A Fourth Order Tensor Statistical Model for Diffusion Weighted MRI
Application to Population Comparison

Theodosios Gkamas$^1$, Félix Renard$^2$, Christian Heinrich$^1$ and Stéphane Kremer$^1$

$^1$ICube UMR 7357, University of Strasbourg – CNRS, Fédération de Médecine Translationnelle de Strasbourg (FMTS), 300 bd Sébastien Brant – CS 10413, 67412 Illkirch cedex, France
$^2$Gipsa-lab, 11 rue des Mathématiques, Grenoble Campus BP 46, 38402 St Martin d’Hères cedex, France

Keywords: Diffusion-weighted MRI, Fourth Order Tensor, Non Euclidean Metric, Nonlinear Dimension Reduction, Permutation Testing.

Abstract: In this communication, we propose an original statistical model for diffusion-weighted magnetic resonance imaging, in order to determine new biomarkers. Second order tensor (T2) modeling of Orientation Distribution Functions (ODFs) is popular and has benefited of specific statistical models, incorporating appropriate metrics. Nevertheless, the shortcomings of T2s, for example for the modeling of crossing fibers, are well identified. We consider here fourth order tensor (T4) models for ODFs, thus alleviating the T2 shortcomings. We propose an original metric in the T4 parameter space. This metric is incorporated in a nonlinear dimension reduction procedure. In the resulting reduced space, we represent the probability density of the two populations, normal and abnormal, by kernel density estimation with a Gaussian kernel, and propose a permutation test for the comparison of the two populations. Application of the proposed model on synthetic and real data is achieved. The relevance of the approach is shown.

1 INTRODUCTION

Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI) is a unique way to probe the diffusion of water molecules in the human brain in vivo. The corresponding data provide information about the underlying white matter fiber tracts. Diseases such as multiple sclerosis, Alzheimer’s disease, and stroke can be monitored using DW-MRI (Horsfield and Jones, 2002).

Data stemming from DW-MRI require mathematical modeling. Diffusion Tensor Imaging (DTI) is a popular description tool for such data (Jones, 2011). The Orientation Distribution Function (ODF) of a given voxel in white matter is accounted for by a six parameter model, also known as second order tensor (T2) model. The major limitation of the T2 model is its inability to account for crossing fibers. Nevertheless, accounting for crossing fibers is important since about half of the brain voxels are hosting crossings. To this end, the T2 ODF model has been generalized to a fourth order tensor (T4) model, comprising 15 parameters. This model encapsulates the T2 model as a particular case, and is able to account for fiber crossings (Weldeselassie et al., 2012). This versatility has led us to rely on T4 modeling (Ozarslan and Mareci, 2003; Weldeselassie et al., 2012; Tuch, 2004).

The goal of the present study is to propose an original statistical model related to T4 parameterization, in order to provide more efficient diagnosis and follow up tools for pathologies of interest. In particular, we aim at early diagnosis and at the determination of new biomarkers. Many statistical description models have been devised in the T2 framework (see e.g. (Arsigny et al., 2006)). The main concern is to account for the particular geometry of the underlying space in order to propose powerful tools. A popular model is the log-Euclidean one (Arsigny et al., 2006). The contribution of the present work is to propose a dedicated metric and corresponding statistics in the T4 framework.

A T4 ODF is parameterized by a vector in $\mathbb{R}^{15}$. This working space will be equipped with a suited metric. Besides, a 15-dimension space will be sparsely filled with data: for a given voxel of interest, we have a number of points corresponding to the number of patients (abnormal data) and controls (normal data), which is orders of magnitude below the number of points required to fill $\mathbb{R}^{15}$. Reduction of dimension of this space will improve the robustness of the subsequent statistical tests. We resort to a nonlinear dimension reduction procedure (Isomap (Tenenbaum...
et al., 2000)) which is now widely used. Isomap has already been used for group comparison with classical statistical tests in the T2 framework (Verma et al., 2007). In the reduced space (the space of features), we propose here an original statistical discriminative test between the two populations. This test allows to detect the voxels of interest, permitting the discrimination between normal and abnormal and discovering voxels where the populations are significantly different (i.e. biomarker).

This communication is organized as follows. The proposed model is presented in section 2. We detail in particular the choice of the metric, the reduction of dimension, and the statistical test in the reduced space. Experimental results on synthetic data and on real data are shown in section 3. Conclusion is drawn in section 4.

2 PROPOSED STATISTICAL MODEL

This section presents the considered data parameterization, the chosen data normalization procedure, and the proposed statistical model.

2.1 T4 ODF Parameterization

A second order semi-definite Cartesian tensor writes $f(g) = g^T D g$, where $g$ is a unit vector of $\mathbb{R}^3$ corresponding to a given direction in space and $D$ is a matrix containing the coefficients of the tensor. Such a model may be used to represent ODFs (T2 ODF model).

The generalization to fourth order semi-definite Cartesian tensors writes:

$$f(g) = \sum_{i+j+k=4} D_{i,j,k} g_i g_j g_k$$

(1)

where $(g_x, g_y, g_z)$ are the components of the unit vector $g$ and the $D_{i,j,k}$’s are the coefficients of the T4 tensor. Such a model may be used to represent ODFs (T4 ODF model). The coefficients are found by minimizing a quadratic cost function under constraints that imposes also the positivity constraint to the estimated tensor model (see (Weldesellasse et al., 2012) for details, see Fig. 1 for a T2–T4 comparison).

2.2 Data Normalization

DW-MRI data sets have to be normalized in a common reference space. Such a spatial normalization is particularly cumbersome for diffusion data, since care must be taken of the underlying fibers.

2.3 T4 Parameter Space Metric

We have to define a metric in $\mathbb{R}^{15}$, the T4 parameter space. An Euclidean distance would not be convenient, since it would not take into consideration the specificity of the corresponding underlying ODF.

A point in $\mathbb{R}^{15}$ corresponds to an ODF on the unit sphere. The metric in $\mathbb{R}^{15}$ should thus correspond to a distance between two positive valued functions on the unit sphere. Such a positivity constraint has led to log-based distances (see e.g. (Tarantola, 2005)), as is for example the case in the Kullback-Leibler divergence or in the T2 log-Euclidean framework (Arsigny et al., 2006).

Following (Tarantola, 2005), we define the distance between two ODFs $d_1$ and $d_2$ as:

$$\mathcal{D}(d_1, d_2) = \int \int |\log \frac{d_1(\theta, \phi)}{d_2(\theta, \phi)}| \sin \theta \, d\theta \, d\phi$$

(2)

where $\theta$ and $\phi$ are the angular parameters on the sphere. This distance induces a metric in $\mathbb{R}^{15}$, with $d_1$ and $d_2$ issued from two points in $\mathbb{R}^{15}$. This distance is original in the DW-MRI context, to the best of our knowledge. Other (i.e. non log-based) distances have been proposed for T4s (see e.g. (Barmpoutis et al., 2007; Du et al., 2014)).
We now have a matrix of interpoint distances attached to a voxel of interest \((i.e.)\) to a \(5 \times 5 \times 5\) ROI. Relying on T4s, the underlying space would be \(\mathbb{R}^s\), with \(s = 15\) for a mere voxel (no patch), or \(s = 3 \times 3 \times 3 \times 15 = 405\) for a 27-voxel patch. Given the number of data points (in the order of 100) in usual studies, the space is sparsely filled. Reducing the dimension will provide more robust and reliable results.

We tested several nonlinear dimension reduction methods, among which Isomap (Tenenbaum et al., 2000), maximum variance unfolding (Weinberger and Saul, 2006), and locality preserving projection (He and Niyogi, 2003; He et al., 2005). There was no significant difference between the aforementioned methods, from a discrimination point of view. We finally chose Isomap, which enables to reproduce the initial structure of points in a reduced space, say \(\mathbb{R}^2\) or \(\mathbb{R}^3\), by preserving the geodesic interpoint distances. Scree plots lead us to work in \(\mathbb{R}^2\), as is the case in (Verma et al., 2007).

### 2.6 Statistical Reduced Space Model

Each of the two populations (normal and abnormal) corresponds to a set of points in the reduced space. We represent the probability densities \(p_1\) and \(p_2\) corresponding to the populations using kernel density estimation, with a Gaussian kernel (Hastie et al., 2011). One Gaussian kernel is attached to each point. There is one parameter to be determined for each population, which is the covariance matrix attached to each point of this population. These matrices are determined using Scott’s rule (Scott, 1992).

The discrepancy between densities \(p_1\) and \(p_2\) has to be quantified. The Kullback-Leibler divergence and its symmetrized version are popular choices. Nevertheless, in the case of a mixture of Gaussian distributions corresponding to our kernel density estimation, there is no closed formula available and the numerical estimation of the Kullback-Leibler divergence is time consuming. We rely on another discrepancy measure, noted \(P\), which provides good results with a low computation time (Sfikas et al., 2005). The computation of \(P\) is about 150 times faster than the computation of the symmetrical Kullback-Leibler distance in the present case. The discrepancy \(P\) writes:

\[
P(p_1, p_2) = -\log \left[ \frac{2 \int p_1(x)p_2(x) \, dx}{\int (p_1(x))^2 \, dx + \int (p_2(x))^2 \, dx} \right]
\]  

(3)

Considering the mixtures

\[
\begin{align*}
p_a(x) &= \sum_{j=1}^{I} \pi_j^{(a)} N \left( \mathbf{x}; \mu_j^{(a)}, \Sigma_j^{(a)} \right) \\
p_b(x) &= \sum_{j=1}^{J} \pi_j^{(b)} N \left( \mathbf{x}; \mu_j^{(b)}, \Sigma_j^{(b)} \right)
\end{align*}
\]  

(4)

a straightforward computation yields

\[
\int p_a(x)p_b(x) \, dx = \sum_{i,j} \pi_i^{(a)} \pi_j^{(b)} \left( \frac{\exp(k) |\mathbf{V}|}{(2\pi)^{N_x}|\Sigma_i^{(a)}|^{1/2} |\Sigma_j^{(b)}|^{1/2}} \right)
\]  

(5)

where \(N_x\) is the dimension of \(\mathbf{x}\) and with
The discrepancies $P$ are a statistic. We have to estimate the corresponding p-value $\nu^*$, i.e. the probability of getting a distance $P$ larger than the reference one $P(p_1, p_2)$ under the hypothesis that both populations are indiscernible. This will be achieved resorting to permutation testing. Besides, we want to determine a confidence interval associated to the estimation of $\nu^*$. This interval will depend on the number of permutations (i.e. label shufflings) used. To address these issues, we will resort to the Bayesian framework.

We shuffle $N$ times the labels of the points in the reduced space, and for each shuffling we compute the value of $P$. Comparison of the $n^{th}$ sample of $P$ with the reference $P(p_1, p_2)$ yields a binary value $q_n$, which is a sample of the Bernoulli distribution of (unknown) parameter $\nu^*$. The problem is now the estimation of $\nu^*$ from binary samples $q_1, \ldots, q_N$. Using a uniform prior for $\nu$, we derive the posterior $p(\nu | q_1, \ldots, q_N)$. The confidence interval (or credibility interval in $\nu$ is taken here as being the smallest interval capturing 99 % of the a posteriori mass, which is known as Highest Probability Density (HPD) interval in the Bayesian framework. We thus obtain a confidence interval for the p-value. If needed, the length of the confidence interval may be reduced by using more shufflings.

Finally, the voxels are sorted according to increasing p-values. The most dissimilar voxels (the biomarkers) appear on top of the list. We do not address multiple comparison issues here, since we do not provide a list of voxels considered as significantly different between populations. This is left for future work.

3 EXPERIMENTAL RESULTS

In this section, we present experimental results based on synthetic data and on real data. The pathology of
Figure 3: Representation of probability densities in the reduced space. Green dots: normal cases; red squares: abnormal cases. (a) dissimilar populations; (b) and (c) similar populations. Left column: representation of the probability density corresponding to the normal population; right column: representation of the probability density corresponding to the abnormal population (blue: low density; red: high density. The same color scale is used across all subfigures.). The densities correspond to kernel density estimation with a Gaussian kernel.

Table 1: P-value HPD intervals of the cases depicted in Fig. 3. We used 1000 shuffles. It took approximately 3 minutes on a standard computer to process each voxel. Most of the computation time was devoted to determining the patch correspondence.

<table>
<thead>
<tr>
<th>Cases</th>
<th>p-value HPD interval</th>
<th>width of interval</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>[0, 0.0046]</td>
<td>0.0046</td>
<td>Dissimilar</td>
</tr>
<tr>
<td>(b)</td>
<td>[0.35, 0.43]</td>
<td>0.08</td>
<td>Similar</td>
</tr>
<tr>
<td>(c)</td>
<td>[0.988, 0.999]</td>
<td>0.011</td>
<td>Similar</td>
</tr>
</tbody>
</table>

T2-T4 comparison on Synthetic Data. We consider a case of two fibers crossing at 90 degrees. Both fibers have the same amplitude. The two populations differ by a rotation angle of 5 degrees affecting both fibers likewise. Noise was added to the synthetic data. The T4 model is able to discriminate the two populations, whereas the T2 model is not able to do so (this fiber crossing is seen by the T2 model as an isotropic ODF, independently of the rotation angle).

Comparison of the proposed Statistical Test with the Hotelling Test on Synthetic Data. We consider a case of crossing fibers, where the angle between the two fibers differs from one population to the other (the values considered are 90 degrees for the normal population, and 85 degrees for the abnormal population). The amplitude of the fibers are different as well (the amplitude is 1 for the fiber common to both populations; for the fiber with a different orientation in both populations, the amplitude is 0.35 for the normal population, and 0.45 for the abnormal population).

The proposed statistical test is able to discriminate both populations. Besides, we perform a Hotelling test. This test takes place in the reduced space. Each population is modeled by a Gaussian distribution, and means are compared, looking for a statistical difference. The Hotelling test was not able to discriminate both populations.

T2-T4 comparison on Real Data. We compared the results (similar/dissimilar) issued from T2 model and from T4 model. Whenever the results disagreed, it happened that the T4 model gave the correct answer and that the T2 model was not able to capture all information conveyed by the data. This was validated by testing the residuals (the dissimilarity information is contained in the T2 residual). Using a log-Euclidean distance on second order tensors to reduce the dimension of the data yields globally the same results as those issued by the metric we propose, when applied on T2 profiles.

P-value Distribution on Real Data. On Fig. 4, we represent the distribution of p-values (one p-value for each voxel), for a region of interest (ROI). The ROI is identified as being sensitive to the disease of interest (NMO). This is confirmed on the plot, where a peak of low p-values appears.

A Fourth Order Tensor Statistical Model for Diffusion Weighted MRI - Application to Population Comparison
are concerned as well.

4 CONCLUSION

We proposed an original statistical model for fourth order tensor ODF modeling of DW-MRI. This model incorporates a suitable metric in the parameter space. It relies on nonlinear dimension reduction, and on an original statistic in the reduced space, allowing to compare two populations and to extract biomarkers. This approach has shown better ability to discriminate two populations, as compared to models relying on other metrics, on T2 ODF profiles, or on the Hotelling test. It has thus a potential for early diagnosis.

Resorting to T4 ODF profiles will also enable us to identify more precisely the changes effectively taking place between the two populations, as compared to changes identified by T2 models where the profiles are described less accurately. This is left for future work.

REFERENCES


