Which Manifold Should be Used for Group Comparison in Diffusion Tensor Imaging?

A. Bouchon[1], V. Noblet[1], F. Heitz[1],
J. Lamy[1], F. Blanc[1,2], and J.-P. Armspach[1]

[1] ICube, University of Strasbourg, CNRS,
Fédération de Médecine Translationnelle de Strasbourg (FMTS), France
[2] Neuropsychology Service, Department of Neurology,
University Hospital of Strasbourg, France

Abstract. Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) is a modality which allows to investigate the white matter structure by probing water molecule diffusion. A common way to model the diffusion process is to consider a second-order tensor, represented by a symmetric positive-definite matrix. Currently, there is still no consensus on the most appropriate manifold for handling diffusion tensors. We propose to evaluate the influence of considering an Euclidean, a Log-Euclidean or a Riemannian manifold for conducting group comparison in DT-MRI. To this end, we consider a multi-linear regression problem that is solved on each of these manifolds. Statistical analysis is then achieved by computing an F-statistic between two nested (restricted and full) models. Our evaluation on simulated data suggests that the performance of these manifolds varies with the kind of modifications that has to be detected, while the experiments on real data do not exhibit significant difference between the methods.

1 Introduction

Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) is a modality commonly used to investigate the cerebral white matter integrity. There is a great need in the neuroscientific community for efficient tools to compare DT-MRI across cohorts of subjects according to clinical or cognitive data. Most studies focus on the comparison of scalar images derived from DT-MRI such as Fractional Anisotropy (FA) or Mean Diffusivity (MD) using either the voxel-based analysis framework [1] provided in SPM[1] or the tract-based spatial statistics (TBSS) method [2] provided in FSL[2]. However, these methods do not exploit all the information contained in tensor images and thus cannot detect all kind of changes. Different statistical frameworks have been proposed to compare the whole diffusion tensor information [3,4,5,6,7]. Only a few take into account covariates (e.g. age, sex, cognitive scores) [4,6,7]. A convenient way to introduce

[1] `http://www.fil.ion.ucl.ac.uk/spm/`
[2] `http://fsl.fmrib.ox.ac.uk/fsl/`

N. Navab et al. (Eds.): MICCAI 2015, Part I, LNCS 9349, pp. 158–165, 2015.
DOI: 10.1007/978-3-319-24553-9_20

these covariates is to consider a linear model [4,6,7]. The regression estimation can be done by considering different metrics to compute the residuals. In this work, we propose to compare the influence of considering either an Euclidean, a Log-Euclidean or a Riemanian metric in the regression problem. A few works have already evaluated the impact of these metrics for various image processing problems [8,9], but there is still no consensus on the most appropriate one, especially in the context of group comparison. To compare the influence of these metrics, a simulation framework has been set up, based on DT-MRI acquisitions of healthy subjects in which different kinds of lesions have been introduced. This study has also been complemented by results on a cohort of patients suffering from neuromyelitis optica (NMO).

2 Methods

2.1 Pre-processing

To conduct group studies at the voxel level, all images should first be registered in a common space. This is done by registering all FA images derived from DT-MRI on a common template using an affine followed by a non-rigid registration method [10]. Spatial transforms estimated from FA images are then applied on tensor images using the Preservation of Principal Direction (PPD) reorientation strategy.

2.2 Multi-linear Regression

Let $\mathcal{M}$ be a manifold and $\{y_i\}_{i\in[1..N]} \in \mathcal{M}$ the observations from N individuals, each individual being characterized by K explanatory variables $\{x_{i,j}\}_{j\in[1..K]}$ such as age, gender or group affiliation. The regression problem consists in estimating a function $f : \mathbb{R}^K \mapsto \mathcal{M}$ that best fits all the couples $(\{x_{i,1} \dots x_{i,K}\}, y_i)$. A simple parametric model is to consider a multi-linear function:

$$y_i = \alpha + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + \beta_K x_{i,K} + \varepsilon_i \tag{1}$$

where α is the intercept, β_i are the regression coefficients and ε_i are the residuals.

Case of Scalar Observations. If y_i are scalar observations (*i.e.,* $\mathcal{M} \subseteq \mathbb{R}$), this regression problem amounts to the general linear model commonly used to perform voxelwise group comparison on scalar images derived from DT-MRI [1]. Let $Y = [y_1 \dots y_i \dots y_N]^t$, $X\,[i,j] = x_{i,j}$ and $B = [\beta_1 \dots \beta_j \dots \beta_K]^t$. If the residuals ε_i are assumed to be independent and identically distributed according to a normal distribution, then the least squares estimate of B can be obtained analytically:

$$\hat{B} = \arg\min_{B\in\mathbb{R}^K} \|Y - XB\|^2 = (X^tX)^{-1}X^tY \tag{2}$$

To extend this framework to tensors images, several strategies can be implemented depending on the assumption made on the manifold $\mathcal{M}$.

Euclidean Framework. A diffusion tensor can be represented by $D^i = [D^i_{xx}\ D^i_{xy}\ D^i_{xz}\ D^i_{yy}\ D^i_{yz}\ D^i_{zz}]^t \in \mathbb{R}^6$. The previous regression problem can straightforwardly be extended to the multivariate case by assuming the noise variance on each tensor element to be identical (homoscedasticity assumption) [7]. This method will be referred in the sequel as General Linear Model for Diffusion Tensors (*GLM-DT*).

The basic idea is to concatenate all tensor components of the N individuals in a single vector $Y \in \mathbb{R}^{6N}$. For each explanatory variable, six regressors are estimated, one associated with each tensor component. This is done by constructing a new design matrix $X[i,j] = x_{i,j}$, for $i = 1 \ldots N \times 6$ and $j = 1 \ldots K \times 6$, where each explanatory variable is replicated in six columns: the first column is composed of the values of the explanatory variable for the entries corresponding to the first tensor components D^i_{xx} and of zeros for the others entries, and so on for the five others columns. With that formulation, the least squares estimate of B has also a closed-form solution and leads to the estimation of K regressors of six components [7].

Log-Euclidean Framework. The Euclidean framework does not take into account that diffusion tensors are symetric positive-definite matrices, *i.e.* $\mathcal{M} = \mathrm{Sym}^+(3)$, which is in fact only a subset of $\mathbb{R}^6$. Consequently, the estimated regression can possibly map some sets of explanatory variables to vectors that do not correspond to positive-definite matrices. A way to circumvent this limitation is to conduct the regression on the logarithm of tensors instead on the tensors directly. Indeed, the logarithmic transformation enables to map the space of symmetric positive-definite matrices $\mathrm{Sym}^+(3)$ to the space of symmetric matrices $\mathrm{Sym}(3)$. Thus, the same framework as in [7] can be used for the regression estimation. The corresponding regression model is then given by:

$$Y = \exp(XB + \varepsilon) \tag{3}$$

where $\exp(\cdot)$ stands for the matrix exponential. This method will be referred as General Linear Model for Diffusion Log-Tensors (*GLM-LOG-DT*).

Riemannian Framework. Recently, a multiple geodesic regression model based on Riemannian geometry has been proposed [6]. This framework enables to account for the positive-definite nature of diffusion tensors. The corresponding regression model is given by:

$$y_i = \mathrm{Exp}(\mathrm{Exp}(\alpha, \sum_{j=1}^{N} \beta^j x_{i,j}),\ \varepsilon_i) \tag{4}$$

where Exp refers to the exponential map. The regressors β and the intercept α are simultaneoustly estimated using a gradient descent scheme by minimizing the distance between the data y_i and the estimate $\hat{y}_i = \mathrm{Exp}(\alpha, \sum_{j=1}^{N} \beta^j x_{i,j})$. To this end, the following geodesic distance is used:

$$d(y_i, \hat{y}_i) = \sqrt{\langle \mathrm{Log}(y_i, \hat{y}_i),\ \mathrm{Log}(y_i, \hat{y}_i) \rangle_{y_i}} \tag{5}$$

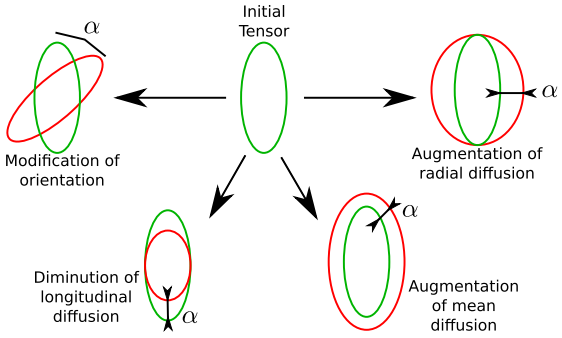

Fig. 1. Representation of the four kinds of lesions. The green and the red tensors represent respectively the initial tensor and the modified tensor.

The implementation of this method has been made available on `https://www.nitrc.org/projects/riem_mglm`. This method will be referred in the sequel as Manifold General Linear Model for Diffusion (*MGLM*).

2.3 Statistical Test

The objective of the statistical test is to evaluate whether a given explanatory variable has a significant contribution in the regression model. To this end, an F-test is used to compare the Residuals Sum of Squares (RSS) of two nested models: a full model taking into account all the covariates (RSS_2) and a restricted model where the covariate of interest is discarded (RSS_1):

$$F = \frac{\frac{RSS_1 - RSS_2}{p_2 - p_1}}{\frac{RSS_2}{N - p_2}} \tag{6}$$

with p_2 and p_1 representing respectively the number of covariates for both models. Assuming that the residuals follow a normal distribution, F follows a Fisher distribution with $p_2 - p_1$ and $N - p_2$ degrees of freedom under the null hypothesis that model 2 does not provide a significantly better fit than model 1. The Gaussianity assumption of the residuals is a reasonable hypothesis for the Euclidean and Log-Euclidean frameworks, but it is no longer valid for the Riemannian framework. Permutations may be used to obtain corresponding p-values in the latter case [6], but it was not done in this work because of prohibitive computational time.

3 Validation Framework

The Euclidean, the Log-Euclidean and the Riemannian frameworks are compared on both simulated and real data. All DT-MRI images were acquired on a 1.5T Siemens scanner with 30 encoding gradients (b-value of 1000 s/mm^2) and two baseline images (*i.e.,* b-value of 0 s/mm^2). The image dimensions are $128 \times 128 \times 41$ and the spatial resolution is $1.8 \times 1.8 \times 3.5\ mm^3$.

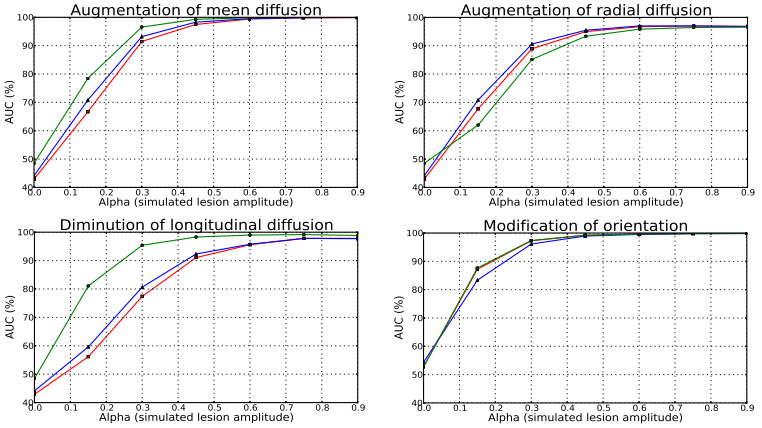

Fig. 2. Area under the ROC curve (AUC) *vs* lesion amplitude (α) (Legend: ● = GLM-DT, ▲ = GLM-LOG-DT, ■ = MGLM)

3.1 Synthetic Data

A set of 11 images with simulated lesions in the corpus callosum (CC) are generated from the DT-MRI acquisitions of 11 healthy subjects. We focused on the CC since it consists of a single direction white matter tract, where the tensor assumption holds. These simulated images, considered in the experiments as the patient group, are compared with the group of 11 images of healthy subjects, introducing age, gender, and group affiliation as covariates. For the statistical test, the group affiliation is discarded in the restricted model. Four kinds of lesions are simulated: mean and radial diffusion augmentation, longitudinal diffusion diminution and diffusion orientation modification. A parameter $\alpha \in [0 : 0.9]$ controls the amplitude of the simulated lesions (Fig.1). An image is generated for each α and each kind of lesions, leading to $4 \times 19 = 76$ groups of 11 patients images. The three first kind of lesion are consistent with real case scenarios [11]. The last one is rather a toy example, since a realistic diffusion orientation modification would imply the deflection of the whole trajectory of a fiber bundle.

Methods are compared using ROC analysis, which allows to compare the statistical maps with the ground truth for various statistical thresholds. The area under the ROC curve (AUC) is computed for each experiment. An AUC of 50% corresponds to a random detection and 100% to a perfect detection. AUCs are plotted with respect to the lesions amplitude α for each kind of simulated lesion (Fig. 2). This way, it is possible to assess the performance of each method for major changes as well as smaller ones and for different types of lesions.

3.2 Neuromyelitis Optica Cohort

The Neuromyelitis Optica (NMO) is an inflammatory disease characterized by alterations of the normal-appearing white matter in relation to several disorders [12]. A group of 34 patients suffering from NMO are compared to a group

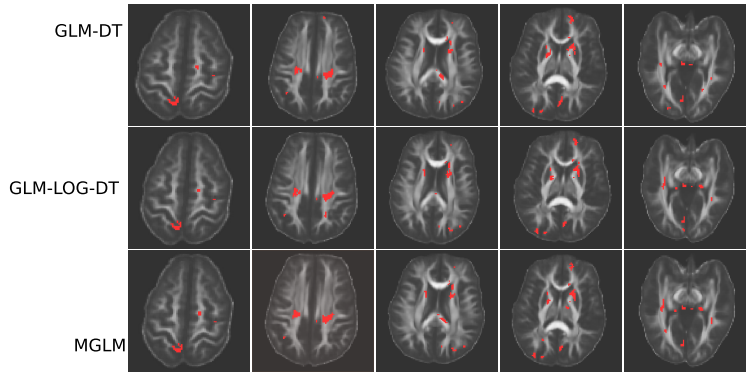

Fig. 3. Results obtained for the three methods on the comparison of NMO patients ($n = 34$) to healthy subjects ($n = 22$). The statistical thresholds have been chosen to detect the 5% most significant voxels within the white matter mask for each method. Clusters of less than 10 voxels are discarded.

of 22 healthy subjects by considering age, gender, and group affiliation as covariates. For the statistical test, the group affiliation is discarded in the restricted model. Since the F-statistic maps may not follow the same distribution for the three methods, they cannot be compared by using the same statistical threshold. To obtain comparable detection maps, the threshold is adjusted for each method to obtain the same number of detected voxels within the white matter mask for each method. To compare the similarity of the detection maps obtained by two methods, the Dice coefficient is computed. This is done for a wide range of thresholds (adjusted for each method) in order to compare the behavior of the three methods for different levels of sensitivity (Fig. 4).

4 Results

4.1 Results on Synthetic Data

The results on simulated data are summed up in Fig. 2. They highlight that the performance of the methods varies with the kind of simulated changes. On one hand, the *GLM-DT* outperforms the two other methods for detecting a mean diffusivity augmentation and a longitudinal diffusion diminution. On the other hand, the *GLM-LOG-DT* and the *MGLM* exhibit better results for detecting an augmentation of the radial diffusion. Notice that all the three methods show similar performance for detecting modifications of orientation. Finally, it has to be pointed out that the Log-Euclidean and Riemannian frameworks lead to very close results.

4.2 Results on NMO Patients

Fig.3 shows the results obtained by the three methods for the comparison of 34 NMO patients to 22 healthy subjects. The statistical thresholds have been

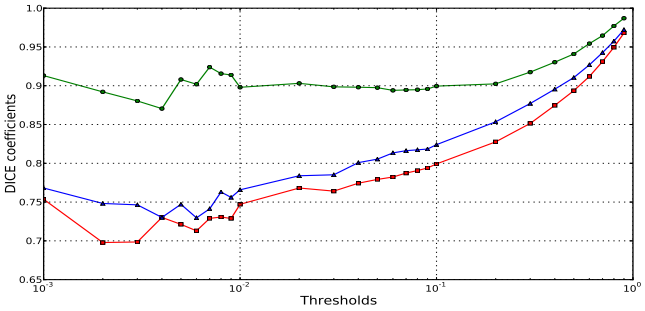

Fig. 4. Dice coefficient *vs* percentage of detected voxels in logarithmic scale (Legend: ● = MGLM / GLM-LOG-DT, ▲ = GLM-DT / GLM-LOG-DT, ■ = MGLM / GLM-DT)

chosen to detect the 5% most significant voxels within the white matter mask for each method. Visually, all three methods lead to very close results, which does not allow us to assert that one method outperforms the others. They succeed to detect the regions involved in the NMO pathology [12]: the visual cortex, the corticospinal tract and the CC that are respectively related to the visual, motor and cognitive disorders induced by the pathology. To strengthen the claim that there is no significant difference between the methods on these data, the Dice coefficient for the three couples are presented in Fig. 4 for several statistical thresholds. It shows a good agreement between the methods, with a Dice coefficient always greater than 70%. The best agreement is obtained between the *GLM-LOG-DT* and *MGLM* frameworks (Dice coefficient greater than 90%), which is in accordance with the conclusion on simulated data.

5 Conclusion and Discussion

In this paper, the Euclidean, the Log-Euclidean and a Riemannian framework have been compared in the context of a DT-MRI group comparison. The experimental results on simulated lesions highlight that all methods can efficiently detect the four kinds of lesions, with performance varying for each framework according to the kind of simulated lesions. The results on the real database confirm the good ability of the three methods to detect the regions affected by the NMO, without exhibiting any significant difference between the detection maps. Typical computation time on a standard workstation is about 2 min for the Euclidean and Log-Euclidean approaches as compared to more than 80 hours for the Riemannian framework on the real database with the provided implemantation. This time gap may probably be reduced by optimizing the implementation of the Riemannian framework, for instance by rewriting the matlab code in C++ language. But in any case, the Riemannian framework would still be much more computationally intensive than the Euclidean and Log-Euclidean approaches. In conclusion, this study suggests that, despite the mathematical elegance offered by the Riemannian framework, its superiority on both simulated and real clinical

data could not be demonstrated compared to the Euclidean and Log-Euclidean frameworks. For a similar performance, these two latter methods present the advantage to be easily implementable and computationally effective. Bearing in mind the limitation of tensors to model crossing fibers, we plan to investigate more complex diffusion models.

Acknowledgements. We thank Pr Jérome de Sèze for the recruitment of NMO patients and Pr Stéphane Kremer for image acquisitions.

References

1. Penny, W., Friston, K., Ashburner, J., Kirbel, S., Nichols, T.: Statistical Parametric Mapping: The Analysis of Functional Brain Images. Elsevier LTD, Oxford (2006)
2. Smith, S., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T., Mackay, C., Watkins, K., Ciccarelli, O., Cader, M., Matthews, P., Behrens, T.: Tract based spatial statistics: voxelwise analysis of multi-subjects diffusion data. NeuroImage 31(4), 1487–1505 (2006)
3. Whitcher, B., Wisco, J.J., Hadjikhani, N., Tuch, D.S.: Statistical group comparison of diffusion tensors via multivariate hypothesis testing. Magn. Reson. Med. 57(6), 1065–1074 (2007)
4. Zhu, H., Chen, Y., Ibrahim, J., Li, Y., Hall, C., Lin, W.: Intrinsic regression models for positive-definite matrices with applications to Diffusion Tensor Imaging. Journal of the American Statistical Association 104(487), 1203–1212 (2009)
5. Schwartzman, A., Dougherty, R., Taylor, J.: Group comparison of eigenvalues and eigenvectors of diffusion tensors. Journal of the American Statistical Association 105(490), 588–599 (2010)
6. Kim, H., Bendlin, B., Adluru, N., Collins, M., Chung, M., Johnson, S., Davidson, R., Singh, V.: Multivariate General Linear Models (MGLM) on Riemannian Manifolds with Applications to Statistical Analysis of Diffusion Weighted Images. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2705–2712 (June 2014)
7. Bouchon, A., Noblet, V., Heitz, F., Lamy, J., Blanc, F., Armspach, J.P.: General linear models for group studies in diffusion tensor imaging. In: Proceedings of International Symposium on Biomedical Imaging: From Nano to Macro, Beijing, China (2014)
8. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-euclidean metrics for fast and simple calculus on diffusion tensors. In: Magnetic Resonance in Medicine (2006)
9. Pasternak, O., Sochen, N., Basser, P.J.: The effect of metric selection on the analysis of diffusion tensor MRI data. NeuroImage 49(3), 2190–2204 (2010)
10. Noblet, V., Heinrich, C., Heitz, F., Armspach, J.P.: Retrospective evaluation of a topology preserving non-rigid registration method. Medical Image Analysis 10(3), 366–384 (2006)
11. Harsan, L.A., Poulet, P., Guignard, B., Steibel, J., Parizel, N., de Sousa, P., Boehm, N., Grucker, D., Ghandour, M.S.: Brain dysmyelination and recovery assessment by noninvasive in vivo diffusion tensor magnetic resonance imaging. Journal of Neuroscience Research 83(3), 392–402 (2006)
12. Yu, C., Lin, F., Li, K., Jiang, T., Qin, W., Sun, H., Chan, P.: Pathologenesis of normal-appearing white matter damage in neuromyelitis optica: Diffusion-Tensor MR Imaging. Radiology 246(1), 222–228 (2008)