MULTIVARIATE HYPOTHESIS TESTING OF DTI DATA FOR TISSUE CLUSTERING

Raisa Z. Freidlin[1,2], Yaniv Assaf[3], Peter J Basser[4]

[1]TAIS / DCB / CIT / NIH, Bethesda, MD, USA.
[2]Dept. of Electrical and Computer Engineering, George Washington University, Washington, DC, USA.
[3]Dept. of Neurobiochemistry, Tel Aviv University, Ramat Aviv, Israel.
[4]STBB / LIMB / NICHD / NIH, Bethesda, MD, USA.

ABSTRACT

In this work we investigate the feasibility and effectiveness of unsupervised tissue clustering and classification algorithms for DTI data. Tissue clustering and classification are among the most challenging tasks in DT image analysis. While clustering separates acquired data into objects, tissue classification provides in-depth information about each region of interest.

The unsupervised clustering algorithm utilizes a framework proposed by Hext and Snedecor, where the null hypothesis of diffusion tensors arising from the same distribution is determined by an *F*-test. Tissue type is classified according to one of three possible diffusion models (general anisotropic, prolate, or oblate), which is determined with a parsimonious model selection framework. This approach, also adapted from Snedecor, chooses among different models of diffusion within a voxel using a series of *F*-tests.

Both numerical phantoms and DWI data obtained from excised rat spinal cord are used to test and validate these tissue clustering and classification approaches.

Keywords: DTI, diffusion tensor, parsimonious, clustering, hypothesis testing

1. INTRODUCTION

Diffusion Tensor Magnetic Resonance Imaging[1] (DT-MRI or MRI) provides noninvasive quantitative measurements of the apparent diffusion tensor of water molecules in tissue. In an anistropic medium, the signal attenuation in diffusion-weighted images (DWI) depends on the underlying tissue structures. Most work in DTI segmentation is based on applying thresholding criteria to tensor-derived scalar quantities, such as the Trace of the diffusion tensor (*Tr*), the Fractional Anisotropy (*FA*), and the Relative Anisitropy (*RA*). However, these scalars are generally subject to bias usually due to background noise [2, 3]. Also, they do embody for all the information available in 3×3 diffusion tensor. To reliably identify regions of interest (ROI), we propose using information contained in the entire diffusion tensor to perform clustering. In this work we propose the use of statistical hypothesis testing adapted from Hext[4] and Snedecor[5, 1] to perform unsupervised tissue classification. This method uses the *F*-test for assessing similarities between tensors in different voxels. To mitigate the effect of variability between voxels with different *FA*, we choose uniform seed voxels that have been preselected according to their local diffusion properties as determined by a hierarchical model selection framework, which is also based on *F*-tests.

One advantage of using statistical hypothesis testing lies in performing tests on the entire diffusion tensor, which containes information about *Tr*, *FA* and diffusion orientation. Another advantage is that one can assess errors in ROI selection and choose the confidence levels for each test. Finally, these tests are rapid and easy to perform voxel-by-voxel, even for large DTI data sets.

2. THEORY

2.1. Diffusion Tensor Imaging

The relationship between observed echo attenuation [6, 7], caused by applying diffusion sensitizing gradients along various directions, and the diffusion tensor, $\mathbf{D}$, can be characterized by

$$S(\mathbf{G}) = S(0)e^{-\mathbf{b}:\mathbf{D}}, \quad (1)$$

where $S(\mathbf{G})$ is observed signal, $S(0)$ is a signal in the absence of the diffusion-weighted gradient, $\mathbf{D}$ is a symmetric (3x3) 2nd-order diffusion tensor and $\mathbf{b}$ is a matrix computed by:

$$b_{ij} = \gamma^2 G_i G_j \delta^2 \left[\Delta - \frac{\delta}{3} \right], \quad (2)$$

where G_i is the diffusion gradient applied in i^{th} direction with duration δ, and Δ is the diffusion time.

2.2. Parameter estimation framework

Diffusion tensor estimation is performed using a non-linear least square minimization method, proposed by Koay et al.[8], applied to the function Eq. 1. The design matrix, $\mathbf{B}$, consists

of a list of **b**-matrix elements for a series of n trials or DWI acquisitions:

$$\mathbf{B} = \begin{bmatrix} b_{x_1}^2 & b_{y_1}^2 & b_{z_1}^2 & 2b_{x_1y_1} & 2b_{x_1z_1} & 2b_{y_1z_1} & -1 \\ b_{x_2}^2 & b_{y_2}^2 & b_{z_2}^2 & 2b_{x_2y_2} & 2b_{x_2z_2} & 2b_{y_2z_2} & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{x_n}^2 & b_{y_n}^2 & b_{z_n}^2 & 2b_{x_ny_n} & 2b_{x_nz_n} & 2b_{y_nz_n} & -1 \end{bmatrix} . \quad (3)$$

Below, $\hat{\mathbf{D}}$ is the estimated diffusion tensor for the general anisotropic model written as a (7×1) column vector. Here we estimate six independent parameters of $\mathbf{D}$ in Eq. 1 and the log of the signal in the absence of the diffusion-weighted gradient, $log[S(0)]$:

$$\hat{\mathbf{D}} = [D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}, log[S(0)]]', \quad (4)$$

We minimize the Residual Sum of Squares (*RSS*) with respect to the 7 free parameters of the diffusion model:

$$RSS = \sum_{i=1}^{n} \left(S_i(\mathbf{G}) - e^{-\mathbf{B}_i\hat{\mathbf{D}}} \right)^2, \quad (5)$$

where $S_i(\mathbf{G})$ and $e^{-\mathbf{B}_i\hat{\mathbf{D}}}$ are the observed and estimated signals, respectively, n is the number of data points in each voxel.

3. CLUSTERING BASED ON PARAMETER DISTRIBUTION OF DIFFUSION TENSORS

Parsimonious model selection methods are used as a pre processing for segmenting voxels based on diffusion properties within the voxels, however these methods do not provide any information about the homogeneity of the tissue, i.e., whether diffusion tensors within a given ROI have the same model type, and if so, whether their parameters are similar to those of their neighboring voxels. Such information can further improve tissue segmentation.

In order to justify the use of the *F*-test hypothesis testing framework for tissue segmentation, the assumptions of the normally distributed residuals and homoscedasticity (i.e., uniformity of the variance within an ROI) have to be satisfied. It has been shown that the residuals are asymptotically normally distributed at *SNR* greater than 7 in an experiment otherwise free of systematic artifacts. However, the variance may not be homogeneous among neighboring voxels. To overcome this problem, we select locally homogeneous regions in the model map having the same model type as determined by the parsimonious model selection method.

The parsimonious model selection framework proposed previously selects the one of four models that provides the best fit to the DWI data using the fewest parameters. In this work we are interested in segmenting white matter regions, i.e., presumed to have their fractional anisotropy, *FA*, greater than 0.5. Thus the hierarchy of models from which we choose

the seed points excludes the isotropic model. Remaining models describe transverse isotropy (prolate and oblate ellipsoids) and full anisotropy.

3.1. Test between voxels

Once the diffusion model is chosen in each voxel, we select the seed region from 6 to 9 neighboring voxels to performt tissue clustering based on the diffusion model type of the seed voxels. The null hypothesis assumes that the difference between diffusion tensors for m voxels of the same model type is statistically insignificant. To test this hypothesis, we take following steps, adapted from Hext:

1. Combine m sets of acquired signals, S_{CAS}, into an $[n \cdot m \times 1]$ array, where n is the number of experimental data points in each voxel;

2. Combine m sets of n individually estimated signals, S_{CES} ($[n \cdot m \times 1]$);

3. Estimate the average diffusion tensor for m voxels, $\hat{\mathbf{D}}_{Avg}$ by a non-linear least square minimization method, applied to *RSS* (Eq.5), using the combined acquired signal vector, S_{CAS}, and the augmented $[n \cdot m \times 7]$ design matrix, $\mathbf{B}_C$;

4. Estimate the average signal vector, S_{Avg}, $[n \cdot m \times 1]$ using $S_{Avg}(\mathbf{G}) = S(0)e^{-\mathbf{B}_C\hat{\mathbf{D}}_{Avg}}$.

5. Apply the *F*-test of the null hypothesis.

We use Snedecor's *F*-test[1] to assess the similarity among variances within the voxels:

$$F_0 = \frac{(RSS_{Avg} - RSS_{CES})/(fp \cdot (m-1))}{(RSS_{CES})/(m \cdot (n-fp))} \quad (6)$$

where $fp = 7$ is the number of the free parameters in the general anisotropic model, m is a number of voxels with n experimental data points each; RSS_{CES} and RSS_{Avg} are the residual sums of squares of the combined estimated signals and the results from the fit for the estimated average diffusion tensor, $\hat{\mathbf{D}}_{Avg}$, respectively.

4. METHODS

4.1. Simulations

To evaluate the parsimonious model selection approaches, synthetic phantoms were generated in MATLAB (The MathWorks, Inc.) by varying the fractional anisotropy, *FA*, from 0.2 to 1.0 with a step size of 0.1, and signal to noise, *SNR*, from 5 to 23, for a fixed signal intensity, (I_0=1000). The trace of **D**,

---

[1]A typo appears in "Statistical Methods" by Snedecor and Cohran[5] in the formula given on page 344 describing the *F*-test for comparing two nested models. The corrected formula is given in Eq. 6 above

Tr, was set to $2100e^{-6}$ mm^2/sec and $2400e^{-6}$ mm^2/sec for white and gray matter respectively, which are typical values for living brain tissue[9] . Normally distributed random noise with $\sigma = 1$ and zero mean, was added to the signal intensity and the diffusion weighted images were calculated and scaled. This model assumes that noise is added to the real and imaginary channels independently, and that the MR signal is rectified. The *b*-matrix was calculated with the imaging parameters described in the Excised Rat Spinal Cord DTI Experiments subsection. The hierarchical methods for parsimonious model selection were applied to the set of 47 reconstructed diffusion-weighted images with 1 non diffusion weighted image.

4.2. Excised Rat Spinal Cord DTI Experiments

In addition to simulation, we demonstrate our results on experimental MRI data obtained from an excised rat spinal cord fixed with 4% paraformaldehyde solution. DWIs were obtained using a PGSE DWI sequence with pulse duration $\delta = 2.5$ ms, diffusion time $\Delta = 70$ ms, repetition time (TR) = 3500 ms, and echo time (TE) = 14.7 ms. Other imaging parameters were: in-plane resolution 200x200μm^2, slice thickness = 2mm, number of averages (NEX): n = 1, bandwidth = 50 kHz. Forty DWIs per slice were acquired during 28 hours of scanning. Thirty-one of these were attenuated by diffusion gradients $\mathbf{G} = (G_x, G_y, G_z)$ and 9 were not attenuated ($|\mathbf{G}| = 0$). In each direction approximate b-values was 2000 s/mm^2. At each voxel location in the raw image, the apparent diffusion tensor, $\mathbf{D}$, was calculated[1]. Tensor-derived parameters, such as the *Tr*, *FA*, principal directions and principal diffusivities, λ_1, λ_2, and λ_3 were all calculated and passed to the parsimonious model selection algorithm.

5. RESULTS

Since the residuals from the phantom and the excised rat spinal cord experiments are asymptotically normally distributed, and the variance of each measurement is unchanging (homoscedasticity), testing of one model against another, presented below, is well grounded. The confidence interval for all tests was set to 95%.

5.1. Simulations

The parsimonious model selection results obtained at *SNR*=25 and *FA*=0.6 showed 95% success for the isotropic model ($\lambda_1 = \lambda_2 = \lambda_3$), 99% in identifying the general anisotropic model ($\lambda_1 > \lambda_2 > \lambda_3$), 90% for the oblate model ($\lambda_1 = \lambda_2 > \lambda_3$) , and 93% for the prolate model ($\lambda_1 > \lambda_2 = \lambda_3$). Within regions of oblate or prolate symmetry, we created 4 distinct regions with varying degrees of prolateness and oblateness, as well as, directions of $\boldsymbol{\epsilon}_{o/p}(\theta, \varphi)$ (Eq. 7) where θ was set

to 10°, 20°, 30°, and 40°, where the normalized eigenvector parallel to the axis of symmetry for the oblate or prolate model is:

$$\boldsymbol{\epsilon}_{o/p}(\theta, \varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)'. \quad (7)$$

However, the parsimonious model selection algorithm segments tissue based only on the diffusion properties within a voxel, i.e., presence of the transverse symmetry. Thus, for example, all prolate fibers having different diffusion parameters and/or orientation would still be marked as "prolate", while the multivariate hypothesis testing based clustering algorithm succesfully separated ROIs with different degrees of prolateness and oblateness and differentiated between diffusion tensosr with at least a 10° directional difference. It also performed denoising of the data. Overall oblate and prolate model segmentation improved from 90% to 97% and from 93% to 99%, respectively.

Excised Rat Spinal Cord DTI Experiment

Fig. 1 shows (A) the T_2-weighted amplitude image and the orientationally invariant B) *FA* and C) *Tr* maps. By examining amplitude image and the maps we can only distinguish white from gray matter groups, athough the white matter itself consists of several different fiber compartments.
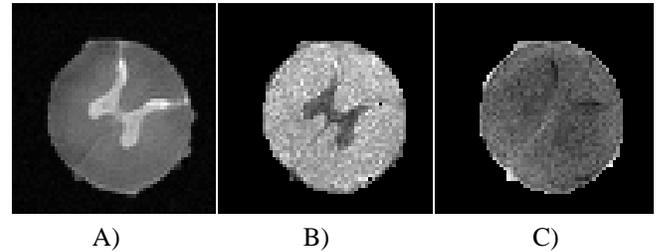

A) B) C)

Fig. 1. Excised rat spinal cord: A) the T_2-weighted amplitude image; B) the Fractional Anisotropy (*FA*) map; C) the Trace (*Tr*) map.

The parsimonious model selection method consistently segmented prolate regions in the white matter (Fig. 2A), but does not reveal different fiber organization patterns within white matter. The multivariate hypothesis testing based clustering algorithm, however, identified 7 distinct prolate regions within white matter (Fig. 2B)). The red arrow points to two non-symmetric regions, which, after closer examination of the fixed spinal cord, revealed that the fibers in these areas were compressed during specimen preparation.

6. DISCUSSION AND CONCLUSIONS

The primary goal of this work is to investigate the feasibility of using a multivariate statistical hypothesis testing framework with DTI data to perform tissue clustering and classifi-
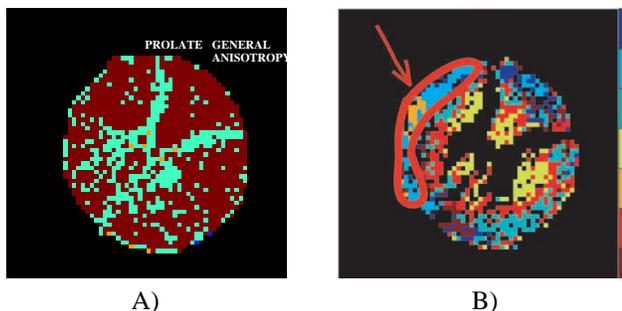
Fig. 2. Excised rat spinal cord: A) Parsimonious model selection map; B) 7 ROIs represent areas with different fiber bundles.

cation. The maps produced by the proposed multivariate hypothesis testing framework provide useful information about the distribution of different fiber types within tissues. Using numerical and spinal chord phantoms we demonstrated that the anisotropic regions with subtle differences in diffusion type (oblate, prolate or full anisotropy) and model parameters (e.g., degree of prolateness or oblateness and orientation of axis of symmetry) could be resolved. Numerical phantom results showed the ability to separate tensors with at least a $10°$ difference in the orientation of their axis of symmetry at SNR=25 and FA=0.6 or greater. Such results increase our confidence in clustering based upon statistical hypothesis tests.

F-testing for tissue clustering and classification applications is both efficient and powerful. The current approach was succesfully applied to MRI microscopy of fixed samples in which imaging artifacts can be significantly reduced and assumptions of normal residuals and uniform variance for each voxel within the DWI data can be assured. Thus, when applied to *ex vivo* tissue specimens, where background noise is the primary artifact and other systematic artifacts can be remedied, this approach should work robustly.

Provided that the conditions for normally distributed residuals and uniform variances for DWIs within each voxel are met, multivariate hypothesis testing could be used with *in vivo* data as well. In clinical applications, however, where other systematic artifacts can corrupt DWI data, this approach may be problematic. When using DWI data from living tissue, tests for Gaussianity of the distribution of residuals and a careful assessment of the degree of homoscedasticity must be performed prior to applying this segmentation approach to ensure its integrity. Our expectation is that applying model selection procedures prior to segmentation may improve automatic region of interest (ROI) delineation and classification of different tissue types in DT-MRI volume data sets.

7. REFERENCES

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

[2] H.R. Van Der Vaart, "Some results on the probability distribution of the latent roots of a symmetric matrix of continuously distributed elements, and some applications to the theory of response surface estimation," Report issued by the Institute of Statistics, University of North Carolina, 1958.

[3] C. Pierpaoli and P. J. Basser, "Toward a quantitative assessment of diffusion anisotropy.," *Magn Reson Med*, vol. 36, no. 6, pp. 893–906, Dec 1996.

[4] G. R. Hext, "The estimation of second-order tensors, with related tests and designs," *Biometrika*, vol. 50, pp. 353–357, 1963.

[5] G.W. Snedecor and W.G. Cochran, *Statistical Methods*, Blackwell Publishing, 8th edition, 1989.

[6] H.Y. Carr and E.M. Purcell, "Effects of diffusion on free precession in nuclear magnetic resonance experiments," *Phys. Rev*, vol. 94, no. 3, pp. 630–638, May 1954.

[7] E.O. Stejskal and J.E. Tanner, "Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient," *J Chem Phys*, vol. 42, no. 1, pp. 288–292, January 1966.

[8] C.G. Koay, L.C. Chang, J.D. Carew, C. Pierpaoli, and P.J. Basser, "Aunifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging.," *J Magn Reson*, vol. 182, pp. 115–125, Jul 2006.

[9] C. Pierpaoli, P. Jezzard, P. J. Basser, A. Barnett, and G. Di Chiro, "Diffusion tensor MR imaging of the human brain.," *Radiology*, vol. 201, no. 3, pp. 637–648, Dec 1996.