

Manifold Learning Techniques in Image Analysis of High-dimensional Diffusion Tensor Magnetic Resonance Images

Parmeshwar Khurd¹, Sajjad Baloch¹, Ruben Gur², Christos Davatzikos¹ and Ragini Verma¹

¹Section for Biomedical Image Analysis,
Dept. of Radiology, University of Pennsylvania
Philadelphia, PA 19104

²Brain Behavior Laboratory,
Dept. of Psychiatry, University of Pennsylvania
Philadelphia, PA 19104

{khurdp, balochs, davatzic, vermar}@uphs.upenn.edu; gur@bb1.med.upenn.edu

Abstract

Diffusion Tensor magnetic resonance imaging (DT-MRI) provides a comprehensive characterization of white matter (WM) in the brain and therefore, plays a crucial role in the investigation of diseases in which WM is suspected to be compromised such as multiple sclerosis and neuropsychiatric disorders like schizophrenia. However changes induced by pathology may be subtle and affected regions of the brain can only be revealed by a group-based analysis of patients in comparison with healthy controls. This in turn requires voxel-based statistical analysis of spatially normalized brain DT images, as in the case of conventional MR images. However this process is rendered extremely challenging in DT-MRI due to the high dimensionality of the data and its inherent non-linearity that causes linear component analysis methods to be inapplicable. We therefore propose a novel framework for the statistical analysis of DT-MRI data using manifold-based techniques such as isomap and kernel PCA that determine the underlying manifold structure of the data, embed it to a manifold and help perform high dimensional statistics on the manifold to determine regions of difference between the groups of patients and controls. The framework has been successfully applied to DT-MRI data from patients with schizophrenia, as well as to study developmental changes in small animals, both of which identify regional changes, indicating the need for manifold-based methods for the statistical analysis of DTI.

1. Introduction

Diffusion Tensor Imaging (DTI) [1] has gained wide acceptance as an MR modality that provides a non-invasive characterization of white matter (WM) micro-structural integrity [2], and anatomical connectivity [3], by quantifying the magnitude and directionality of microscopic water diffusion in the WM, and is being increasingly used in the investigation of WM diseases, especially in the study of neuropsychiatric disorders such as schizophrenia [4], where

conventional radiological evaluations fail to detect substantial WM differences. The growing clinical importance of DTI in disease investigation has prompted large population longitudinal and cross-sectional studies of WM changes in the brain, which can lead to early diagnosis of disease, to a more effective monitoring of treatment or to study subtle developmental changes. Group-based analysis entails statistical analysis of brains of subjects that have been spatially normalized to a template. This has therefore generated a need for sophisticated techniques for statistically analyzing spatially normalized DTI data, in order to identify and quantify complex patterns of structural changes associated with inter-individual variability and those induced by pathology or developmental changes. While such techniques have been successfully developed for analysis of conventional structural MR images, their development for DTI data is challenging as there is a tensor associated with each voxel with the data. The high dimensionality and complex non-linear underlying structure makes the commonly used voxel-wise linear statistics [5] or principal components analysis [6], used for other modalities which have a scalar value associated with each voxel, inapplicable to DTI. The statistics of tensors need to be performed along a linear manifold that appropriately parameterizes the non-linear space of the tensors under consideration at a voxel. Such faithful linear manifold representations may be identified using techniques such as Isomap [7], that preserve geodesic distances between tensors, or by using kernel-based techniques such as kernel PCA [6].

Existing work on statistics of DT-MRI either does not incorporate the underlying manifold structure [8], or does not address the issue of faithful linear representations critical to analysis of non-linear tensor data. In order to circumvent the lack of statistical methods for tensorial data, current day studies perform statistical analyses on scalar (fractional anisotropy and diffusivity [1, 2]) and vector (orientation of tensor [9]) measures, computed from the tensor data. However statistical analysis of these simpler measures requires the prior knowledge as to which of these measures is affected by disease, hence it limits their general

applicability. Recognizing the shortcomings of the existing DTI analysis approaches, in this paper we try to address and alleviate these issues by developing a paradigm for a full voxel-based analysis of the tensor data. Following spatial normalization of the DTI data, the crux of our framework lies in our novel method for voxel-based statistical analysis of tensors that learns the underlying structure of the tensor data using manifold learning methods (isomap [7], extended isomap [10], kernel PCA [6] and kernel Fisher Discriminants (FD) [6]) and

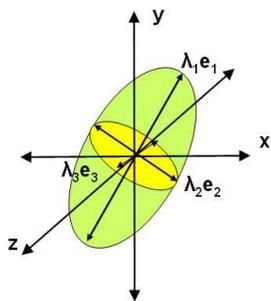


Fig. 1: Ellipsoidal representation of a tensor where eigen vectors represent the orientation of the axes and eigen values their magnitudes

embeds it into a linear manifold on which linear statistics can be applied to identify regions of difference in group-based analysis. The application of our framework to a population of schizophrenia patients has demonstrated its ability to identify regions of change, while the specificity of changes depends on the type of manifold identification method used. The framework is also used to determine regions of subtle developmental changes in mouse brains. Hence it is demonstrated that different manifold learning methods cater to different applications in which DTI data has been used.

2. Manifold Structure of Tensors

2.1 Mathematical model of DTI

In conventional DTI [1, 2], each voxel is associated with a measure of diffusivity that is mathematically represented by a 3×3 symmetric positive definite matrix $D = (d_{ij})$, $i, j = 1, 2, 3$, $D = D^T$. The eigen decomposition of D yields $D = \lambda_1 e_1 e_1^t + \lambda_2 e_2 e_2^t + \lambda_3 e_3 e_3^t$, where $\lambda_i, i = 1, 2, 3$ are the eigen values that are positive and $e_i, i = 1, 2, 3$ are the respective eigen vectors. Physically, the tensor is represented by an ellipsoid whose axes and their magnitudes represent the directions and magnitudes of water diffusivity respectively (see Fig.1). The positive definiteness of the matrix characterizes positiveness of the diffusivity.

Impact of structure of DTI data on analysis: As each voxel in the DTI data is associated with a tensor (instead of a scalar value in structural MR images), these images

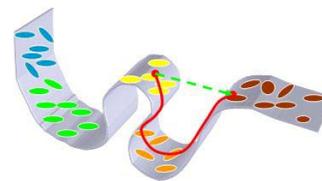


Fig. 2: Manifold structure of tensors. The black surface represents the non-linear manifold fitted through the tensors represented as ellipses. The green line represents the Euclidean distance between tensors treated as elements of R^6 and the red line represents the geodesic distance along the manifold that will be used for all tensor manipulations.

require specialized tensor mathematics for operations involving combination of tensors and performing statistics on them. Although tensor matrices have 6 variable elements $[d_{11} \ d_{22} \ d_{33} \ d_{12} \ d_{13} \ d_{23}]$, tensors cannot be

treated as elements in R^6 because the symmetry and positive definiteness and the structural property of anisotropy, constrain the tensor datasets to lie on a non-linear sub-manifold of R^6 . Therefore any analysis pertaining to tensors needs to identify the underlying manifold structure (as seen in Figure 2). We will approach this problem of manifold structure identification in DTI in two ways: **1) Geodesic-based:** This approach involves identification of the manifold and defining a geodesic distance on this manifold. Fig. 2 shows tensors on their underlying non-linear manifold. The green dotted line shows the Euclidean distance between tensors treated as elements of R^6 , that does not lie along the manifold, while the red line depicts the geodesic distance along the manifold. Mathematical operations on the tensors are performed using geodesic distances. There are several manifold learning techniques that can be adopted for determining the structure of the underlying manifold [7, 11] and defining a graph-based geodesic distances on these. We use Isomap and extended Isomap [10] for our purposes; and **2) Kernel-based:** In this approach, a carefully selected kernel is used to obtain a rich linear representation of the tensor in a high-dimensional space. Geodesic distances are not used. A potential advantage of the kernel-based approach is that separability using linear statistics can be more easily identified in the higher-dimensional space. Another potential advantage is that it is straightforward to incorporate neighborhood information from tensors belonging to a spatial neighborhood around a voxel into kernel-based approaches, whereas the estimation of geodesic distances in the neighborhood-based data space can be less reliable. We use the kernel PCA [6] and kernel FDA [6] methods in this category. Details of all these

methods will be provided in section 2.2. While one of the methods (geodesic or kernel based) may be better for datasets that require a more sensitive identification of regions as in the group-based analysis of non-focal diseases such as schizophrenia, the other may be preferable for detecting specific changes as observed in growth.

2.2 Voxel-based analysis using Manifold Learning

Let $\text{Image}(i)$, $i = 1, \dots, N$ be the datasets (images) that need to be statistically analyzed. Of these, N_1 are of group 1 and N_2 are group 2 ($N_1 + N_2 = N$). Depending on the application the two groups could be groups of healthy and diseased subjects or groups based on age (or any other parameter that needs to be tested for group differences). These datasets have been spatially normalized to the template [12]. Then a set has been created at each voxel comprising of all the tensors at that voxel across all subjects, $\{D_i\}, D_i \in \text{Image}\{i\}, i = 1 \text{ to } N$.

We then apply different manifold learning techniques to this dataset to embed this to a manifold, by learning the underlying structure of a set of tensors. Then geodesic-based or kernel-based techniques are applied to capture the manifold structure on a voxel-by-voxel basis, collected across the whole population, and the tensors are flattened to a lower dimensional space. Multivariate statistics such as Hotelling's T^2 -test are then applied to this embedded dataset at each voxel of the brain, to determine regions of significant difference between the two groups. It may be noted that Hotelling's test cannot be applied to non-linear high dimensional data, and hence the data was embedded to a linear sub-manifold prior to the application of the statistical test. This process is repeated on each voxel to obtain a voxel-wise map of differences between the two groups. We now explain the application of each of the methods to the DTI dataset. We shall first present the main ideas behind the geodesic-based and kernel-based manifold learning approaches in Sec. 2.2.1 and Sec. 2.2.2 respectively and then discuss how the performance of these two approaches can be optimized to identify group separability in Sec. 2.2.3.

2.2.1 Geodesic-based analysis of DTI data using Isomap

We determine the underlying manifold structure using Isomap in the following way:

1. Construction of a tensor neighborhood graph: This determines which points from the tensor dataset are neighbors, using a tensor-based metric [8] and creates a graph with the tensors as nodes. The application of tensor metrics instead of Euclidean metrics in determining the geodesic distance, distinguishes these tensor isomaps from traditional isomaps.

2. Defining the geodesic distance and creation of Distance Matrix: The geodesic distance between two faraway tensors on the graph is determined as the graph based shortest

distance between nodes, which is computed using the Floyd-Warshall algorithm [13] and is used to create the distance matrix relating all tensors.

3. Fitting the manifold: Multidimensional scaling (MDS) [11] is used to extract the manifold fitted to the tensors. In the eigen decomposition of the distance matrix, let $\lambda_i, i = 1, \dots, N$ be the eigen values of the $N \times N$ distance matrix in descending order of magnitude and $e_i, i = 1, \dots, N$ be the corresponding eigen vectors. The dimension d of the underlying manifold is equal to the rank of the centered distance matrix. The top d eigen vectors are then chosen to obtain the representation of these tensors in the d -dimensional lower dimensional space. Let λ_p be the p -th eigen value (in decreasing order) of the distance matrix and let e_p^i be the i -th component of the p -th eigen vector. Then by setting the p -th component of the d -dimensional coordinate vector \mathcal{Y}^i equal to $\sqrt{\lambda_p} e_p^i$, we obtain the lower dimensional representation. Then the residual variance or the fraction of the "unexplained residuals" as captured by a set of d eigen vectors is defined as:
$$RV(d) = \frac{\sum_{i=d+1}^N \lambda_i}{\sum_{i=1}^N \lambda_i}.$$

This provides a measure of the information captured by the data.

The fitted manifold retains the underlying structure of the data. More mathematical and computational details of the isomap algorithm can be found in [11, 14].

2.2.2 Kernel-based analysis of DTI data using kernel PCA

Kernel-based techniques are alternative ways to learn the linear manifold representations of the non-linear tensors under consideration. Rather than using the notion of geodesic distances to obtain linear representations, they use a carefully selected kernel (typically, a Gaussian radial basis function with an appropriate kernel width parameter) to obtain a rich linear representation of the tensors in a high-dimensional reproducing kernel Hilbert space (RKHS). These linear representations are found by computing projections onto appropriate RKHS directions using the "kernel trick" [6].

In conventional PCA, we find principal directions in the vector space of the samples that maximize the variance of the projections of the samples along those directions and which also minimize the least-squares representation error for the samples. In kernel PCA, we find similar principal eigen directions in the higher-dimensional RKHS by solving an appropriate eigenvalue problem as described below. Let x_i denote the R^6 representation of D_i (x_i could also be a vector formed using the tensors in a spatial neighborhood

around a voxel) and let $\Phi(x_i)$ denote its nonlinear mapping into the RKHS. Let us denote the underlying kernel by $k(\cdot, \cdot)$, where $k(x_i, x_j)$ equals the inner product $\langle \Phi(x_i), \Phi(x_j) \rangle$. Since a principal eigenvector v in the higher-dimensional Hilbert space lies in the span of the vectors $(\Phi(x_i) - \bar{\Phi})$ (where the bar super-script denotes the mean operation), it can be conveniently represented as $v = \sum_i \alpha_i (\Phi(x_i) - \bar{\Phi})$, where α is an N-dimensional vector. The entire kernel PCA procedure is summarized below [11]:

1. Form the kernel matrix K , where $K_{ij} = k(x_i, x_j), i = 1, \dots, N, j = 1, \dots, N$
2. Center the kernel matrix to obtain K_C .
3. Eigen-decompose K_C to obtain its eigenvectors α_i and eigenvalues $\lambda_i, i = 1, \dots, N$.
4. Normalize the eigenvectors α^i to have length $1/\sqrt{\lambda_i}$ so that the eigenvectors v^i in the RKHS have unit length.
5. The i th kPCA component for training sample x_k is given by: $\langle \Phi(x_k) - \bar{\Phi}, v^i \rangle = \lambda_i \alpha_k^i$
6. For a general test point x , the i th kPCA component is: $\langle \Phi(x_k) - \bar{\Phi}, v^i \rangle = \sum_m \alpha_m^i k(x, x_m) - \frac{1}{N} \sum_{m,n} \alpha_m^i k(x, x_n)$

2.2.3. Incorporation of separability criteria into manifold learning approaches

Often the goal is not merely to obtain a linear representation of the data, but also to obtain a representation that can discriminate between the tensors belonging to different classes in a statistically optimal manner. It is difficult to *a priori* predict the discriminatory component(s) in the Isomap-based and the kernel PCA-based approaches described earlier and using non-discriminatory components can reduce the statistical significance of the subsequent separability tests. The kernel Fisher discriminant analysis (FDA) technique incorporates separability criteria into kernel-based manifold learning approaches in a principled manner. Kernel FDA finds a direction in the higher-dimensional RKHS so that a projection along this direction maximizes a separability measure known as the Rayleigh coefficient (or the Fisher discriminant ratio). This direction can be found by solving an appropriate linear algebraic problem [6], as in the case of kernel PCA. Geodesic-based approaches can also be extended to incorporate separability criteria, albeit in a less principled manner, as in the extended Isomap method [10]. In this

method, the geodesic distances between D_i are themselves used as features and a linear discriminant maximizing the Rayleigh coefficient in this feature space is computed. A disadvantage of the kernel FDA and the extended Isomap methods is that they require cross-validation using computationally expensive permutation tests since they use the class labels while computing the p-value maps.

In summary, our framework for the DTI analysis of large population studies consists of three steps: 1) spatial normalization of the DTI to a template, followed by 2) application of geodesic-based/kernel-based analysis to tensors voxel-wise across the whole population, and 3) applying Hotelling's T-test to these linearized datasets to identify regions of significant difference. In the next section, we apply our framework to study a dataset of schizophrenia patients and matched healthy controls, as well as in small animal imaging studies to determine difference in growth patterns.

3. Experimental results and discussion

In this section, we present the results of applying both the Isomap-based and kernel-based framework for the analysis of DTI data. We apply the methods to a dataset of human brains to determine regions of change between patients with schizophrenia and healthy controls, where differences are subtle and non-focal, that is, they are distributed over the whole brain. The second dataset consists of developing mouse brains. Manifold-based analysis is applied to it to determine regions of change as a result of maturation and brain development. The experiments aim at demonstrating the applicability of our manifold-based framework for group-based statistical analysis of high dimensional DTI data. We will apply both the isomap-based and kernel-based methods to these datasets to determine regions of difference and indicate the applicability of these datasets in each scenario.

3.1 Manifold-based group analysis of a dataset of patients with schizophrenia

DTI-based research in schizophrenia [4] has revolved around studying WM changes and possible abnormalities in inter-hemispheric connectivity through the corpus callosum and intra-hemispheric connectivity between frontal, temporal and occipital lobes via association fibers. Study is mainly based on anisotropy and diffusivity values of specific regions under study or voxel based analysis of scalar maps. We will determine regional changes in a group of schizophrenia patients and controls using our manifold learning techniques.

The dataset consists of 34 patients (21 male and 13 female) and 36 healthy controls (17 male and 19 female). The

controls are matched to the patients by age, sex and ethnicity. We applied our isomap-based voxel-wise analysis framework (as described in section 2.2.1) to the spatially normalized DTI datasets [12]. Subsequently on applying Hotelling's T^2 -test we obtain a p-map of the levels of change. Fig. 3 shows regions with $p < 0.1$ indicating trends of change. We see changes in the splenium of the corpus callosum (Fig. 3(a), (b)) and the inferior fronto-occipital fasciculus (ifo) (Fig. 3(c)) when all the patients and controls are analyzed, combined with respect to gender. In addition, differences are seen in the cingulum (Fig. 3(d)), the interior capsule (Fig. 3(a)) and in the temporal lobe (Fig. 3 (b)) and the insular cortex (Fig. 3(a)). The same analysis can be carried out on groups divided on the basis of gender. The application of the manifold-based method for tensor analysis produces a comprehensive map of change, both in the white matter and the gray matter. Since the regions identified compares with the regions hypothesized to change in schizophrenia [4], it establishes the significance of the full

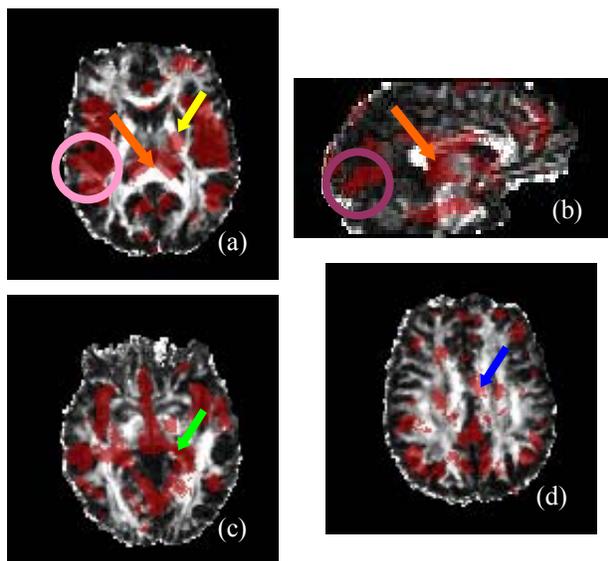


Fig. 3: Isomap-based analysis of the schizophrenia dataset: P-values < 0.1 are overlaid on the template FA map. Regions of difference: Corpus callosum (red arrows), internal capsule (yellow arrow), fronto-occipital fasciculus (green arrow), cingulum (blue arrow), temporal lobe (purple circle) and insula (pink circle).

tensor manifold-based analysis.

The same datasets were analyzed using voxel-wise kernel PCA (section 2.2.2), extended Isomap (section 2.2.3) and kernel FDA (section 2.2.3). The results are shown in Figure 4. The kernel PCA-based p-value map (Fig. 4(a)) exhibits regions of differences very similar to the Isomap-based p-value map (Fig. 3), but these regions are somewhat larger and smoother. P-value maps using the discriminant-based methods, kernel FDA and extended Isomap, are shown in

Fig. 4(b) and Fig. 4(c), respectively, with both methods used without any form of cross-validation. Since the discriminant analysis methods use the class labels and cross-validation was not performed, a lower threshold has been used for these methods while determining regions with significant differences. At this lower threshold, the extended Isomap method found fewer regions with significant differences than kernel PCA or Isomap, but kernel FDA found larger significant differences than the other three methods, indicating that it is the more sensitive method for this dataset.

3.2 Manifold-based group analysis of developing mouse brains

The dataset consists of 30 brains (18 young (group 1) and 12 old (group 2)). The hypothesis was that there were differences induced by maturation of white matter and gray matter that will be demonstrated by a group-based

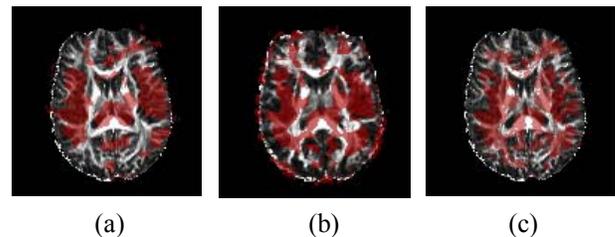


Fig. 4: Schizophrenia dataset: Axial slices of (a) Kernel PCA p-value map thresholded at 0.1. (b) Kernel FDA p-value map thresholded at 0.01 (c) Extended ISOMAP-based p-value map thresholded at 0.01. Lower thresholds have been used for the discriminant-based methods.

analysis. An Isomap analysis, similar to the schizophrenia case, was performed in order to identify differences between young and old mice brains. The resulting p-value map is shown in Fig. 5. Changes were identified in the cerebellum, the cortex, corpus callosum (cc), the hippocampus (hpc) and the internal (ic) and external capsules (ec). This indicates that the white matter tracts (ic, ec and cc) showed differences in anisotropy and became more oriented, while the gray matter structures (hpc, cortex) lost some of their orientation with development.

Results of applying the kernelized methods and the extended Isomap method to the mouse datasets are shown in Figure 6. Figure 6(a) shows an axial slice of a kernel PCA-based p-value map. A comparison with the corresponding slice in Fig. 5 reveals that similar group differences were identified in the kernel PCA-based method and the geodesic-based method. As in the case of the schizophrenia data, the p-value maps when evaluated using the kernel FDA-based method and the extended Isomap-based method, a lower threshold

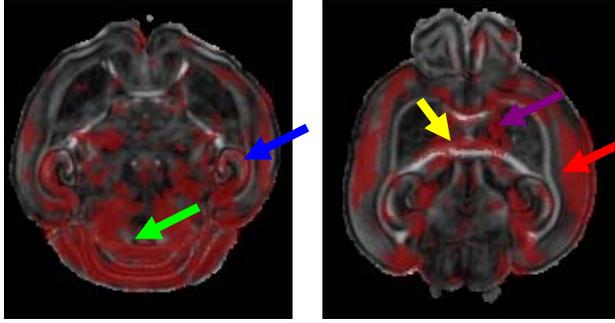


Fig. 5: Isomap- based analysis applied to the mouse data set. P-value map thresholded at 0.005 is overlaid on the template FA map. Regions of large change are the cortex (red arrow), hippocampus (blue arrow), corpus callosum (yellow arrow), internal capsule (purple arrow) and cerebellum (green arrow).

was used to determine significance (shown in Figs. 6(b)-(d)) since no cross-validation was used.

3.3 Discussion

It is somewhat difficult to compare group analyses techniques on the schizophrenia and mouse datasets since the ground truth is not known. We are currently working on a rigorous comparison of the different group analysis techniques on simulated DTI datasets with known ground truth. In future work, we plan to use computationally expensive permutation tests to evaluate more meaningful, cross-validated p-value maps, corresponding to the results in Figs. 4(b)-(c) and Figs. 6(b)-(d).

Nevertheless, an application of the geodesic-based and kernel-based methods to the above datasets enhanced the understanding about these methods in several ways. Comparing Figures 3 and 4, we observe that more regions are identified by the kernelized methods, indicating that these are more sensitive to subtle changes. This makes these methods ideal for group-wise analysis of diseases like schizophrenia. This was also observed in some other studies [15], albeit on simulated datasets [16]. A comparison of the application of these methods to the mouse dataset (Figures 5 and 6) showed that the regions identified by both types of methods were almost the same. This indicates that the geodesic-based methods are better applicable to studies where changes are more global and easily observable (unlike small subtle changes in schizophrenia).

4. Conclusions and Future Work

The need for methods of tensorial statistics has been established by the fact that the full impact of pathology or growth on tensors cannot be characterized just by analyzing the simple scalar and vector measures derived from the

tensors. We have presented a methodology for statistical analysis of tensorial image data, by estimating the structure of the nonlinear manifold on which these data lie, thereby enabling the proper definition of statistical measures. In view of the broad range of research and clinical applications of DT-MRI, we anticipate that the proposed methods will open exciting new avenues for exploring the full richness of diffusion tensor measurements, and for understanding the ways in which such measurements are affected by genotype,

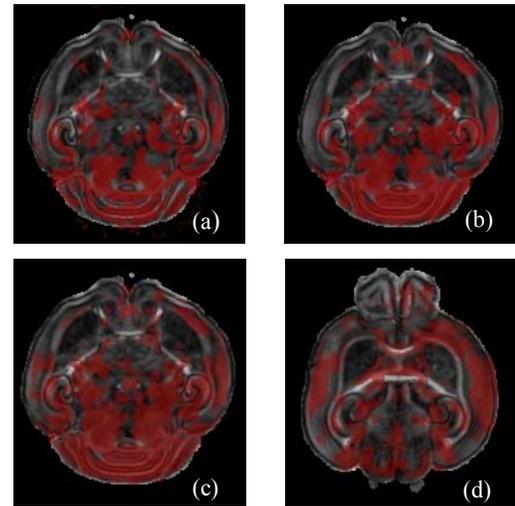


Fig. 6: Mouse dataset (a) Axial KPCA p-value map thresholded at 0.005. (b) Axial slice of KFDA p-value map thresholded at 0.0001. (c)-(d) Axial slices of extended ISOMAP-based p-value maps thresholded at 0.0001. Lower thresholds have been used for the discriminant-based methods.

pathologies and treatments. In particular, it can be used to analyze pathological differences in disease and localize regions of maximum effect as has been demonstrated in the schizophrenia dataset. It can also be used to investigate new avenues in probabilistic atlasing of brains, as statistical measures defined along the manifold provide a biologically and mathematically consistent way to summarize anatomical variability within groups of patients or research subjects, as has been shown for the mouse datasets.

In summary, we expect that a comprehensive examination of any disease or developmental changes using the DTI processing pipeline that we propose, will elucidate subtle regional changes and subtle disruptions of connectivity. We expect that in long-term, these tools will be used for prognosis and for studying subtle temporal white matter changes which may be an indicator of pathology. In the future, we will use the embedded lower dimensional tensors to generate features that can be used for region-based full classification of these datasets.

5. Acknowledgements

This work was supported by the National Institute of Health via grants R01MH070365, R01MH079938 and R01MH060722.

6. References

- [1] D. Le Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat, "Diffusion tensor imaging: concepts and applications,," *Journal of Magnetic Resonance Imaging*, vol. 13, pp. 534-546, 2001.
- [2] P. J. Basser and C. Pierpaoli, "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI," *Journal of Magnetic Resonance, Series B*, vol. 111, pp. 209-219, 1996.
- [3] S. Wakana, H. Jiang, L. M. Nagae-Poetscher, P. C. M. van Zijl, and S. Mori, "Fiber Tract-based Atlas of Human White Matter Anatomy," *Radiology*, vol. 230, pp. 77-87, 2004.
- [4] M. Kubicki, R. W. McCarley, C. F. Westin, H. J. Park, S. Maier, R. Kikinis, F. A. Jolesz, and M. E. Shenton, "A review of diffusion tensor imaging studies in schizophrenia," *Journal of Psychiatric Research*, vol. 41, pp. 15-30, 2007.
- [5] K. J. Friston, A. P. Holmes, K. J. Worsley, J. B. Poline, C. D. Frith, and R. S. J. Frackowiak, "Statistical Parametric Maps in functional imaging: A general linear approach," *Human Brain Mapping*, vol. 2, pp. 189-210, 1995.
- [6] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond (Adaptive Computation and Machine Learning)*, 1st edition (December 15, 2001) ed: The MIT Press, 2001.
- [7] J. B. Tenenbaum, V. d. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, pp. 2319-2323, 2000.
- [8] P. T. Fletcher and S. Joshi, "Principal Geodesic Analysis on Symmetric Spaces: Statistics of Diffusion Tensors," presented at Computer Vision Approaches to Medical Image Analysis, 2004.
- [9] Y.-C. Wu, A. S. Field, M. K. Chung, B. Badie, and A. L. Alexander, "Quantitative Analysis of Diffusion Tensor Orientation:Theoretical Framework," *Magnetic Resonance in Medicine*, vol. 52, pp. 1146 -1155, 2004.
- [10] M. Yang, "Extended Isomap for Classification," presented at International Conf. Pattern Recognition, 2002.
- [11] C. J. C. Burges, "Geometric Methods for Feature Extraction and Dimensional Reduction," in *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers'*, O. Maimon, Ed.: Kluwer Academic Publishers, 2005.
- [12] D. Xu, S. Mori, D. Shen, P. C. M. v. Zijl, and C. Davatzikos, "Spatial Normalization of Diffusion Tensor Fields," *Magnetic Resonance in Medicine*, vol. 50, pp. 175-182, 2003.
- [13] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*: MIT Press, 1990.
- [14] J. B. Tenenbaum, V. d. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, pp. 2319-2323, 2000.
- [15] P. Khurd, R. Verma, and C. Davatzikos, "On Characterizing and Analyzing Diffusion Tensor Images by Learning their

Underlying Manifold Structure," presented at IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis, New York, NY, 2006.

[16] P. Khurd, R. Verma, and C. Davatzikos, "Kernel-based Manifold Learning for Statistical Analysis of Diffusion Tensor Images," presented at Information Processing in Medical Imaging, pp. To Appear, 2007.