

Classification of Tensors and Fiber Tracts Using Mercer-Kernels Encoding Soft Probabilistic Spatial and Diffusion Information

Radhouène Neji^{1,2,3}

Nikos Paragios^{1,2}

Gilles Fleury³

Jean-Philippe Thiran⁴

Georg Langs⁵

¹Laboratoire MAS, Ecole Centrale Paris, Châtenay-Malabry, France

²Equipe GALEN, INRIA Saclay - Île-de-France, Orsay, France

³Département SSE, Ecole Supérieure d'Electricité, Gif-sur-Yvette, France

⁴Signal Processing Laboratory 5, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

⁵ CIR lab, Department of Radiology, Medical University of Vienna, Vienna, Austria

Abstract

In this paper, we present a kernel-based approach to the clustering of diffusion tensors and fiber tracts. We propose to use a Mercer kernel over the tensor space where both spatial and diffusion information are taken into account. This kernel highlights implicitly the connectivity along fiber tracts. Tensor segmentation is performed using kernel-PCA compounded with a landmark-Isomap embedding and k-means clustering. Based on a soft fiber representation, we extend the tensor kernel to deal with fiber tracts using the multi-instance kernel that reflects not only interactions between points along fiber tracts, but also the interactions between diffusion tensors. This unsupervised method is further extended by way of an atlas-based registration of diffusion-free images, followed by a classification of fibers based on nonlinear kernel Support Vector Machines (SVMs). Promising experimental results of tensor and fiber classification of the human skeletal muscle over a significant set of healthy and diseased subjects demonstrate the potential of our approach.

1. Introduction

Diffusion Tensor Imaging (DTI) is a modality that has been used mainly in the study of the connectivity in the human brain [2]. Diffusion tensors correspond to a field of 3×3 symmetric positive definite matrices that model the uncertainty information (covariance) of the displacements of water protons in the tissues.

The existing diffusion tensor clustering algorithms can be subdivided into three groups. The first class uses a variational approach along with an adequate distance over the

manifold of tensors. For instance in [13], a 3D surface is evolved using a level set representation to segment a region of interest where the distributions of the tensors in each region are modeled as Gaussian. In [28], the Mumford-Shah functional is minimized using a distance between tensors derived from the Burg divergence. Similarly in [12], several similarity measures are investigated and guide the evolution of coupled level sets. A level set technique is also used in [16] to extract the cingulum based on the Finsler metric.

The second class of methods uses common clustering algorithms. In [1], the authors propose to use the Log-Euclidean metric to obtain a kernel density estimate of the probability distribution of tensors and include it in a fuzzy k-means framework where the spatial interactions are handled using local Gaussian kernels with a fixed bandwidth. In [5], mean-shift clustering is applied for the segmentation of the thalamus using Gaussian kernels both in the tensor and in the position space.

The third class of methods consists of graph theoretical approaches and manifold learning techniques. In [29], a graph-cut approach is used with seed-point initialization. Spectral clustering is performed in [30] through the eigenanalysis of an affinity matrix between tensors based on a selected similarity measure. In [9], Locally Linear Embedding (LLE) is considered with different choices of tensor metrics and spatial connectivity is ensured by considering isotropic neighborhoods.

Similar concepts were investigated for fiber clustering. For instance, a Gaussian model of the fibers and a normalized-cut approach based on the Euclidean distance between the obtained moments is presented in [3]. In [19], spectral clustering along with the Hausdorff distance between fibers is considered. In [25], another manifold learning technique is proposed based on a graph-based distance

that captures local and global dissimilarities between fibers and uses LLE for clustering of the tracts. Curve modeling has attracted attention and was handled in [14] by defining a spatial similarity measure between curves and using the Expectation-Maximization algorithm for clustering. More recently, fibers were represented in [22] using their differential geometry and frame transportation and a consistency measure was used for clustering. Of particular interest in the field are the supervised methods that try to achieve a segmentation consistent with a predefined atlas. Registration of B0 images and a hierarchical classification of fibers are performed in [15] using the B-spline representation of fibers. The method proposed in [19] is further extended in [20] by means of a Nystrom approximation of the out-of-sample extension of the spectral embedding.

In the existing work, fiber clustering methods discard the tensor information and rely on the obtained tracts, creating an artificial gap between tensor and fiber clustering. In this paper we introduce a kernel that unifies these two viewpoints, and handles tensor and fiber clustering in a concise probabilistic framework. The method we propose works on tensors and is easily extended to segment fibers while taking into account the tensor field. The kernel quantifies not only the dissimilarity between tensors, but also their mutual spatial relationship. It exploits connectivity along fiber tracts implicitly, i.e. in the feature space provided by the kernel embedding, tensors which are aligned are closer than tensors which do not lie on the same fiber tract. It is extended to deal with fiber tracts by way of the summation kernel, which takes into account both the interactions between the points (as spatial positions) and the information provided by the whole tensor field. We also give an interpretation of the fiber kernel as a comparison between soft representations of the fiber tracts and show that it provides a natural generalization for Gaussian kernel correlation. To overcome the limitation inherent to unsupervised clustering where the clusters may not correspond to prior knowledge (atlas), we show how the algorithm can include expert knowledge by using atlas-based registration and kernel SVM classification.

The remainder of this paper is organized as follows: in section 2, we introduce the kernel over the tensor space. Using the summation kernel concept, this kernel is extended to deal with fiber tracts in section 3, where we also propose a supervised version of the algorithm based on kernel SVMs. Section 4 is dedicated to the experimental results using images of the human skeletal muscle and we discuss the perspectives of this work in section 5.

2. Kernel-based Clustering of Tensors

Diffusion tensors measure the motion distribution of water molecules. More explicitly, they refer to the covariance of a Gaussian probability over the displacements \mathbf{r} of the

water protons given a diffusion (mixing) time t :

$$p(\mathbf{r}|t, \mathbf{D}) = \frac{1}{\sqrt{\det(\mathbf{D})(4\pi t)^3}} \exp\left(-\frac{\mathbf{r}^t \mathbf{D}^{-1} \mathbf{r}}{4t}\right) \quad (1)$$

Given a diffusion tensor \mathbf{D} localized at voxel \mathbf{x} , we can obtain the probability of the position \mathbf{y} of the water molecule previously localized at \mathbf{x} in a straightforward way:

$$p(\mathbf{y}|\mathbf{x}, t, \mathbf{D}) = \frac{1}{\sqrt{\det(\mathbf{D})(4\pi t)^3}} \exp\left(-\frac{(\mathbf{y} - \mathbf{x})^t \mathbf{D}^{-1} (\mathbf{y} - \mathbf{x})}{4t}\right) \quad (2)$$

Therefore, a natural way to define a kernel over the tensor space where position is taken into account is to consider the expected likelihood kernel [11]. Let us consider two tensors \mathbf{D}_1 and \mathbf{D}_2 localized at \mathbf{x}_1 and \mathbf{x}_2 respectively, and a diffusion time t . The expected likelihood kernel $k_t(\tau_1, \tau_2)$ between the pairs $\tau_1 = (\mathbf{D}_1, \mathbf{x}_1)$ and $\tau_2 = (\mathbf{D}_2, \mathbf{x}_2)$ is defined as the expectation of Gaussian probability $p_2(\mathbf{y}|\mathbf{x}_1, t, \mathbf{D}_1)$ under the probability law of $p_1(\mathbf{y}|\mathbf{x}_2, t, \mathbf{D}_2)$ and is given by the following expression:

$$\begin{aligned} k_t(\tau_1, \tau_2) &= E_{p_2(\mathbf{y}|\mathbf{x}_2, t, \mathbf{D}_2)}(p_1(\mathbf{y}|\mathbf{x}_1, t, \mathbf{D}_1)) \quad (3) \\ &= \int p_1(\mathbf{y}|\mathbf{x}_1, t, \mathbf{D}_1) p_2(\mathbf{y}|\mathbf{x}_2, t, \mathbf{D}_2) d\mathbf{y} \quad (4) \end{aligned}$$

Note that the diffusion time t is a natural scale parameter for this kernel and that it is computationally tractable, scale dependent and handles spatial information. Indeed, based on the derivation in [11] and using the expression provided in 2, we obtain the following diffusion kernel:

$$k_t(\tau_1, \tau_2) = \frac{1}{\sqrt{(4\pi t)^3}} k_1(\mathbf{D}_1, \mathbf{D}_2) k_2(\tau_1, \tau_2) \quad (5)$$

where

$$\begin{aligned} k_1(\mathbf{D}_1, \mathbf{D}_2) &= \frac{1}{\sqrt{\det(\mathbf{D}_1 + \mathbf{D}_2)}} \\ k_2(\tau_1, \tau_2) &= \exp\left(-\frac{1}{4t} (\mathbf{x}_1^t \mathbf{D}_1^{-1} \mathbf{x}_1 + \mathbf{x}_2^t \mathbf{D}_2^{-1} \mathbf{x}_2)\right) \times \\ &\exp\left(\frac{1}{4t} (\mathbf{D}_1^{-1} \mathbf{x}_1 + \mathbf{D}_2^{-1} \mathbf{x}_2)^t (\mathbf{D}_1^{-1} + \mathbf{D}_2^{-1})^{-1} (\mathbf{D}_1^{-1} \mathbf{x}_1 + \mathbf{D}_2^{-1} \mathbf{x}_2)\right) \quad (6) \end{aligned}$$

2.1. Properties of the Tensor Kernel

The kernel is based on an L^2 inner product defined on the Hilbert space of square-integrable functions, to which Gaussian probability densities belong. Therefore the kernel verifies the Mercer conditions, i.e. it is positive definite. We now provide an analysis of this kernel:

- The first term $k_1(\mathbf{D}_1, \mathbf{D}_2)$ may be rewritten as fol-

lows:

$$k_1(\mathbf{D}_1, \mathbf{D}_2) = \frac{1}{\sqrt{\det((\mathbf{Id} + \mathbf{D}_1 \mathbf{D}_2^{-1}) \mathbf{D}_2)}} \quad (7)$$

$$= \frac{1}{\sqrt{\det(\mathbf{D}_2)}} \frac{1}{\sqrt{\prod_{i=1}^3 (1 + \lambda_i)}} \quad (8)$$

where \mathbf{Id} is the 3×3 identity matrix and λ_i are the generalized eigenvalues of the pair of matrices $(\mathbf{D}_1, \mathbf{D}_2)$. This is reminiscent of the geodesic distance on the manifold of 3×3 symmetric positive definite matrices $d = \sqrt{\sum_{i=1}^3 (\log(\lambda_i))^2}$ [21] which is also based on the generalized eigenvalues. The distance (respectively the kernel) is increasing (respectively decreasing) with increasing generalized eigenvalues, which is a reasonable behavior (recall that the kernel reflects similarity). The symmetry is ensured in the geodesic distance by the squared logarithm function. The latter is invariant with respect to the inverse transformation $\lambda_i \mapsto \frac{1}{\lambda_i}$ ¹, while it can be seen that the factor $\frac{1}{\sqrt{\det(\mathbf{D}_2)}}$ has a similar role since it preserves the symmetry. Note that the original expression is clearly symmetric. We can conclude that the term k_1 encodes similarity between tensors.

- Two special cases of the second factor k_2 are interesting:

1. When $\mathbf{D}_1 = \mathbf{D}_2 = \mathbf{D}$

$$k_2(\tau_1, \tau_2) = \exp\left(-\frac{1}{8t} (\mathbf{x}_1 - \mathbf{x}_2)^t \mathbf{D}^{-1} (\mathbf{x}_1 - \mathbf{x}_2)\right) \quad (9)$$

As expected, when the tensors are equal, what appears is the Mahalanobis distance between positions \mathbf{x}_1 and \mathbf{x}_2 with respect to \mathbf{D} . In particular when the tensor \mathbf{D} is isotropic, i.e. $\mathbf{D} = \mu \mathbf{Id}$, $k_2(\tau_1, \tau_2) = \exp\left(-\frac{1}{8\mu t} \|\mathbf{x}_1 - \mathbf{x}_2\|^2\right)$, which is plainly a Gaussian kernel between the points \mathbf{x}_1 and \mathbf{x}_2 .

2. When $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$, $k_2(\tau_1, \tau_2) = 1$, which means that the kernel k_t reduces to k_1 . Again this was expected since the kernel will rely only on tensor similarity if there is no difference in spatial positions.

- The first special case shows that connectivity between tensors which are aligned on the same fiber tract is enhanced. Let us consider the tensor configuration in [Fig. 1 (a)] where all tensors are equal to $\mathbf{D} = \mu(\vec{\mathbf{e}}_1 \vec{\mathbf{e}}_1^t + \nu \vec{\mathbf{e}}_3 \vec{\mathbf{e}}_3^t)$, where $(\vec{\mathbf{e}}_i)_{i=1,2,3}$ are the canonical basis of \mathbb{R}^3 and $\nu > \mu$ are the eigenvalues of \mathbf{D} . The tensors have therefore a principal direction of diffusion along $\vec{\mathbf{e}}_3$. The tensors are all equal, \mathbf{D}_1 is equidistant spatially from \mathbf{D}_2 and \mathbf{D}_3 , yet the second term k_2 allows to affect more affinity between tensors 1 and 2 than between tensors 1 and 3. Indeed, we can compute the kernel values to obtain

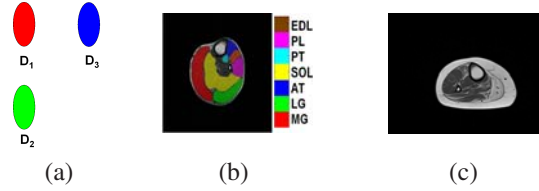


Figure 1. (a) A configuration where the tensor kernel implicitly puts more weight on the connection between \mathbf{D}_1 and \mathbf{D}_2 than between \mathbf{D}_1 and \mathbf{D}_3 , reflecting their alignment. An axial slice of a T1-weighted image showing: (b) a manual segmentation of different muscle groups in the calf (c) the calf of a diseased patient where the zone in hypertension is fat that replaced the muscle.

$k_2(\tau_1, \tau_2) = e^{-\frac{d^2}{8\nu t}}$, $k_2(\tau_1, \tau_3) = e^{-\frac{d^2}{8\mu t}}$ (10)

where $d = \|\mathbf{x}_1 - \mathbf{x}_2\| = \|\mathbf{x}_1 - \mathbf{x}_3\|$. Since $\nu > \mu$, we can see that the similarity between the tensors 1 and 2 is more important than the similarity between the tensors 1 and 3, so the kernel captures locally the fiber structure.

In the next subsection, we discuss how to embed the tensors in a Euclidean space based on the kernel information as a preliminary step to clustering.

2.2. Embedding of the Tensors through Kernel PCA and Landmark Isomap

Let us consider the N pairs $\tau_i = (\mathbf{D}_i, \mathbf{x}_i)_{i=1, \dots, N}$ representing a tensor field. We construct the $N \times N$ kernel matrix \mathbf{K} of entries $\mathbf{K}_{ij} = k_t(\tau_i, \tau_j)$ for a fixed diffusion time t and normalize it to obtain $\tilde{\mathbf{K}}$ such that $\tilde{\mathbf{K}}_{ij} = \frac{\mathbf{K}_{ij}}{\sqrt{\mathbf{K}_{ii} \mathbf{K}_{jj}}}$. These pairs are then embedded in a k -dimensional Euclidean space using an eigenvalue decomposition of $\tilde{\mathbf{K}} = \mathbf{U} \mathbf{S} \mathbf{U}^t$ where \mathbf{U} is an orthogonal $N \times k$ matrix and \mathbf{S} is a $k \times k$ diagonal matrix. The coordinates of the embedded tensors are given by the $N \times k$ matrix $\mathbf{X} = \mathbf{U} \sqrt{\mathbf{S}}$ where $\sqrt{\mathbf{S}}$ is obtained by setting the diagonal elements of \mathbf{S} to their square roots [23]. Each row m of \mathbf{X} holds the coordinates in the feature space of the m -th pair.

Given the k -dimensional representation \mathbf{X} of the tensor field, one has to propagate the local interaction between the tensors and take into account the distance along diffusive pathways, i.e. simulate the water flow along these trajectories. This is done using the Isomap algorithm, which is a

¹Recall that the generalized eigenvalues of the pair $(\mathbf{D}_2, \mathbf{D}_1)$ are the inverse of those of the pair $(\mathbf{D}_1, \mathbf{D}_2)$

manifold learning technique based on two steps [24]: it consists first in approximating the geodesic distances between every two points of the dataset using the Dijkstra algorithm and then perform Multidimensional Scaling (MDS) to obtain the new embedding, i.e. a configuration of points that respects approximately the distance matrix computed in the previous step. Since the kernel enhances fiber connectivity, we expect the new configuration to reflect the diffusion flow in the tissues. Note that in practice we use a faster version of this algorithm called landmark-Isomap [4] that reduces the computational time of the first step by computing the distance of the points to a reduced set of landmarks chosen randomly in the dataset. The clustering is done afterward using a plain k-means algorithm. Note that the normalized kernel k_t implies a Euclidean distance (L^2 norm) given by $d(\tau_1, \tau_2) = \sqrt{2 - 2\tilde{k}_t(\tau_1, \tau_2)}$. We could have used this distance from the outset in the Isomap algorithm, however we opt for a preliminary kernel PCA projection since it amounts to denoising in the feature space and keeping the relevant information present in a few components.

In the following section, we extend in a tractable way the proposed kernel defined over the tensor field to the fiber domain.

3. Extending the Kernel to the Fiber Domain

The fiber trajectories are obtained through the integration of the vector field of principal directions of diffusion. Based on the continuous tensor field approximation (by means of interpolation), we represent each fiber tract as a sequence of tensors localized in spatial positions, i.e. is a set of pairs $\tau_i = (\mathbf{D}_i, \mathbf{x}_i)_{i=1\dots n}$ where n is the number of points lying on the fiber. Note that the tractography already requires tensor interpolation and that the interpolated tensors are therefore kept for kernel computation. So it is natural to extend the tensor kernel using a kernel over sets. We simply use the summation kernel [10] to obtain the following Mercer kernel K_t between two fibers \mathbf{F}_1 and \mathbf{F}_2 :

$$K_t(\mathbf{F}_1, \mathbf{F}_2) = \frac{1}{n_1} \frac{1}{n_2} \sum_{\substack{\tau_i \in \mathbf{F}_1 \\ \tau_j \in \mathbf{F}_2}} k_t(\tau_i, \tau_j) \quad (11)$$

where n_1 (resp. n_2) is the number of points of the fiber \mathbf{F}_1 (resp. \mathbf{F}_2). This kernel sums the interactions between tensors belonging to the fiber tracts. It captures the diffusion and spatial links between diffusive pathways. It is important to notice that while all the interactions are summed, the diffusion time t acts as a scale parameter. Therefore for a suitable choice of t , a tensor interacts only with tensors lying in a local neighborhood and far-away tensors have a negligible impact on the summation. The segmentation of the fiber tracts is achieved using kernel PCA and k-means clustering.

3.1. A Physical Interpretation of the Fiber Kernel

One can see that the summation kernel is simply the expected likelihood kernel between distributions providing a soft representation for fibers. More explicitly, we consider a dynamical system where a particle can be initially at a position \mathbf{x}_i on the fiber tract and moves to a position \mathbf{y} with the following probability

$$p(\mathbf{y}|t, (\mathbf{D}_i)_{i=1\dots n}) = \sum_{i=1}^n p(\mathbf{x}_i) p(\mathbf{y}|\mathbf{x}_i, t, \mathbf{D}_i) \quad (12)$$

With a uniform prior distribution on \mathbf{x}_i , $p(\mathbf{y}|t, (\mathbf{D}_i)_{i=1\dots n}) = \frac{1}{n} \sum_{i=1}^n p(\mathbf{y}|\mathbf{x}_i, t, \mathbf{D}_i)$. By bilinearity of the expected likelihood kernel, it is straightforward to see that the summation kernel is the expected likelihood kernel of the distributions given in (12). If the initial positions \mathbf{x}_i were independent observations (which is not the case because they are the result of the tractography), this would have amounted exactly to an adaptive kernel density estimation of the position of the water molecules along the fiber tracts. In that case, the point-dependent Gaussian kernels of the density estimation would use the diffusion tensors as covariance matrices to model the uncertainty. However the distributions in (12) still provide a soft ("blurry") representation of the fibers and measure the compactness (concentration) of the spatial configuration of the fiber tract according to the diffusion information. More explicitly, instead of considering the fiber as a discrete sum of Dirac distributions localized at its points, i.e. $\frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{x}_i}$, we associate to it its representation via the mapping ϕ to the Reproducing Kernel Hilbert Space (RKHS) provided by the kernel k_t , i.e. $\frac{1}{n} \sum_{i=1}^n \phi(\mathbf{D}_i, \mathbf{x}_i)$.

The unsupervised clustering does not take prior clinical knowledge into account. However, the kernel introduced in (11) accommodates a supervised scheme. In the following, we propose a classification strategy that aims at learning a set of atlas fibers in a discriminative way.

3.2. Supervised Kernel SVM Learning of Atlas Fibers

When considering the special case where the tensor field is constant and isotropic, i.e. all tensors are equal to $\mathbf{D} = \mu \mathbf{Id}$, we obtain the Gaussian kernel correlation KG_t [26]:

$$KG_t(\mathbf{F}_1, \mathbf{F}_2) \propto \frac{1}{n_1} \frac{1}{n_2} \sum_{\substack{\mathbf{x}_i \in \mathbf{F}_1 \\ \mathbf{x}_j \in \mathbf{F}_2}} \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{8\mu t}\right) \quad (13)$$

Therefore, we can see that the kernel K_t deals with a generic tensor field and provides a generalization of the Gaussian kernel correlation. A by-product of this reasoning is that Gaussian kernel correlation is a kernel on fibers that considers only point positions. One could have seen from

the outset that it is a Mercer kernel since it is a summation of Mercer (Gaussian) kernels. This is particularly useful to learn spatial interactions between fibers.

Indeed, given an atlas of fibers segmented by an expert in R regions, we can learn their spatial arrangement using the Gaussian kernel correlation in (13). It can be used as an input in a kernel SVM classifier [27] to learn the training fibers of the atlas. The kernel SVMs provide support vectors (fibers), which are the fibers that define the decision boundaries. Note that the SVMs are used in a one-against-one fashion in order to deal with multiple regions. A point of interest here is that from the initial set of training fibers, we have only to keep a sparse subset of support fibers that will guide the classification process.

Given a new set of fibers that we wish to segment in a manner consistent with the already-defined atlas, we start by finding a spatial transformation that maps the diffusion-free (B0) image of the testing dataset to the corresponding one in the training atlas, as in [15]. This transformation is subsequently used to register the testing set of fibers to the space of the support fibers obtained from the atlas. We are therefore able to compute the scores of the $\frac{R(R-1)}{2}$ pairwise SVM classifiers on the testing dataset and use a simple voting procedure to classify the fibers.

4. Experiments: The Human Skeletal Muscle

The human skeletal muscles and more specifically the lower leg are of particular interest because they present an ordered structure of elongated myofibers where some muscle groups differ only subtly in their direction. In order to study the effect of neuromuscular diseases (myopathies) on water diffusion in the muscles, segmentation in regions consistent with anatomical knowledge is a crucial preliminary step before a localized quantitative study of diffusion properties in the fiber bundles for healthy and diseased tissues. Diffusion-based studies of the human skeletal muscle [7] focused on studying variation across subjects of scalar information like diffusivity (trace), fractional anisotropy, pennation angles (the orientation of muscular fibers with respect to tendons), etc.

Twenty-five subjects (twenty healthy patients and five patients affected by myopathies) underwent a diffusion tensor imaging of the calf muscle using a 1.5 T MRI scanner². The size of the obtained volumes is $64 \times 64 \times 20$ voxels with a voxel resolution of $3.125 \text{ mm} \times 3.125 \text{ mm} \times 7 \text{ mm}$. We acquired simultaneously high resolution T1-weighted images that were segmented manually by an expert into 7 muscle groups to provide the ground truth and fiber trajectories were reconstructed using [6]. To give an idea about

²The following parameters were used : repetition time (TR)= 3600ms, echo time(TE) = 70 ms, slice thickness = 7 mm and b value of 700 s.mm^{-2} with 12 gradient directions and 13 repetitions.

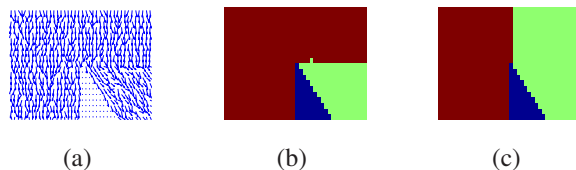


Figure 2. (a) Synthetic noisy field of principal directions of diffusion. (b) Result of spectral clustering. (c) Result of our method

the muscle architecture in the calf, we present in [Fig.1 (b)] a manual segmentation overlaid on an axial slice of a high-resolution T1-weighted image. The following muscle groups are considered: the soleus (SOL), lateral gastrocnemius (LG), medial gastrocnemius (MG), posterior tibialis (PT), anterior tibialis (AT), extensor digitorum longus (EDL), and the peroneus longus (PL).

In the following , we present the obtained experimental results on a synthetic dataset and for tensor classification and both supervised and unsupervised fiber bundling of the lower leg muscles.

4.1. Tensor Clustering

4.1.1 Preliminary Experiments

We first generated a 20×40 lattice of synthetic tensors composed of two close fiber bundles. The first bundle has a vertical principal direction, the second starts with a vertical direction then deviates with a 45° angle [Fig.2 (a)]. We added a Gaussian noise of standard deviation 10° to these directions. The eigenvalues of the tensors were set to $\{2 \cdot 10^{-3}, 1.5 \cdot 10^{-3}, 10^{-3}\}$. We tested the following values of diffusion time t : $\{10^4, 10^5\}$. We compare the behavior of the kernel PCA + Isomap embedding with spectral clustering using the metric $d(\mathbf{D}_1, \mathbf{D}_2) = \arccos(\langle \vec{e}_1, \vec{e}_2 \rangle)$ where \vec{e}_1 (resp. \vec{e}_2) is the principal direction of diffusion of \mathbf{D}_1 (resp. \mathbf{D}_2), as in [30]. Following [30], the scale parameter in the affinity matrix is set as the sample variance of d between neighboring tensors. The clustering is obtained using k-means with 50 restarts in the spectral embedding space. [Fig.2 (c)] shows the segmentation result obtained by our approach (stable across the tested set of diffusion times) and [Fig.2 (b)] shows its counterparts for spectral clustering. We can notice that unlike spectral clustering, the proposed algorithm finds a clustering solution which is more compatible with the tensor arrangement. This is due to the fact that it captures both tensor similarity and spatial connectivity.

To further assess qualitatively the method, we used the proposed kernel PCA + Isomap embedding to see if it is faithful to the known structure of the muscles. Of particular interest is the soleus which is a major part of the calf. It has a bipennate structure where oblique fibers converge towards a central aponeurosis [Fig.3 (c)]. In [Fig.3 (a), (b)], we show

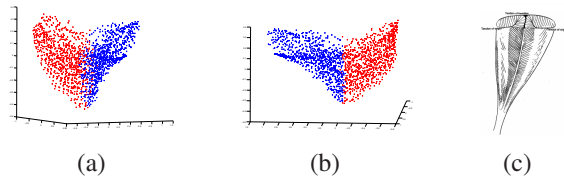


Figure 3. (a), (b) Two views of a three-dimensional embedding of the tensors of the soleus muscle, k-means clustering shows its bipennate structure. (c) Anatomy of the soleus [17].

the (here three-dimensional) proposed embedding for the soleus muscle of one subject for $t = 2 \cdot 10^5$. The obtained points reveal the structure of the muscle, which means that the embedding is faithful to the diffusion flow in the tissues as the points are aligned along the diffusive pathways.

4.1.2 Classification of the Muscle Groups

For each subject, a region of interest (ROI) was manually delineated and we tested the performance of the tensor clustering algorithm both for healthy and diseased subjects. In regions affected by myopathies, the tensors have a relatively small volume since fat replaces the fibers (as can be seen in [Fig. 1 (c)]) and were eliminated through simple thresholding over the determinant. In all the experiments, we set the diffusion time t to $t = 2 \cdot 10^5$ and we used a ten-dimensional embedding. For a quantitative evaluation of the method, we follow the validation protocol proposed in [30]: for all the 25 subjects, the manually delineated ROI was segmented at two different levels (in 7 and 10 classes). The resulting clusters are then classified according to the labels given by the expert. As in [30], several clusters are allowed to have the same label. We also test the algorithm on a section near the knee which is characterized by a higher amount of noise and artifact obtained automatically by a threshold (set to 20) on the diffusion-free images. We report in [Fig. 5] the boxplots of the dice volume overlap with the expert labels for the 25 patients and its counterpart for the spectral clustering as described in the previous subsection. We can see that our algorithm performs slightly better for the case of the manually-delineated ROI, however for the noisy automatic ROI, spectral clustering is misled by isolated points (see [Fig. 4 (c)]), whereas the performance of our algorithm is not significantly worsened. Note also that the thresholding over the determinant removes some of the tensors originating from the artifact. From a qualitative point of view, one can see in [Fig. 4 (a), (b)] that the algorithm was able to segment correctly fine structures like AT and PL and the segmentation result is rather smooth.

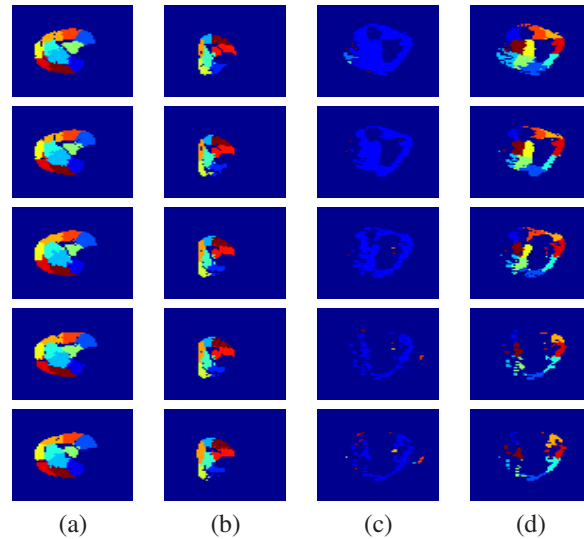


Figure 4. Axial slices of the segmentation of the tensors for (a) a healthy subject in 10 classes, manual ROI (b) a diseased subject in 7 classes where the MG is partially affected, manual ROI (c) noisy automatic ROI of a section near the knee using spectral clustering (d) noisy automatic ROI using our method

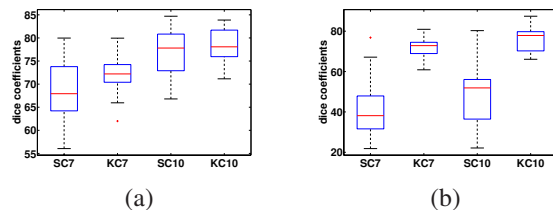


Figure 5. Boxplot of the dice overlap coefficients for the 25 subjects for tensor clustering in 7 and 10 classes for (a) manual ROI and (b) automatic noisy ROI. SC7 (resp. SC10) refers to spectral clustering in 7 (resp. 10) classes and KC7 (resp. KC10) refers to kernel clustering in 7 (resp. 10) classes.

4.2. Fiber Clustering

4.2.1 Unsupervised Fiber Clustering

To test the unsupervised kernel-PCA clustering algorithm, we only kept the fibers which have a majority of points lying in the manually delineated ROI. The number of fibers in the different datasets ranged approximately from 1000 to 2500, which makes the computation and eigenanalysis of the kernel matrix achievable in a rather reasonable time. The diffusion time t was set to $t = 2 \cdot 10^4$ and we used a ten-dimensional embedding for kernel-PCA. [Fig. 6] shows the clustering results in 10 (resp. 7) classes for the fiber tracts of a healthy (resp. diseased) subject. The fibers in [Fig. 6 (a)] correspond to the tensors segmented in [Fig. 4 (a)]. It is interesting to notice that despite the fact that the tractogra-

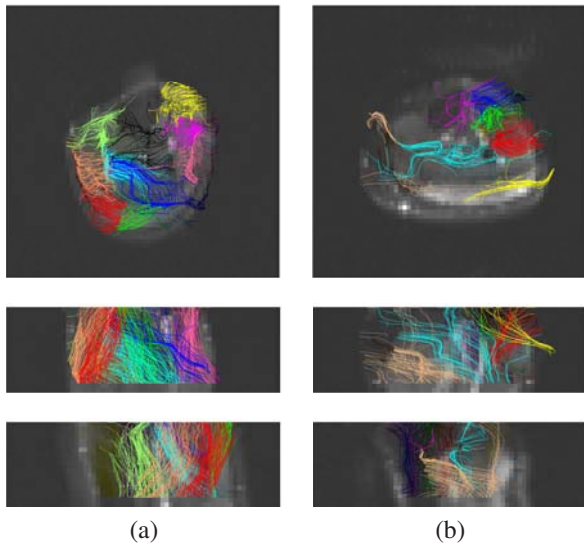


Figure 6. Axial, coronal and sagittal views of fiber segmentation for (a) a healthy subject in 10 classes, (b) a diseased subject in 7 classes.

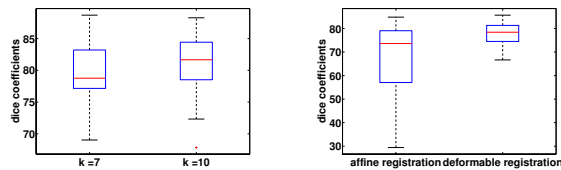


Figure 7. Boxplot of the dice overlap coefficients for the 25 subjects for (a) unsupervised fiber clustering in 7 and 10 classes, (b) supervised fiber clustering using affine and deformable registration.

phy algorithm was unable to recover fiber tracts in the diseased regions due to the presence of degenerate tensors, the clustering algorithm could still segment the fiber tracts of the healthy region in anatomically relevant subgroups. For quantitative assessment, we report in [Fig.7 (a)] the boxplot of the dice overlap measures of the fiber segmentation with the expert labeling for the 25 subjects for 7 and 10 classes. Overall, the algorithm performs well in separating the regions of the calf muscle with a mean dice coefficient of 79.5% (respectively 80.93%) and a standard deviation of 5.04% (respectively 5.14%) for 7 (respectively 10) classes.

4.2.2 Supervised Fiber Clustering

A manually segmented volume was used as an atlas. Atlas fibers were assigned to a class based on a simple voting procedure: each fiber is classified according to the majority vote class of the voxels it crosses. We experimented both with affine [6] and deformable [8] registration to map the

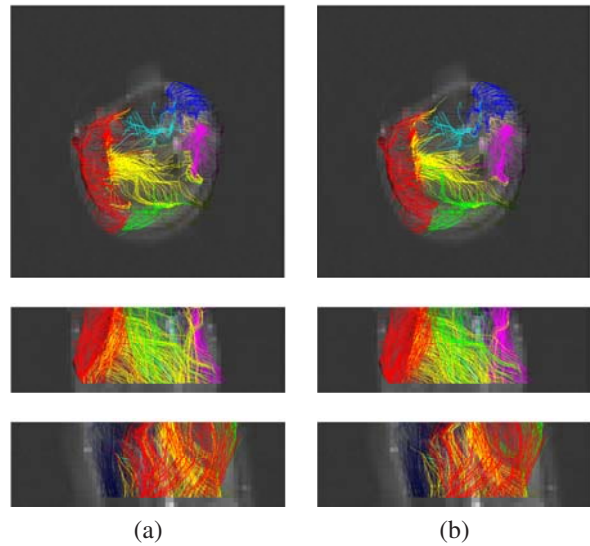


Figure 8. Axial, coronal and sagittal views for (a) supervised classification in 7 classes (b) the ground truth segmentation.

B0 images of the testing case to the B0 images of the atlas. We use kernel SVM classification to learn the fibers of the atlas as explained in section 3.2, using 21 one-against-one pairwise classifiers. The scale parameter in the Gaussian correlation kernel was set to $\sigma = 2\sqrt{\mu t} = 10$. We report in [Fig.7 (b)] the boxplot of the dice overlap coefficients both for deformable and affine registration. We can note that we obtain significantly better results with deformable registration, which was expected given the relatively high inter-patient variability and that muscles are soft tissues, so the anatomy and shape are likely to vary significantly across patients. In [Fig.8 (a)], we show an example of supervised segmentation compared with the ground truth in [Fig.8 (b)]. We can observe that as opposed to the unsupervised setting ([Fig.6 (a)]), the MG is not oversegmented.

5. Conclusion

In this paper, we proposed a kernel-based method for clustering of both tensors and fibers in diffusion tensor images. It exploits the physical interpretation behind the modality and offers a unified approach towards tensor and fiber grouping. The kernel defined over the tensor space encompasses both localization and diffusion information and naturally reflects tensor alignment along fiber tracts. We showed its flexibility by extending it to deal with fibers and gave the physical intuition behind its mathematical definition as a kernel over sets of tensors. We also showed how to include expert knowledge by means of kernel SVMs.

Future research will focus on the use of the defined kernels for deformable registration of diffusion tensors and will

explore the possible extensions of the clustering framework without resorting to manifold embeddings [18]. The motivation of the work is to use DTI in order to detect and monitor the progression of skeletal muscle diseases like myopathies.

Acknowledgments

Data are courtesy of Jean-François Deux, Mezri Maatouk, Guillaume Bassez and Alain Rahmouni, CHU Henri Mondor, Créteil, France. This work was partially supported by Association Française contre les Myopathies (AFM: <http://www.afm-france.org>) under the DTI-MUSCLE project.

References

- [1] S. P. Awate, H. Zhang, and J. C. Gee. A fuzzy, nonparametric segmentation framework for DTI and MRI analysis: With applications to DTI-tract extraction. *IEEE TMI*, 26(11):1525–1536, 2007. **1**
- [2] D. L. Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabrait. Diffusion tensor imaging: Concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:534–546, 2001. **1**
- [3] A. Brun, H. Knutsson, H. J. Park, M. E. Shenton, and C.-F. Westin. Clustering fiber tracts using normalized cuts. In *MICCAI*, 2004. **1**
- [4] V. de Silva and J. B. Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. In *NIPS*, pages 705–712, 2002. **4**
- [5] Y. Duan, X. Li, and Y. Xi. Thalamus segmentation from diffusion tensor magnetic resonance imaging. *Journal of Biomedical Imaging*, 2007(2):1–1, 2007. **1**
- [6] P. Fillard, N. Toussaint, and X. Pennec. Medinria: DT-MRI processing and visualization software. Similar Tensor Workshop, 2006. **5, 7**
- [7] C. J. Galban, S. Maderwald, K. Uffmann, A. de Greiff, and M. E. Ladd. Diffusive sensitivity to muscle architecture: a magnetic resonance diffusion tensor imaging study of the human calf. *European Journal of Applied Physiology*, 93(3):253–262, 2004. **5**
- [8] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios. Dense image registration through MRFs and efficient linear programming. *Medical Image Analysis*, 12(6):731–741, 2008. **7**
- [9] A. Goh and R. Vidal. Segmenting fiber bundles in diffusion tensor images. In *ECCV*, 2008. **1**
- [10] D. Haussler. Convolution kernels on discrete structures. Technical report, 1999. **4**
- [11] T. Jebara, R. Kondor, and A. Howard. Probability product kernels. *Journal of Machine Learning Research*, 5:819–844, 2004. **2**
- [12] L. Jonasson, P. Hagmann, C. Pollo, X. Bresson, C. Richero Wilson, R. Meuli, and J.-P. Thiran. A level set method for segmentation of the thalamus and its nuclei in DT-MRI. *Signal Processing*, 87(2):309–321, 2007. **1**
- [13] C. Lenglet, M. Rousson, and R. Deriche. DTI segmentation by statistical surface evolution. *IEEE TMI*, 25(06):685–700, 2006. **1**
- [14] M. Maddah, W. Grimson, S. Warfield, and W. Wells. A unified framework for clustering and quantitative analysis of white matter fiber tracts. *Medical Image Analysis*, 12(2):191–202, 2008. **2**
- [15] M. Maddah, A. U. J. Mewes, S. Haker, W. E. L. Grimson, and S. K. Warfield. Automated atlas-based clustering of white matter fiber tracts from DTMRI. In *MICCAI*, 2005. **2, 5**
- [16] J. Melonakos, V. Mohan, M. Niethammer, K. Smith, M. Kubicki, and A. Tannenbaum. Finsler tractography for white matter connectivity analysis of the cingulum bundle. In *MICCAI*, 2007. **1**
- [17] N. Milne. Human functional anatomy 213. <http://www.lab.anhb.uwa.edu.au/hfa213/week3/lec3unimuscle.pdf>. **6**
- [18] R. Neji, A. Besbes, N. Komodakis, J.-F. Deux, M. Maatouk, A. Rahmouni, G. Bassez, G. Fleury, and N. Paragios. Clustering of the human skeletal muscle fibers using linear programming and angular hilbertian metrics. In *IPMI*, 2009. **8**
- [19] L. O'Donnell and C.-F. Westin. White matter tract clustering and correspondence in populations. In *MICCAI*, 2005. **1, 2**
- [20] L. O'Donnell and C.-F. Westin. Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE TMI*, 26(11):1562–1575, 2007. **2**
- [21] X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, 2006. **3**
- [22] P. Savadjiev, J. S. W. Campbell, G. B. Pike, and K. Siddiqi. Streamline flows for white matter fibre pathway segmentation in diffusion MRI. In *MICCAI*, 2008. **2**
- [23] B. Scholkopf, A. Smola, and K.-R. Muller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comp.*, 10(5):1299–1319, 1998. **3**
- [24] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000. **4**
- [25] A. Tsai, C.-F. Westin, A. O. Hero, and A. S. Willsky. Fiber tract clustering on manifolds with dual rooted-graphs. In *CVPR*, 2007. **2**
- [26] Y. Tsin. *Kernel Correlation as an Affinity Measure in Point-Sampled Vision Problems*. PhD thesis, Robotics Institute, Carnegie Mellon University, 2003. **4**
- [27] V. Vapnik. *Statistical Learning Theory*. Wiley, 1998. **5**
- [28] Z. Wang and B. C. Vemuri. DTI segmentation using an information theoretic tensor dissimilarity measure. *IEEE TMI*, 24(10):1267–1277, 2005. **1**
- [29] Y. T. Weldeselasie and G. Hamarneh. DT-MRI segmentation using graph cuts. *SPIE Medical Imaging*, 2007. **1**
- [30] U. Ziyen, D. Tuch, and C.-F. Westin. Segmentation of thalamic nuclei from DTI using spectral clustering. In *MICCAI*, 2006. **1, 5, 6**