A new Mahalanobis distance measure for clustering of fiber tracts

C. Koay[1], C. Pierpaoli[1], and P. J. Basser[1]
[1]NIH, Bethesda, MD, United States

INTRODUCTION Data analysis in Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) is highly sophisticated and can be thought of as a "pipeline" of closely connected processing and modeling steps. Cluster analysis of the orientation of the fiber direction and fiber tracts is typically carried on the major eigenvector. This type of cluster analysis is also important in reducing sorting bias in the eigenvalues of the diffusion tensor [1-3]. In this work, we present a simple and novel generalization of Mahalanobis distance measure for the dyadics of the eigenvector for the purposes of clustering fiber tracts and fiber orientation. This approach is built upon a series of works by Koay et al.[4-6], especially the error propagation framework for DT-MRI as presented in [5]. This approach is straight forward and the Mahalanobis distance measure for the dyadics can computed efficiently without ad hoc combinatorial optimization that is typical in the eigenvector-clustering techniques, e.g., [1,2]. The proposed Mahalanobis distance measure is the ideal measure for clustering fiber tracts.

METHODS AND RESULTS The dyadics formalism is a relatively well known technique in DTI [1,7]. A dyadic tensor of a vector, $\mathbf{q} = [q_x, q_y, q_z]^T$ is simply the matrix outer product of $\mathbf{q}$, which is given by:

$$\mathbf{q}\cdot\mathbf{q}^T = \begin{pmatrix} q_x q_x & q_x q_y & q_x q_z \\ q_x q_y & q_y q_y & q_y q_z \\ q_x q_z & q_y q_z & q_z q_z \end{pmatrix} \qquad [1]$$

To construct the Mahalanobis distance measure for the dyadics, we use the recently proposed error propagation framework [5-6] to (a) compute the covariance structure for the eigenvectors, (b) construct the covariance structure for the dyadics from the covariance structure for the eigenvectors, and then (c) express the Mahalanobis distance measure for the dyadics.

Briefly and without loss of generality, let us assume that $\{{}^1\mathbf{q},\cdots,{}^N\mathbf{q}\}$ is a collection of properly oriented eigenvectors with respect to the mean eigenvector and let $\left\langle \mathbf{q}\cdot\mathbf{q}^T \right\rangle = (\sum_{k=1}^{N} {}^k\mathbf{q}\cdot{}^k\mathbf{q}^T)/N$ be the average dyadics. If we imagine that the mean eigenvector is pointing toward the positive z-axis, then an eigenvector is considered properly oriented with respect to the mean eigenvector if its z-component is positive. Let the eigenvalue decomposition of the average dyadics be $\sum_{i=1}^{3}\lambda_i \boldsymbol{\psi}_i \boldsymbol{\psi}_i^T$ where $\lambda_1 \ge \lambda_2 \ge \lambda_3$. It is shown in [5] that the sample covariance of $\{{}^1\mathbf{q},\cdots,{}^N\mathbf{q}\}$, denoted by $\boldsymbol{\Sigma}_{\mathbf{q}}$, can be expressed as:

$$\boldsymbol{\Sigma}_{\mathbf{q}} \cong \frac{N}{N-1}\left(\lambda_2\boldsymbol{\psi}_2\boldsymbol{\psi}_2^T + \lambda_3\boldsymbol{\psi}_3\boldsymbol{\psi}_3^T + (\lambda_1 + 1 - 2\|\mathbf{t}\|)\boldsymbol{\psi}_1\boldsymbol{\psi}_1^T\right) \quad , \qquad [2]$$

where $\mathbf{t} = (\sum_{i=1}^{N}\mathbf{q}_i)/N$, $\hat{\mathbf{t}} = \mathbf{t}/\|\mathbf{t}\|$, and $\hat{\mathbf{t}} \approx \boldsymbol{\psi}_1 \equiv [\psi_{1x}, \psi_{1y}, \psi_{1z}]^T$ is the mean eigenvector.

To construct the covariance matrix of the dyadics from the covariance matrix of the eigenvectors, it is convenience to work with the vectorized form of the average dyadics and of the dyadics. Let us denote the vectorized form of a dyadic tensor, $\mathbf{q}\cdot\mathbf{q}^T$, by $\tilde{\mathbf{q}} \equiv [q_x q_x, q_y q_y, q_z q_z, q_x q_y, q_y q_z, q_x q_z]^T$. This convention of vectorization is consistent with [4-6]. Let

$$\boldsymbol{\Sigma}_{\tilde{\mathbf{q}}(\boldsymbol{\psi}_1)}$$

denotes the covariance matrix of the vectorized dyadics evaluated at $\boldsymbol{\psi}_1$. We note that it can be computed through the following transformational law of error propagation [5]:

$$\boldsymbol{\Sigma}_{\tilde{\mathbf{q}}(\boldsymbol{\psi}_1)} = \mathbf{J}\,\boldsymbol{\Sigma}_{\mathbf{q}}\,\mathbf{J}^T \quad , \qquad [3]$$

where

$$[\mathbf{J}]_{ij} = \partial\tilde{\mathbf{q}}_i / \partial\mathbf{q}_j\Big|_{\mathbf{q}=\boldsymbol{\psi}_1} \quad , \qquad [4]$$

with i=1,…, 6 and j=1,2,3. Furthermore, it can be shown that

$$\mathbf{J}^T = \begin{pmatrix} 2\psi_{1x} & 0 & 0 & \psi_{1y} & 0 & \psi_{1z} \\ 0 & 2\psi_{1y} & 0 & \psi_{1x} & \psi_{1z} & 0 \\ 0 & 0 & 2\psi_{1z} & 0 & \psi_{1y} & \psi_{1x} \end{pmatrix} . \qquad [5]$$

It is clear that once the covariance matrix for the vectorized dyadics is known and the dyadic tensor of the mean eigenvector is available, we can construct the Mahalanobis distance measure, ρ, between a vectorized dyadic tensor, $\tilde{\mathbf{x}}$, and the vectorized mean eigenvector, which is given as follows:

$$\rho(\tilde{\mathbf{x}} \mid \tilde{\boldsymbol{\psi}}_1) \equiv (\tilde{\mathbf{x}} - \tilde{\boldsymbol{\psi}}_1)^T \cdot \boldsymbol{\Sigma}^{+}_{\tilde{\mathbf{q}}(\boldsymbol{\psi}_1)} \cdot (\tilde{\mathbf{x}} - \tilde{\boldsymbol{\psi}}_1) \quad , \qquad [6]$$

where the plus sign denotes the pseudoinverse operation. Note that the pseudoinverse operation employed here is of a special kind in which the pseudoinverse matrix is formed by combining only the first two nonzero singular values because the covariance matrix of the major eigenvector is of second rank, see [5,6].

DISCUSSION & CONCLUSION We proposed a new Mahalanobis distance measure for the dyadics of the major eigenvector. This measure is simple and can be used for sorting bias and it does not require ad hoc combinatorial optimization for selecting the right polarity of the eigenvector. More importantly, with the recent advances in the use of various distance measures for registering tensor data, see e.g., [7], the proposed new measure for the dyadics of the major eigenvectors will be of great important to this area of research. The proposed measure is the ideal measure for clustering of fiber tracts.

REFERENCES [1]. Basser et al. MRM 2000;44:41-50. [2] Martin et al. MRI 1999;17:893-901. [3] Yanasak et al. MRI 2008;26:122-132. [4] Koay et al. JMR 2006;182:115-125. [5]. Koay et al. IEEE TMI 2007;26:1017-1034. [6]. Koay et al . IEEE TMI 2008;27:834-846. [7]. Wu et al. MRM 2004;52:1146-1155. [8]. Irfanoglu et al. MICCAI 2009.