NeuroImage 142 (2016) 99–112

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Which is the most appropriate strategy for conducting multivariate voxel-based group studies on diffusion tensors?

Alix Bouchon [a,*], Vincent Noblet [a], Fabrice Heitz [a], Julien Lamy [a], Frédéric Blanc [a,b], Jean-Paul Armspach [a]

[a] ICube, University of Strasbourg, CNRS, Fédération de Médecine Translationnelle de Strasbourg (FMTS), France

[b] Geriatry service, Memory Resources and Research Center (CMRR), University Hospital of Strasbourg, France

ARTICLE INFO

Article history:
Received 26 October 2015
Revised 12 May 2016
Accepted 13 May 2016
Available online 27 May 2016

Keywords:
Diffusion tensor imaging
General linear model
Group comparison
Tensor-based regression
Euclidean
Log-Euclidean and Riemannian manifolds

ABSTRACT

There is a real need in the neuroscience community for efficient tools to compare Diffusion Tensor Magnetic Resonance Imaging across cohorts of subjects. Most studies focus on the comparison of scalar images such as fractional anisotropy or mean diffusivity. Although different statistical frameworks have been proposed to compare the whole diffusion tensor information, they are still seldom used in neuroimaging studies. In this paper, we investigate on both simulated and real data whether there is a real added value of considering the whole tensor information for conducting voxel-based group studies. Then, we compare two statistical tests dedicated to tensor, namely the Cramér test and a tensor-based extension of the General Linear Model (GLM), the latter presenting the advantage to account for covariates. We also evaluate the impact of different metrics (Euclidean, Log-Euclidean and affine-invariant Riemannian metrics) for estimating the GLM parameters. Finally, we address the problem of interpreting the change detection maps obtained by tensor-based methods by proposing a way to characterize each of the detected clusters according to several scalar indices. Our study suggests that if there is no prior assumption about the nature of the expected changes, it is really preferable to use tensor-based rather than scalar-based statistical analysis. The Cramér test can advantageously be used when no confounding variable hampers the group comparison, otherwise the GLM should be considered. Finally, the different metrics show similar performance in the real scenario, with a significant computational overhead for the Riemannian framework.

Introduction

Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) allows clinicians and neuroscientists to investigate the white matter structure by probing water molecule diffusion. Group studies allow to find out brain regions that are statistically different between two populations or that are significantly correlated with some covariates. As a consequence, there is an increasing need for efficient tools to compare DT-MRI across cohorts of subjects according to clinical or cognitive data. This may help to characterize the damages caused by a pathology and to understand the mechanisms underpinning the disease. Group comparisons of DT-MR images are usually performed using either region-of-interest-based analysis, tract-based analysis, or voxel-based analysis (Cercignani, 2010). In this paper, we will focus on voxel-based analysis, whose major interest is that it does not make any assumption on the spatial location of the expected changes.

Most voxel-based 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 statistical analysis framework provided in SPM[1] (Penny et al., 2006) or the tract-based spatial statistics (TBSS) method provided in the FSL library[2] (Smith et al., 2006). However, these methods do not exploit all the information contained in tensor images and thus cannot detect all kinds of changes. Although different statistical frameworks have been proposed to compare either several scalar indices simultaneously (Chapell et al., 2008), eigenvalues or eigenvectors of diffusion tensors (Schwartzman et al., 2010), or even the whole diffusion tensor information (Whitcher et al., 2007; Zhu et al., 2009; Yuan et al., 2012; Kim et al., 2014; Bouchon et al., 2014), they remain seldom used in neuroimaging studies.

In this paper, we investigate whether there is a real added value of considering the whole tensor information as compared to a single scalar index or several scalar indices simultaneously for conducting voxel-based group studies. To this end, we consider a tensor-based extension of the General Linear Model (GLM), which provides a convenient way to carry out statistical analysis while taking into account several covariates. We evaluate the impact of considering different manifolds for estimating GLM parameters, namely the Euclidean, Log-Euclidean (Arsigny et al., 2006) or the Riemannian framework (Pennec, 1999). We also

* Corresponding author.
E-mail address: a.bouchon@unistra.fr (A. Bouchon).

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

http://dx.doi.org/10.1016/j.neuroimage.2016.05.040
1053-8119/

provide a comparison with an alternative tensor-based test, namely the tensor-based Cramér test, which has already been reported as the best performer among various tensor-based statistical tests (Whitcher et al., 2007). Finally, we deal with the problem of interpreting the change detection map obtained by tensor-based methods by proposing a way to characterize each of the detected clusters according to several scalar indices. To investigate these different points, a simulation framework has been set up, based on DT-MRI acquisitions of healthy subjects in which different kinds of lesions have been introduced. A group study on a cohort of patients suffering from neuromyelitis optica (NMO) is also presented.

The content of the paper is organized as follows. The General framework for DTI group comparison section depicts the whole framework for DTI group comparison by describing the pre-processing steps, the spatial normalization, the regression, the statistical analysis and the characterization of the detected clusters. The evaluation framework is presented in the Validation framework section. Results on both simulated and real data are presented in the Results section and discussed in the Discussion section. Finally, conclusions and perspectives are given.

General framework for DTI group comparison

Overview of the processing pipeline

The goal of voxel-based group comparison is to decide for each voxel $\mathbf{s}\in\mathbb{R}^3$ whether there is a significant difference between two groups of subjects while taking into account some covariates such as age, gender, *etc.* To this end, all subjects should first be registered into a common space according to a deformable registration method (the Spatial normalization section). A prerequisite to this step is to first control the quality of each acquisition to assess whether it can be included in the analysis. This is performed through visual inspection of each DWI image and by using a dedicated tool, DTIPrep[3] (Oguz et al. 2014). Then, each Diffusion-Weighted Image (DWI) is corrected for subject motion and eddy current distortions using FSL's Diffusion Toolkit. Finally, tensors are estimated from the DWI images according to a weighted least square (WLS) estimation procedure (Zhu et al., 2007a).

After the registration step, all tensor images are filtered in the Log-Euclidean space (Arsigny et al., 2006) using a Gaussian kernel (FWHM = 8 mm in all our experiments). This allows to account for the spatial information from the neighboring voxels and also, by the central limit theorem, to render the data more normally distributed (Jones et al., 2005). Then, the statistical analysis is conducted.

A convenient way to carry out statistical analysis while taking into account several covariates is to consider the General Linear Model (GLM). It can be done either on a given scalar image such as FA or MD, or on several scalar images simultaneously, or on tensors directly using either an Euclidean, a Log-Euclidean or a Riemannian metric in the regression (the Multi-linear regression section). A statistical F-test is then used to evaluate whether a given explanatory variable has a significant contribution in the regression model (the Statistical test section). Finally, the statistical map is thresholded and each detected cluster is characterized by a signature obtained as a combination of several scalar indices (the Characterization of the detected clusters section). A graphical overview of the pipeline is presented in Fig. 1.

Spatial normalization

Before performing voxel-based statistical analysis, all images are first registered on the same arbitrary chosen reference image belonging to the control group. To this end, all FA images derived from DT-MRI acquisitions are registered through an affine method followed by a non-rigid method (Noblet et al., 2006) using the sum of squared differences metric. Spatial transforms are then applied on tensor images using the Preservation of Principal Direction (PPD) (Alexander et al., 2001) reorientation strategy. This step is required to ensure the orientational consistency of the warped tensor field. The basic idea is to apply the local affine component F of the deformation field on the three tensor eigenvectors (e_1,e_2,e_3), and then to estimate the rotation that leads to the reoriented eigenvectors $(\tilde{e}_1,\tilde{e}_2,\tilde{e}_3)$, such that Fe_1 and $\tilde{e}_1$ are aligned, and such that Fe_1 and Fe_2 span the same plane as $\tilde{e}_1$ and $\tilde{e}_2$ (Alexander et al., 2001).

Multi-linear regression

In this section, we will present several regression schemes on either scalar images or on tensor fields. Let $\mathcal{M}$ be a manifold and $\{y_i\}_{i\in[1..N]}\in\mathcal{M}$ be the observations at a given voxel $\mathbf{s}\in\mathbb{R}^3$ 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. These observations can be either scalar indices such as FA or MD $\mathcal{M}\subset\mathbb{R}$ or full tensors $\mathcal{M}\subset\mathbb{R}^6$. 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 approach is to consider the multi-linear function:

$$y_i=\alpha+\beta_1x_{i,1}+\beta_2x_{i,2}+\cdots+\beta_Kx_{i,K}+\varepsilon_i \tag{1}$$

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

For the scalar case, this corresponds to the standard GLM, which is commonly used in neuroimaging studies (Penny et al., 2006). For the tensor case, we investigate the use of different metrics to compute the residuals, namely an Euclidean ($\mathcal{M}\subset\mathbb{R}^6$), a Log-Euclidean ($\mathcal{M}\subset\mathrm{Sym}(3)$) and a Riemannian ($\mathcal{M}\subset\mathrm{Sym}^+(3)$) metric, in order to evaluate their impact in the context of group comparison.

Scalar regression

In the scalar case, the General Linear Model (GLM) aims at representing N scalar observations (FA or MD) $Y=[y_1\dots y_i\dots y_N]^t$ as a linear combination of K explanatory variables, whose values for the *ith* observation are stored in the design matrix $X[i,j]=x_{i,j}$, for $i=1\dots N$ and $j=1\dots K$. If the residuals ε_i are assumed to be independent and identically distributed (i.i.d.) according to a normal distribution, then the least squares estimate of B is obtained analytically:

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

The regression methods on scalar images will hereafter be referred to as General Linear Model for FA (*GLM-FA*) or for MD (*GLM-MD*).

Tensor-based regression: Euclidean framework

Scalar regression does not capture all the information embedded in the diffusion tensor, in particular the orientation information. Thus, we extended the GLM to take advantage of the full tensor information (Bouchon et al., 2014).

A 3×3 diffusion tensor matrix is symmetric ($\in\mathrm{Sym}(3)$) and may be represented by a vector: $D^i=[D^i_{xx}\ D^i_{xy}\ D^i_{xz}\ D^i_{yy}\ D^i_{yz}\ D^i_{zz}]^t\in\mathbb{R}^6$. The previous regression framework can straightforwardly be extended to the multivariate case by assuming the noise on the tensor components to be i.i.d. We also assume equal noise variance for all components (homoscedasticity assumption[4]). 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 $i=1\dots N\times6$ and

[3] https://www.nitrc.org/projects/dtiprep/

[4] In a previous work, we have also investigated the heteroscedasticity assumption (*i.e.*, different noise variances for each tensor component), but our experiments did not enable us to exhibit any improvement (Bouchon et al., 2014).