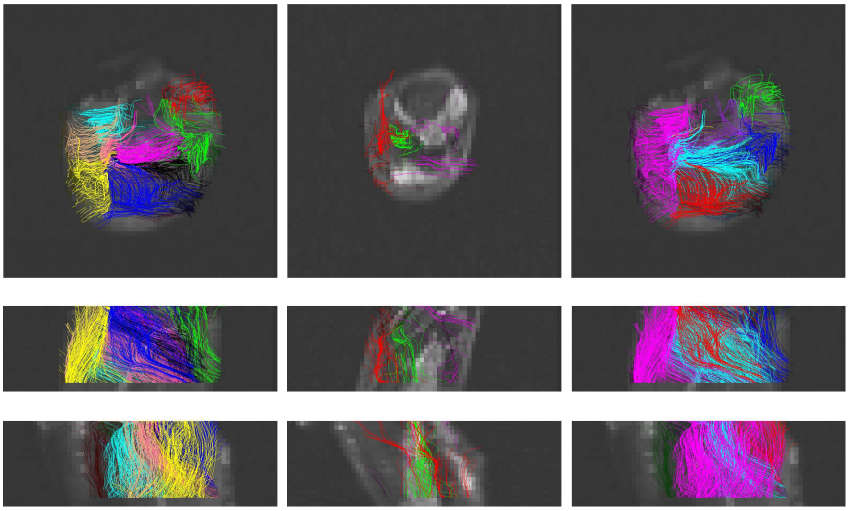
we obtain the following compact expression for $C_\alpha^{spatial}$:

$$C_\alpha^{spatial}(p_1, p_2) = \exp\left(-\frac{\alpha}{4t}(\mathbf{x}_1 - \mathbf{x}_2)^t (\mathbf{D}_1 + \mathbf{D}_2)^{-1} (\mathbf{x}_1 - \mathbf{x}_2)\right) \tag{16}$$

We can see that C^{tensor} is a tensor similarity term and is independent of the parameter α while $C_\alpha^{spatial}$ is a spatial connectivity term where appears the Mahalanobis distance between the locations $\mathbf{x}_1$ and $\mathbf{x}_2$ with respect to the sum of tensors $(\mathbf{D}_1 + \mathbf{D}_2)$. Therefore C_α takes into account the tensor affinity as well as the spatial position. This is crucial since the combination of spatial and diffusion information allows for a better modeling of the interactions between tensors and favors a generalization to the fiber domain, as will be discussed in the next subsection. The diffusion time t is important to weight the contribution of each term and $t \to \infty$ corresponds to the case where the spatial interaction is

not taken into account. Furthermore, there is a striking similarity between the proposed family of measures since α appears as a scale parameter in the exponential function. Given the present formulation, we can conclude that changing the parameter α amounts to a rescaling of the diffusion time t. The derivation of the metrics $d_{\alpha|\mathcal{H}_\alpha}$ is handily done using (10).

In the next subsection, we show how the Mercer property of C_α allows the definition of angular similarities between fiber tracts.



(a) (b) (c)

Fig. 1. Axial, coronal and sagittal views of fiber segmentation obtained with the lower bound approximation for (a) a healthy subject in 10 classes (b) a diseased subject in 3 classes. The parameter β was set to 10 in both cases. In (c) the ground truth segmentation of (a) with the following muscles: the soleus (cyan) , lateral gastrocnemius (red), medial gastrocnemius (magenta), posterior tibialis (yellow), anterior tibialis (green), extensor digitorum longus (purple), and the peroneus longus (blue).

3.3 Angular Similarities between Fibers

A fiber tract is obtained by following the principal directions of diffusion of the tensor field starting from an initial location. It is therefore natural to represent a fiber $\mathbf{F}$ as a sequence of Gaussian probability measures $(p_i)_{i=1\ldots N}$ where N is the number of points of the fiber. Every probability measure (p_i) has a pair of parameters $(\mathbf{x}_i, \mathbf{D}_i)$ where $\mathbf{x}_i$ is the spatial location and $\mathbf{D}_i$ is the tensor at $\mathbf{x}_i$ when the tensor field is supposed to be continuous. When considering the mapping ϕ_α of these measures in the RKHS $\mathcal{H}_\alpha$, we can represent $\mathbf{F}$ as a

weighted average of $(\phi_\alpha(p_i))_{i=1\ldots N}$, i.e. $\mathbf{F} = \sum_{i=1}^{N} w_i \phi_\alpha(p_i)$. A straightforward choice of weights is $\forall i, w_i = \frac{1}{N}$.

Let us consider a fiber $\mathbf{F}_1$ (resp. $\mathbf{F}_2$) represented using a set of probabilities $(p_i^{(1)})_{i=1\ldots N_1}$ (resp. $(p_i^{(2)})_{i=1\ldots N_2}$) and weights $w_i^{(1)}$ (resp. $w_i^{(2)}$). The angular similarity $\widehat{C}_\alpha$ between $\mathbf{F}_1$ and $\mathbf{F}_2$ is defined as follows:

$$\widehat{C}_\alpha(\mathbf{F}_1, \mathbf{F}_2) = \frac{< \sum_{i=1}^{N_1} w_i^{(1)} \phi_\alpha(p_i^{(1)}), \sum_{j=1}^{N_2} w_j^{(2)} \phi_\alpha(p_j^{(2)}) >_{\mathcal{H}_\alpha}}{\left\| \sum_{i=1}^{N_1} w_i^{(1)} \phi_\alpha(p_i^{(1)}) \right\|_{\mathcal{H}_\alpha} \left\| \sum_{j=1}^{N_2} w_j^{(2)} \phi_\alpha(p_j^{(2)}) \right\|_{\mathcal{H}_\alpha}} \tag{17}$$

Using the bilinearity of the inner product $< ., . >_{\mathcal{H}_\alpha}$, we can express $\widehat{C}_\alpha$ using C_α :

$$\widehat{C}_\alpha(\mathbf{F}_1, \mathbf{F}_2) = \frac{\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} w_i^{(1)} w_j^{(2)} C_\alpha(p_i^{(1)}, p_j^{(2)})}{\left\| \sum_{i=1}^{N_1} w_i^{(1)} \phi_\alpha(p_i^{(1)}) \right\|_{\mathcal{H}_\alpha} \left\| \sum_{j=1}^{N_2} w_j^{(2)} \phi_\alpha(p_j^{(2)}) \right\|_{\mathcal{H}_\alpha}} \tag{18}$$

where $\left\| \sum_{i=1}^{N_k} w_i^{(k)} \phi_\alpha(p_i^{(k)}) \right\|_{\mathcal{H}_\alpha} = \sqrt{\sum_{i=1}^{N_k} \sum_{j=1}^{N_k} w_i^{(k)} w_j^{(k)} C_\alpha(p_i^{(k)}, p_j^{(k)})}$ for $k = \{1, 2\}$. Again the corresponding Hilbertian metric between fibers is derived in a similar way to (10). Note that the present formulation endows the fiber domain with an Euclidean structure without resorting to a dimensionality reduction step.

4 Experimental Validation

Thirty subjects (twenty healthy subjects and ten patients affected by myopathies) underwent a diffusion tensor imaging of the calf muscle using a 1.5 T MRI scanner with the following parameters : repetition time (TR)= 3600 ms, echo time (TE) = 70ms, slice thickness = 7mm and b value of 700$s.mm^{-2}$ with 12 gradient directions and 13 repetitions. The size of the obtained volumes is $64 \times 64 \times 20$ voxels with a voxel resolution of 3.125 $mm \times 3.125$ $mm \times 7$ mm. T1-weighted volumes were simultaneously acquired, so they are naturally registered to the diffusion images. They were afterwards manually segmented by an expert in 7 classes [Fig.1 (c)].

Fiber tracts were reconstructed using [21], based on a manual region of interest. To obtain the ground-truth class of each fiber, we counted the number of voxels belonging to each muscle group that the fiber crosses and assigned the latter to the majority class. In our experiments we set the diffusion time to $t = 2\ 10^4$ and the parameter α in the fiber metric to $\alpha = 1$. To compute the Hilbertian metrics between fiber tracts, the weights w_i of each fiber $\mathbf{F}$ in (18) were chosen as the inverse of the number of points in $\mathbf{F}$. We selected 30% of the fibers as landmarks and for the computation of the geodesic distances using the Dijkstra algorithm, we considered a k-NN graph where k was set to $k = 12$. The cost $g(\mathbf{F})$ of choosing a fiber $\mathbf{F}$ as a class center in (3) was set to a constant $g = \beta\ \mu_{\frac{1}{2}}\ (d_g(\mathbf{F}_i, \mathbf{F}_j)_{i \neq j})$ where $\mu_{\frac{1}{2}}$ is the statistical median. We tested the following values of β: $\{7, 10, 13\}$. For a quantitative assessment of the

method, we measure the dice overlap coefficient between the obtained segmentation using the proposed method and the ground-truth segmentation provided by the expert. For the sake of comparison, we evaluate also the performance of k-means clustering using the same metric and a manifold embedding. The dimensionality of the embedding is chosen to be the number of clusters obtained by our method, which is a common choice in embedding-based approaches. The k-means algorithm is run 50 times and each time we compute the dice overlap of the clustering result with the ground-truth segmentation. We consider both the average dice coefficients over the restarts of the k-means algorithm and the dice coefficient of the clustering with the least distortion. We run the following experiments:

1. We compute all the geodesic distances between every pair of points and use them for linear programming clustering. We compare the obtained result with an Isomap embedding followed by k-means.
2. We compute the geodesic distances to a set of landmarks and use the lower (resp. upper) bound approximation for linear programming clustering. We compare the obtained result with a landmark-Isomap embedding followed by k-means.

We provide in [Fig.2 (a), (b), (c)] the boxplots showing the distributions of the dice coefficients for the thirty subjects using different values of β for our algorithm, compared with k-means after manifold embedding. We can note that linear programming clustering performs significantly better than the average score achieved by k-means both for a full and landmark-based computation of the geodesic distances. Furthermore, it achieves results equivalent to the best k-means with an average dice coefficient of approximately 0.8 and in some cases it improves marginally the dice overlap. The advantage is that our result is reproducible, i.e. unlike k-means it is not sensitive to initialization. When comparing the three versions of linear programming clustering, we can see in [Fig.2 (d)] that the lower bound and upper bound approximations perform similarly apart from the case $\beta = 10$ where the lower bound approximation performed better, which may be explained by the metricity of the corresponding cost. The full computation yields slightly better results than the approximations. This corroborates the analysis provided is section 2. For qualitative evaluation, we show in [Fig.1 (a)] (resp. [Fig.1 (b)]) a clustering result obtained for a healthy (resp. diseased) subject for $\beta = 10$. Ground truth segmentation for the healthy subject is provided in [Fig.1 (c)]. There are too few fibers in [Fig.1 (b)] because the tractography fails to recover fibers through the manual region of interest. This is due to the presence of tensors with very low determinant (low diffusion). It is interesting to note that with the same parameter $\beta = 10$, the algorithm found ten clusters for the healthy subject while it found only three for the diseased patient, which seems to reflect the advantage of letting the number of clusters a variable of the optimization problem. Note also how the soleus (in cyan in [Fig.1 (c)]) is subdivided in an anterior and a posterior part in [Fig.1 (a)], which is consistent with its anatomy of oblique fibers converging towards a central aponeurosis.

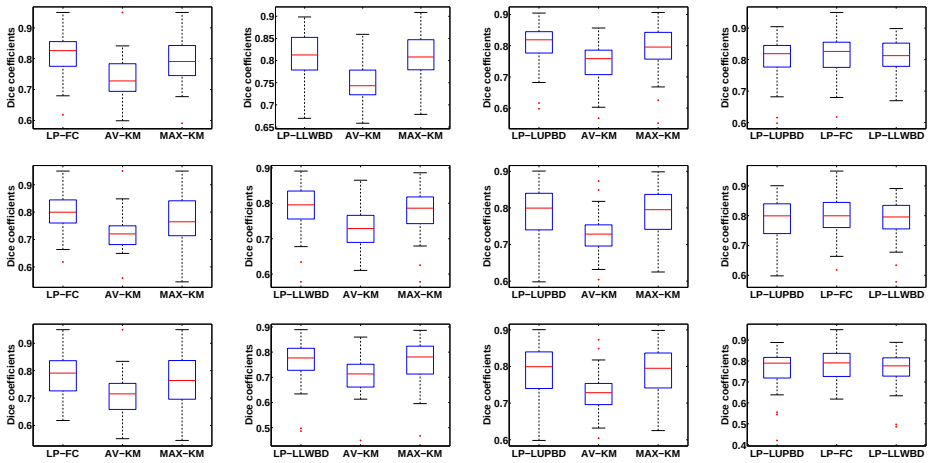


Fig. 2. Boxplots of dice overlap coefficients for the thirty subjects. Each row corresponds to a value of β, from top to bottom β takes the following values 7 , 10 and 13. (a) LP clustering using full computation of distances (LP-FC), comparison is done with respect to the average score of k-means (AV-KM) and the score of the k-means clustering with least distortion (MAX-KM) after manifold embedding. (b) LP clustering using lower bound approximation (LP-LLWBD). (c) LP clustering using upper bound approximation (LP-LUPBD). (d) Comparison between LP-FC, LP-LLWBD and LP-LUPBD.

5 Conclusion

In this paper, we proposed a novel manifold-based fiber clustering approach where there is no need to perform an embedding in a low dimensional space or to select the number of clusters. We applied the method to the bundling of the fibers of the human skeletal muscle. We also developed the theoretical aspects of angular distances between multivariate normal distributions that model local diffusion processes and showed that the RKHS formulation allows for the definition of corresponding metrics between fiber tracts. These metrics were used to approximate the geodesic distances on the fiber manifold using the Dijkstra algorithm. A procedure of landmark selection should be investigated based on the bounds tightness in (6) and (7), as well as other metrics and structures over the fiber domain. Based for example on the metric in (10), the method can also be used for a clustering at the tensor level.

References

1. Bihan, D.L., Mangin, J.F., Poupon, C., Clark, C.A., Pappata, S., Molko, N., Chabrait, H.: Diffusion tensor imaging: Concepts and applications. Journal of Magnetic Resonance Imaging **13** (2001) 534–546

2. Gilbert, R.J., Napadow, V.J.: Three-dimensional muscular architecture of the human tongue determined in vivo with diffusion tensor magnetic resonance imaging. Dysphagia **20** (2005) 1–7
3. Galban, C.J., Maderwald, S., Uffmann, K., de Greiff, A., Ladd, M.E.: Diffusive sensitivity to muscle architecture: a magnetic resonance diffusion tensor imaging study of the human calf. European Journal of Applied Physiology **93**(3) (2004) 253 – 262
4. Damon, B., Ding, Z., Anderson, A., Freyer, A., Gore, J.: Validation of diffusion tensor MRI-based muscle fiber tracking. Magnetic Resonance in Medicine **48** (2002) 97–104
5. Brun, A., Knutsson, H., Park, H.J., Shenton, M.E., Westin, C.F.: Clustering fiber tracts using normalized cuts. In: MICCAI. (2004)
6. ODonnell, L., Westin, C.F.: Automatic tractography segmentation using a high-dimensional white matter atlas. IEEE TMI **26**(11) (2007) 1562–1575
7. Brun, A., Park, H.J., Knutsson, H., Westin, C.F.: Coloring of DT-MRI fiber traces using Laplacian eigenmaps. In: EUROCAST. (2003)
8. Tsai, A., Westin, C.F., Hero, A.O., Willsky, A.S.: Fiber tract clustering on manifolds with dual rooted-graphs. In: CVPR. (2007)
9. Maddah, M., Grimson, W., Warfield, S., Wells, W.: A unified framework for clustering and quantitative analysis of white matter fiber tracts. Medical Image Analysis **12**(2) (2008) 191–202
10. Wassermann, D., Deriche, R.: Simultaneous manifold learning and clustering: Grouping white matter fiber tracts using a volumetric white matter atlas. In: MICCAI 2008 Workshop - Manifolds in Medical Imaging: Metrics, Learning and Beyond. (2008)
11. Zvitia, O., Mayer, A., Greenspan, H.: Adaptive mean-shift registration of white matter tractographies. In: ISBI. (2008)
12. Savadjiev, P., Campbell, J.S.W., Pike, G.B., Siddiqi, K.: Streamline flows for white matter fibre pathway segmentation in diffusion MRI. In: MICCAI. (2008)
13. Maddah, M., Mewes, A.U.J., Haker, S., Grimson, W.E.L., Warfield, S.K.: Automated atlas-based clustering of white matter fiber tracts from DTMRI. In: MICCAI. (2005)
14. Verma, R., Khurd, P., Davatzikos, C.: On analyzing diffusion tensor images by identifying manifold structure using isomaps. IEEE TMI **26**(6) (2007) 772–778
15. Vedaldi, A., Soatto, S.: Quick shift and kernel methods for mode seeking. In: ECCV. (2008)
16. Neji, R., Paragios, N., Fleury, G., Thiran, J.P., Langs, G.: Classification of tensors and fiber tracts using Mercer-kernels encoding soft probabilistic spatial and diffusion information. In: CVPR. (2009)
17. Komodakis, N., Paragios, N., Tziritas, G.: Clustering via LP-based stabilities. In: NIPS. (2008)
18. de Silva, V., Tenenbaum, J.B.: Global versus local methods in nonlinear dimensionality reduction. In: NIPS. (2002)
19. Deriche, R., Tschumperlé, D., Lenglet, C., Rousson, M.: Variational Approaches to the Estimation, Regularization and Segmentation of Diffusion Tensor Images. In Paragios, Chen, Faugeras, eds.: Mathematical Models of Computer Vision: The Handbook. Springer (2005)
20. Jebara, T., Kondor, R., Howard, A.: Probability product kernels. Journal of Machine Learning Research **5** (2004) 819–844
21. Fillard, P., Toussaint, N., Pennec, X.: Medinria: DT-MRI processing and visualization software Similar Tensor Workshop, 2006.