FADTTS

FADTTS represents **F**unctional **A**nalysis of **D**iffusion **T**ensor **T**ract **S**tatistics. The aim of this tool is to implement a functional analysis pipeline, called FADTTS, for delineating the structure of the variability of multiple diffusion properties along major white matter fiber bundles and their association with a set of covariates of interest, such as age, diagnostic status and gender, in various diffusion tensor imaging studies. The FADTTS integrates five statistical tools: a multivariate varying coefficient model for allowing the varying coefficient functions to characterize the varying association between fiber bundles diffusion properties and a set of covariates, a weighted least squares estimation to estimate the varying coefficient functions, a functional principal component analysis to delineate the structure of the variability in fiber bundles diffusion properties, a global test statistic to test hypotheses of interest, and a simultaneous confidence band to quantify the uncertainty in the estimated coefficient function. FADTTS can be used to facilitate understanding normal brain development, the neural bases of neuropsychiatric disorders, and the joint effects of environmental and genetic factors on white matter fiber bundles.

Motivation: Diffusion Tensor Imaging (DTI), which can track the effective diffusion of water in the human brain in vivo, has been widely used to map the structure and orientation of the white matter fier tracts of the brain (Basser et al., 1994b,a). In the current literature, three major approaches to the group analysis of diffusion imaging data are region-of-interest (ROI) analysis, voxel based analysis, and fiber tract based analysis (Smith et al., 2006; O'Donnell et al., 2009; Snook et al., 2007). The ROI analysis used in some neuroimaging studies (Bonekam et al., 2008; Gilmore et al., 2008) primarily suffers from the difficulty in identifying meaningful ROIs. Voxel based analysis is used more commonly than ROI analysis in neuroimaging studies (Chen et al., 2009; Focke et al., 2008; Camara et al., 2007; Snook et al., 2005). The major drawbacks of voxel based analysis include the issues of alignment quality and the arbitrary choice of smoothing extent (Hecke et al., 2009; Ashburner and Friston, 2000; Smith et al., 2006; Jones et al., 2005). With the drawbacks mentioned of the ROI and voxel based analyses, there is a growing interest in the DTI literature in developing fiber tract based analysis of diffusion properties (Smith et al., 2006; O'Donnell et al., 2009; Yushkevich et al., 2008; Goodlett et al., 2009; Zhu et al., 2010b). Statistically, diffusion properties along fiber bundles are functional data and its analysis requires advanced functional data analysis methods (Li and Hsing, 2010; Yao and Lee, 2006; Hall et al., 2006; Ramsay and Silverman, 2005, 2002).

There are several developments on the use of functional data analysis methods for the statistical analysis of diffusion properties along fiber tracts, which all are smoothing first, then estimation" procedures. However, their methods are not capable of delineating the structure of the variability in fiber bundles diffusion properties and for quantifying the uncertainty in the estimated coefficient functions. To specifically address the limitations in (Goodlett et al., 2009; Zhu et al., 2010b), FADTTS presents a functional analysis pipeline for delineating the structure of the variability of multiple diffusion properties along major white matter fiber bundles and their association with a set of covariates of interest, such as age, diagnostic status and gender, in various diffusion tensor imaging studies.

How to use FADTTS GUI

FADTTS GUI is a MATLAB graphical user interface software developed to do data processing by using FADTTS. There are four button groups, which are supposed to be executed in order. The four groups are **Load Raw Data**, **Basic Plots**, **Load Test Data**, and **P-value Plots.** There are three raw data sets, namely, **tract data**, **design data** and **diffusion data**. The test data sets include test design matrix and null hypothesis vector. All data sets must be in .mat. The package includes a sample Matlab code pre_address_data.m on how to set up data. After loading all raw data, GUI will transfer the raw data, estimate the coefficients, do spectral decomposition and estimate confidence bands. Then you can plot the raw tract data, the coefficient functions, spectral decomposition and confidence bands by pushing the corresponding buttons. If you want to do a test, you need to load the test design data. There are two types of test. One is to test individually and the other one is test all the diffusion properties together. Once you loaded the test design data, GUI will display what test type you requested. The test calculation may take a while. After Matlab finishes the computation, GUI will report the global test statistics and p-values. You also have the option to plot the local p-values.

Load Raw Data

tractData: the text file containing (x; y; z) coordinates of all locations on a given fiber tract. The data set should start from one end to the other end. tractData is an L 3 matrix, where L denotes the number of locations. 3 denotes the three coordinates.

designData: the text file containing covariates of interest. Please always include the intercept in the first column. designData is an n p matrix, where n denotes the number of subjects and p denotes the number of covariates.

Diffusion: an m 1 cell containing the names of all fiber diffusion properties files, where m is the number of features. Each fiber bundle diffusion properties should contain an L n matrix. Rows correspond to the columns in **tractData**, while columns correspond to the columns in **designData**.

Basic Plots

Diffusion: plot an m 1 vector of scales for each property, where m is the number of features.

Coefficients: plot a p 1 vector of coefficient functions, p-1 is the number of covariates.

Eigens: plot an L_0 (L_0+1) m matrix of eigenvalues of individual covariance matrix of etas.

CBands: plot a 2p L_0 m matrix of estimated confidence bands.

Load Test Data

CMatrix: an r mp design matrix for characterizing the r linear constraints among mp parameters.

B0vector: an r L_0 vector for hypothesis testing.

P-value Plots: plot p-values.

Click FADTTS button in SSPM interface, or run *FADTTS_GUI* directly, the FADTTS GUI occurs, as shown in Figure 6. Click buttons in order from left to right, following the directions to input files and data, and plot each kind of figure. The results will present in FADTTS Output panel.

For more details about FADTTS, see *FADTTS_refference.pdf*: "*Matlab Tool: Functional Analysis of Diffusion Tensor Tract Statistics* " in FADTTS folder.

Figure 6