Published in final edited form as:

Neuroimage. 2009 April 15; 45(3): 832–844. doi:10.1016/j.neuroimage.2008.12.023.

Tract-Based Morphometry for White Matter Group Analysis

Lauren J. O'Donnell^{a,b}, Carl-Fredrik Westin^b, and Alexandra J. Golby^{a,b}

^aGolby Surgical Brain Mapping Laboratory, Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston MA, USA

^bLaboratory for Mathematics in Imaging, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston MA, USA

Abstract

We introduce an automatic method that we call tract-based morphometry, or TBM, for measurement and analysis of diffusion MRI data *along* white matter fiber tracts. Using subject-specific tractography bundle segmentations, we generate an arc length parameterization of the bundle with point correspondences across all fibers and all subjects, allowing tract-based measurement and analysis. In this paper we present a quantitative comparison of fiber coordinate systems from the literature and we introduce an improved optimal match method that reduces spatial distortion and improves intra- and inter-subject variability of FA measurements. We propose a method for generating arc length correspondences across hemispheres, enabling a TBM study of interhemispheric diffusion asymmetries in the arcuate fasciculus (AF) and cingulum bundle (CB). The results of this study demonstrate that TBM can detect differences that may not be found by measuring means of scalar invariants in entire tracts, such as the mean diffusivity (MD) differences found in AF. We report TBM results of higher fractional anisotropy (FA) in the left hemisphere in AF (caused primarily by lower λ_3 , the smallest eigenvalue of the diffusion tensor, in the left AF), and higher left hemisphere FA in CB (related to higher λ_1 , the largest eigenvalue of the diffusion tensor, in the left CB). By mapping the significance levels onto the tractography trajectories for each structure, we demonstrate the anatomical locations of the interhemispheric differences. The TBM approach brings analysis of DTI data into the clinically and neuroanatomically relevant framework of the tract anatomy.

Keywords

diffusion MRI; tractography; white matter; asymmetry

1 Introduction

By measuring water diffusion in the brain, diffusion tensor MRI (DTI) gives information about the orientation and integrity of fiber tracts, the major neural connections in the white matter. DTI tractography (Basser et al., 2000) follows directions of maximal water diffusion to estimate the trajectories of the fiber tracts. Clinical and neuroscientific questions regarding white matter pathways may be addressed by analyzing DTI data in regions of specific white matter tracts

© 2008 Elsevier Inc. All rights reserved.

odonnell@bwh.harvard.edu (Lauren J. O'Donnell).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

(Johansen-Berg and Behrens, 2006). Many tractography-based analyses (Pagani et al., 2005; Heiervang et al., 2006; Jones et al., 2006; Wakana et al., 2007) have calculated the mean in the entire tract of a scalar value such as the fractional anisotropy (FA) or the mean diffusivity (MD). Other studies have measured mean FA or other scalars in regions of interest (ROIs) within tracts (e.g. Pierpaoli et al. (1996); Kubicki et al. (2003)). Due to anatomical factors such as crossing fibers or nearness to cerebrospinal fluid or gray matter, as well as tissue microstructural factors such as packing densities and axon diameters (Pierpaoli et al., 1996), FA and MD vary spatially along tract trajectories (Xue et al., 1999). Thus analysis of their mean values may not be optimal for localization of group differences in the white matter.

Therefore we propose to perform quantitative analysis of DTI data along the white matter tracts, an approach that we call *tract-based morphometry* (TBM). In this work we present an automatic TBM method for white matter analysis in groups. First DTI tractography from multiple subjects is analyzed to produce common arc length coordinates (point correspondences) across fibers from all subjects, then statistical analysis is performed in this tract-based coordinate system. Our inspiration for naming TBM was the analogy between the voxel-based morphometry method (Ashburner and Friston, 2000) in which local statistical analyses are performed on features derived from scalar MRI (gray and white matter concentrations), and our local statistical analyses of features derived from DTI. For simplicity, to develop the TBM method in this paper we focus only on measurements of white matter microstructural morphometry such as FA, however macroscopic morphological features regarding the entire tract may also be studied with the TBM method. This includes features measured for individual fibers, such as curvatures, as well as shape features (widths, volumes, areas) of the entire bundle.

Other groups have described methods for tract-based analysis of DTI. The methods differ in the coordinate system that is overlaid on the fiber tracts (voxel-based/fiber-based/skeleton-based), in whether they are able to handle data from multiple subjects, and in whether statistical results have been demonstrated in group data.

The first class of methods performs voxel-based measurements along tracts. These methods are limited to fiber tracts that are approximately perpendicular to the axial, sagittal, or coronal image planes. Several groups have quantified diffusion in this way, for example along the trajectory of the corticospinal tract on consecutive axial images (Reich et al., 2006; Lin et al., 2006; Wakana et al., 2007), or in the anterior thalamic radiation, inferior longitudinal fasciculus (Wakana et al., 2007), and cingulum bundle (Gong et al., 2005) using coronal images. In the slice-based measurement framework, the problem of registering the measured data across subjects has been addressed using anatomical landmarks along tracts (Wakana et al., 2007; Gong et al., 2005). Little statistical analysis has been performed along the fiber tracts, with the exception of the cingulum study of interhemispheric FA differences (Gong et al., 2005).

Another class of methods generates coordinate systems based on fibers, and thus is more able to handle arbitrarily shaped fiber tracts. These methods find trajectory correspondences along the length of the fibers. One approach for fiber tract parameterization by arc length employed human interaction to define a corresponding point on all fibers, then assumed point correspondences based on the distance along the fiber from the selected point (Corouge et al., 2006). Other work assumed endpoint correspondences were sufficient to align fibers, even across subjects (Batchelor et al., 2006). Another approach, demonstrated on the corona radiata and cingulum, estimated three level sets to produce a fiber coordinate system (Niethammer et al., 2006). A principled statistical bundle model with point correspondences along fibers was constructed using a unified method for fiber clustering and measurement (Maddah et al., 2008). These more advanced fiber coordinate system methods have so far been demonstrated on fibers from single subjects, but not extended to group data.

A different approach for statistical analysis of DTI uses spatial skeletons to define locations likely to correspond to central parts of fiber bundles. In the tract-based spatial statistics method, in each subject locally high FA values were aligned to a group FA skeleton and voxel-based morphometry was performed (Smith et al., 2006). However, by working in a voxel coordinate system the method could mix information from nearby but differently oriented tracts. Finally, a recent method that blended the spatial skeleton and fiber approaches performed group analysis of MD by sampling each subject's data using a medial model derived from tractography in atlas space (Yushkevich et al., 2008).

Inspired by the highly successful voxel-based morphometry method (Ashburner and Friston, 2000), we have coined the term *tract-based morphometry* (TBM) to describe the statistical analysis of medical image data in a white matter tract coordinate system. Our TBM method can be simply understood in terms of its inputs and outputs. The input to the method is a "bundle" of tractography data from all subjects to be studied. (We will refer to each trajectory as a "fiber" and to each anatomical structure as a "bundle".) The output from the method includes a mean fiber trajectory for the bundle, as well as diffusion measurements and statistical analysis in the coordinate system of the mean fiber. In contrast with other methods for white matter analysis in groups, in which diffusion data is sampled in locations defined by an atlas (e.g. Smith et al. (2006); Yushkevich et al. (2008)), in TBM we employ subject-specific tractography segmentations.

The methodological contributions of this work include: quantitative testing of tract-based coordinate systems from the literature, an improved optimal match method for arc length parameterization of fibers that reduces spatial distortion and variability of measured data, a method for generating arc length correspondences across hemispheres, and an investigation into the appropriate size scale for statistical analysis of diffusion data along tracts.

This paper is divided into two main sets of experiments. The first experiments are performed to optimize our TBM implementation. Then, using the final implementation, we demonstrate the TBM method using a study of hemispheric asymmetry of FA, MD, and the three eigenvalues of the diffusion tensor ($\lambda_1 \geq \lambda_2 \geq \lambda_3$) in the cingulum bundle (CB) and arcuate fasciculus (AF) of normal right-handed subjects. TBM results regarding FA in the cingulum replicate those measured manually in Gong et al. (2005), while the other data are quantified for the first time along the AF and CB tracts.

2 Methods

Our TBM method has five steps: fiber bundle definition and registration, prototype fiber calculation, bundle arc length parameterization, measurement of descriptive statistics, and statistical analysis in the group. In the following sections we explain each step, including experiments performed to optimize its implementation. Finally we describe a TBM study of interhemispheric diffusion differences in the cingulum bundle (CB) and arcuate fasciculus (AF).

2.1 Data

DTI data (EPI, 30 directions + 5B0 images, 2.5mm isotropic voxels) from 35 normal right-handed subjects was seeded with tractography in the whole brain, where $c_L \geq 0.2$ on a 1.5mm grid using random jitter of seed points of up to 0.75mm. Runge-Kutta order two interpolation was used with a step size of 0.5mm. FA images were generated.

2.2 Fiber Bundle Definition and Registration

To enable group analysis of tractography, we chose to align tracts from all subjects using a state-of-the-art unbiased group image registration method. We performed affine group registration of subject FA images using the congealing method (Zollei et al., 2005), then we applied the resulting transformation to the fibers from whole-brain tractography. Fiber bundles were then segmented simultaneously in all subjects via group fiber clustering (O'Donnell and Westin, 2007). Corresponding clusters across hemispheres were automatically identified during the clustering process.

Bilateral multisubject fiber clusters in the arcuate fasciculus (AF) and cingulum regions (CB) were chosen to give bundles for further analysis (Figure 1). The bundles were visually inspected for each subject, and 10 subjects (AF) were excluded due to missing or very short fibers. The total numbers of fibers per bundle were: 2459 (AF L), 2408 (AF R), 10533 (CB L), and 7464 (CB R). The means and standard deviations of the numbers of fibers per subject were: 98 ± 65 (AF L), 96 ± 75 (AF R), 301 ± 54 (CB L), and 213 ± 62 (CB R).

2.3 Prototype Fiber Calculation

To place coordinates on all fibers in a bundle, the method that we have pursued in this work is the propagation of the coordinates of a representative fiber to all other fibers. This idea was proposed by Maddah et al. (2006) and Batchelor et al. (2006), and was applied to fibers from multiple subjects (O'Donnell et al., 2007). We call this representative fiber the *prototype fiber* for the bundle ¹.

2.3.1 Choice of Prototype—In order to investigate the sensitivity of TBM to the choice of prototype and determine a good method for its selection, we compared three possible alternatives for calculating the prototype fiber. Method (1) chose the longest fiber in the bundle as proposed in Maddah et al. (2006), except that here fibers from all subjects were considered. Method (2) attempted to avoid long outlier fibers by choosing the fiber with the greatest length weighted by local “fiber density”. To implement this, fiber density was calculated as the number of tractography trajectories (from all subjects) passing through each voxel, the density was sampled at the (equally spaced) points defining each fiber, and finally the density was integrated along each fiber. Method (3) chose the fiber most representative of the bundle's trajectory across all subjects according to a fiber affinity metric, as proposed in O'Donnell et al. (2007). All fibers were spectrally embedded as points, producing a point cloud whose mean would correspond to the most representative fiber trajectory. Due to the high sample size, the mean was approximated with the point closest to the mean, whose corresponding fiber was chosen as the prototype.

2.3.2 Symmetric Coordinates Across Hemispheres—We developed a simple method to enable comparison of tracts across hemispheres by generating symmetric bundle prototypes. First, each bilateral fiber bundle was augmented with its reflection across the midsagittal plane, generating a symmetric bundle. Then prototype selection was performed, giving a fiber likely to be representative of the bundle's trajectory in both hemispheres. Finally, this prototype was copied and reflected across the midsagittal plane to give two symmetric prototype fibers, with known point correspondences across hemispheres. In the following steps the arc length parameterizations for all fibers naturally corresponded across hemispheres, due to the initial correspondence of the prototypes.

¹Note that because within a bundle fibers have varying lengths and their point correspondence is not known a priori, it is not possible to directly average fiber coordinates to calculate a mean fiber.

2.3.3 Prototype Generation Experiment—The three types of prototype fibers (Section 2.3.1) were generated for AF and CB. For investigation of the appropriate spatial discretization, additional prototype fibers were generated by downsampling, giving prototypes with discretizations of 1mm, 2mm, 3mm, 4mm, 6mm, 8mm, and 10mm.

2.4 Bundle Arc Length Parameterization

In this step, arc length coordinates defined by the prototype fiber were propagated to all other fibers by solving a point correspondence or matching problem. In order to investigate the sensitivity of the results to the choice of matching method, we implemented three versions (described in more detail below). The optimal point match (OP) method performed optimal assignment based on a cost function of our design. The distance map (DM) method, the current state of the art in the literature, searched perpendicular to the prototype (Maddah et al., 2008). Finally, we implemented the most well-known method from the literature. The cutting plane (PL) method used the prototype only to define a plane perpendicular to its midpoint; then coordinates were assigned according to the distance along each fiber from this plane (Corouge et al., 2006). Our investigation into methodology was motivated by the intuition that the performance might differ for example in regions of fiber spreading or high fiber curvature (Figure 2).

With any matching method, the logical constraint is that each fiber point should match to only one point on the prototype fiber (should map to only one arc length). This is important to allow inversion of the mapping: statistical significances are calculated versus arc length, and this constraint is needed to allow mapping of these p-values back into original subject space.

2.4.1 Optimal Point Match Method (OP)—An optimal unique match was produced for each point on the prototype using the Hungarian algorithm (Kuhn, 1955) for optimal assignment. The cost (or distance) function was designed to reward matches found in directions nearly perpendicular to the prototype. We defined the distance as the metric distance D_{ij} between the prototype point i and the fiber point j

$$D_{ij} = \sqrt{\mathbf{v}_{ij}^T \mathbf{M}_i \mathbf{v}_{ij}} \quad (1)$$

where $\mathbf{v}_{ij}$ is the vector between these points, and the metric tensor $\mathbf{M}_i = \mathbf{t}_i \mathbf{t}_i^T$ where $\mathbf{t}_i$ is the tangent to the prototype at point i . The effect of this metric is to have zero cost to match in directions perpendicular to the prototype, where the distance and tangent vectors are perpendicular, and increasing cost as the angle varies from perpendicular. Effectively we are measuring the distance only in the direction tangent to the prototype. In addition, a threshold on this distance limited matching to points within 40% of the prototype's discretization in mm. This gave a search window (in $\pm$ the tangent direction) of 80% of the spacing between prototype points, appropriately limiting the search region to ensure truly non-optimal matches were left as non-matches (otherwise the Hungarian algorithm would try to match all possible point pairs, producing unexpected results near ends of short fibers). After optimal matching, in the case where the fiber being matched had more densely spaced points than the prototype, as a second step any unmatched fiber points located between successful matches were assigned an arc length using linear interpolation followed by rounding to the nearest arc length.

2.4.2 Distance Map Method (DM)—For comparison we implemented the method that we consider to be the current state of the art: closest point matching via efficient search in a perpendicular direction using distance maps as proposed by Maddah et al. (2008). In this method, first the prototype fiber was used to generate a distance map and a map of the nearest point on the prototype (we represented the maps using volumes of cubic voxels measuring 0.3

$\times 0.3 \times 0.3 \text{ mm}^3$). Then, to find point correspondences for a new fiber, for each point on the fiber the nearest voxel was calculated and the nearest prototype point and the distance to that point were found by direct lookup using the maps.

2.4.3 Cutting Plane Method (PL)—To test our assumption that matching should be performed generally perpendicular to the overall tract shape, we implemented the most well-known method from the literature. The method does not constrain matches according to tract shape; rather a cutting plane is defined that crosses all fibers, giving an origin for the coordinate system (Corouge et al., 2006). Next, coordinates are assigned to each fiber based only on the (positive or negative) distance along the fiber from the cutting plane. In our implementation, instead of defining the cutting plane interactively, we used the plane perpendicular to the midpoint of the highest resolution fiber prototype. For comparison to other methods, after assigning coordinates to all subjects' fibers using the PL method we converted the arc lengths to be all positive numbers, and we quantized the arc lengths to correspond to the discretization of each fiber prototype.

2.4.4 Error Checking—There were two potential matching errors that were checked for each fiber. First, in the DM method, any fibers that were longer than the prototype would repeatedly match to the last point on the prototype. This caused inaccuracies at endpoints and

was addressed by limiting the number of matches of the prototype endpoint to $n = \frac{d_p}{d_f}$ where d_p is the discretization of the prototype in mm and d_f is the discretization of the fiber. This means that if the points on the prototype were 3mm apart, and the fiber points had 0.5mm spacing, then only $\frac{3}{0.5} = 6$ points were permitted to match to the prototype endpoint. More distant points at the ends of fibers were defined to be non-matches.

The second potential error with the OP or DM method was if the arc length coordinates produced were not strictly nondecreasing or nonincreasing (for example if the fiber were to loop back and re-match a section of the prototype). This was uncommon but could happen if for example a fiber took an erratic course, perhaps traversing part of another structure. To ensure this type of error did not affect the analysis, only the longest possible fiber segment of correct (nondecreasing or nonincreasing) arc length coordinates was used for further analysis.

2.4.5 Euclidean Distance Threshold—We implemented a threshold to restrict matches to regions of the fiber within a certain Euclidean distance of the prototype. This was a simple way to remove parts of fibers that exited the main body of the tract structure, likely due to partial voluming (very common in CB). We preferred this approach to the removal of entire fibers, because it was highly probable that the retained parts of the fibers did form part of the structure being analyzed. This was due to the fact that all fibers were automatically segmented, and there were no other anatomical structures with similar shapes near AF or CB that the (partially) errant fibers could be considered part of.

2.4.6 Coordinate System Quality Experiment—To determine the effect of the implementation and its parameters on the TBM results, we applied the OP, DM, and PL methods to generate arc length parameterizations for both bundles (AF and CB) using all fiber prototypes generated in Section 2.3.3. Each fiber was matched to the symmetric prototype fiber in the appropriate hemisphere. This was repeated with Euclidean distance thresholds (on distance from the prototype) of 10mm and 20mm.

To quantify the distortion of the output bundle coordinate system, we defined the matched segment length, $Length_{MS}$, as the length of the input fiber segment that mapped to each output arc length. The intuition behind this measure was that in order to avoid any bias regarding

measurement of data across subjects, a successful coordinate system should have a similar-sized region mapped to each arc length coordinate from each subject, i.e. there should be as little distortion as possible. $Length_{MS}$ was measured and visualized for all arc length parameterizations.

2.5 Measurement of Descriptive Statistics

After arc length correspondences had been calculated the measurement of quantities of interest in the new coordinate system was straightforward. We describe it here using FA as an example. First, for each fiber, a vector of FA values vs. arc length (FA_{fiber}) was calculated using the matching result from the previous section: for each arc length coordinate the point(s) matching to it on the fiber contributed the average of their FA values. Next, for each subject, the mean FA across all fibers was calculated for each arc length coordinate, separately for bundles in the right and left hemisphere, giving the vectors $FA_{bundleR}$ and $FA_{bundleL}$. Finally, in the group the pointwise (at each arc length) mean and standard deviation of the subject FA_{bundle} vectors was calculated. We applied the same method to the other scalar invariants and then to the (x,y,z) coordinates for each arc length, giving a per-subject and group mean fiber.

Note that to study FA (MD, etc.) along fibers we performed calculations with FA measurements, not tensors. This made sense because the subjects' tensor orientation information was taken into account by using the fiber coordinate system, so the tensor eigensystem should not further impact the study of the bundle's diffusion properties.

To select the region of data suitable for group statistics along the fibers, the common tract region across all subjects was identified. Using the per-subject vector of mean FA calculated in the previous steps, the common region was defined as the range of coordinates with FA values from all subjects.

2.5.1 Data Variance vs. Method Experiment—To determine how the implementation affected the quality of the data alignment both within and across subjects, we analyzed the mean and variance of the output bundle data for all methods and size scales. First all mean fibers were compared, then the coefficient of variance of FA was quantified both within and across subjects for each method.

2.6 Statistical Analysis in the Group

In this paper, we have applied TBM to a study of interhemispheric differences. Thus within the tract region common across all subjects, for each arc length, a two-tailed paired t-test was employed, where the pairs were the left and right hemisphere in each subject.

2.6.1 Multiple Comparison Correction—To correct for multiple comparisons along the fiber, permutation testing (Nichols and Holmes, 2002) was performed by exchanging the labels for right and left hemisphere within each subject. 10000 permutations were used to estimate the distribution of the summary statistic (the maximum value along the arc length of the absolute value of the paired t-test) under the null hypothesis (that the values in the hemispheres do not differ). Using this estimated distribution, p-values were calculated for each arc length coordinate.

2.6.2 Statistical Significance vs. Scale Experiment—The tractography trajectories were calculated at subvoxel step sizes but the group registration was not expected to be this precise, therefore it did not make sense to perform statistical analysis at such a high resolution. Unlike in voxel-based analysis, our input data were assumed to all belong to the same anatomical structure. Because there was no inherent mixing of anatomical structures problem in blurring or downsampling the bundle data, we were free to choose a discretization that was

comparable in size to the significant regions in the data. However, similar to smoothing in fMRI, there was a tradeoff between spatial resolution and significance. In order to empirically establish an appropriate size scale we repeated the FA data analysis using all discretizations of prototype fiber described in Section 2.3.3.

2.7 Visualization of Results

A novel method was employed for visualization of group results. Output p-values were overlaid on a sample of input fibers from all subjects, allowing anatomical visualization of group differences. The p-values were calculated in the coordinate system of the mean fiber, so to map them back to the input fiber space, each input fiber point was painted with the p-value for the arc length coordinate to which it had been assigned.

2.8 Final TBM Experiment in AF and CB

The results of all previous experiments were used to define the best TBM implementation (from among the methods studied), and interhemispheric differences were studied in AF and CB. First, the means of the scalar invariants FA, MD, λ_1 , λ_2 , and λ_3 were measured in the *entire* AF and CB bundles and significant interhemispheric differences were investigated with a two-tailed paired t-test. Next the final TBM analysis was performed to localize interhemispheric differences along fibers. Within the tract region that was common across subjects, for each arc length coordinate, interhemispheric differences in FA, MD, λ_1 , λ_2 , and λ_3 were studied using two-tailed paired t-tests. Multiple comparisons along the fiber were corrected with permutation testing.

3 Results

The results section is divided into two parts: first we present results from the experiments related to the implementation of TBM, followed by the results of the final AF and CB study using the optimal implementation.

3.1 TBM Implementation Experiments

Results are organized in order according to the steps in the method.

3.1.1 Prototype Fiber Calculation—The results of the fiber prototype generation experiment (Section 2.3.3) are shown in Figure 3 (AF) and 4 (CB). The fiber-density-weighted longest fiber was the most representative of the anatomical trajectory in both AF and CB. In both cases the longest fiber was an outlier that traversed additional regions not part of the anatomical structure. The fiber chosen based on spectral embedding was representative of the structure but was very short in the cingulum.

3.1.2 Bundle Arc Length Parameterization—Here we present results of the coordinate system quality experiment (Section 2.4.6). Output arc length parameterizations generated using the fiber-density-weighted longest fiber prototype and all alignment methods are demonstrated in Figure 5 and Figure 7. The coordinates from the PL method were not constrained to increase/decrease perpendicular to the fibers, and this is especially noticeable in AF (Figure 5). The DM method was sensitive to curvature of the prototype fiber which could introduce shocks in the arc length labeling. At fine discretizations the OP method was more successful than the DM method, due to its objective of assigning all arc length coordinates to each fiber. However, at discretizations 6mm and above, the OP and DM two methods produced visually similar results.

In both AF and CB a threshold on Euclidean distance from prototype was found to be necessary, and was employed for all results given in this paper. In the arcuate, the 20mm threshold successfully limited inclusion of outlier fiber regions, while the 10mm threshold was too near

the bundle radius. In the cingulum, the 10mm threshold was preferred due to the narrower bundle radius.

The matched segment length, $Length_{MS}$, (indicating distortion as described in Section 2.4.6) is displayed in Figure 6 and Figure 8. Due to its better handling of curved prototype fibers, the optimal match (OP) method generally shows reduced distortion compared to the distance map (DM) method. The PL method assigns coordinates based on fiber length, thus it produces no distortion of lengths and is not included in these figures.

3.1.3 Measurement of Descriptive Statistics—Measurements from all fibers for a single subject and from all subjects in the group are plotted in Figure 9 and Figure 10. For all prototypes and matching methods (Section 2.5.1) group mean fibers were found to be similar (Figure 11). An effect of implementation was seen on the coefficient of variation of FA within subject/in group, where the optimal point match (OP) method generally reduced variability of the FA data (Figure 12 and Figure 13). Note that these variability measurements were made only in the tract region common across subjects as (Figure 11).

3.1.4 Statistical Analysis in the Group—Results from the statistical significance vs. scale experiment (Section 2.6.2) demonstrate consistent results across analysis scales (Figure 14 and Figure 15). Empirically we determined that a discretization of 4mm was suitable for further analyses because it represented well the pattern of the FA plot from smaller size scales (rightmost column in Figure 14 and Figure 15) while increasing statistical significance, especially in the AF.

3.1.5 Final Implementation Choices—To summarize the results of previous experiments used to determine the final method implementation, the fiber-density-weighted longest prototype was

3.2 Final TBM Results for Arcuate and Cingulum

The whole bundle results are shown in Table 1, indicating significant inter-hemispheric differences in FA, λ_1 , λ_2 , and λ_3 but not in MD.

TBM results from statistical analysis in the group are presented in Figure 16 and Figure 17, showing tract-based measurements and locations of significant inter-hemispheric differences in the cingulum bundle and arcuate fasciculus, respectively.

4 Discussion

4.1 Discussion of Results in AF and CB

The AF, traditionally thought to connect Broca's and Wernicke's language areas, has known asymmetries in tractography-based measures of size (Powell et al., 2006; Vernooij et al., 2007) and mean FA (Powell et al., 2006). Other language-related asymmetries include prevalent leftward functional language lateralization and perisylvian structural differences (Toga and Thompson, 2003). The CB also has known asymmetry in FA (Park et al., 2004; Gong et al., 2005). Our findings agree with known FA asymmetries and we expand on previous work by giving a more complete picture of diffusivity along the tracts than has been previously measured. In particular, we have found that the FA asymmetries can be attributed in AF primarily to increased λ_3 in the right hemisphere, while in CB the pattern of FA asymmetry parallels a region of higher λ_1 in the left hemisphere. Interestingly, examination of the eigenvector associated with λ_3 in AF shows good consistency of its orientation across fibers, subjects, and hemispheres, implying that the diffusion differences could relate in some meaningful way to the overall geometry of this tract.

TBM results can be compared to other types of analysis that have been applied to the study of white matter asymmetries. FA asymmetry has been studied using VBM, where differences were found in AF (left FA > right FA) but not in CB in one study (Buchel et al., 2004), and in both CB (left FA > right FA) and AF (surprisingly, left FA < right FA) in another study (Park et al., 2004). The TBM results can also be compared to our study of asymmetry of mean values in entire tracts which failed to detect all differences that were seen using TBM. In AF, there was no difference found in mean MD (Table 1), but the pattern in the left and right hemispheres was significantly different (Figure 16). Finally, our automatic TBM method can be compared to a more labor-intensive tractography study of asymmetry. The study of FA laterality differences measured using a manually segmented cingulum-specific arc angle coordinate system (Gong et al., 2005) produced very similar results to those reported here, localizing differences in FA primarily to the anterior cingulum.

4.2 Discussion of Method

In this work we focused on performing quantitative evaluation of the coordinate system placed on the fibers, because this is the most fundamental step toward making accurate measurements. Our within-subject results (Figure 12) showed that the proposed optimal point match (OP) method performs well in comparison with other methods. However improved cross-subject alignment is of interest for future work, as the high-frequency variations seen in individual subjects (Figure 9 and Figure 10) could be of interest and are smoothed in the group average.

Regarding the impact of registration on the TBM pipeline, the affine registration (applied after tractography was performed) was adequate for segmentation of tractography into bundles. Then, in TBM there were two registration problems of interest: the first was alignment of all fibers within a subject, and the second was alignment of the measurements across subjects. The first problem was addressed by the optimal point match method (OP), while the second cross-subject alignment was produced both by the initial affine congealing registration and the OP coordinate system generation. In the case of anatomical abnormalities such as atrophy, if the effect is limited to a “thinning” of the tract and the core of the bundle retains its overall shape, this should be handled relatively well by the current method. In atrophy, because the method analyzes fibers not voxels, TBM would limit potential problems due to partial voluming near enlarged ventricles. However, pathology that changes the shape of a tract in some subjects could invalidate the assumption that one mean fiber (one coordinate system) can represent the population. In this case a registration with higher degrees of freedom would be necessary.

The method proposed for generating a symmetric coordinate system may be too simplistic for some tracts due to anatomical differences between hemispheres. The assumption of the method is that the underlying “mean shape” across hemispheres is similar enough that the coordinate system generation produces anatomically corresponding regions across hemispheres. This does not directly assume low variability of quantities under study such as FA, or even low variability of bundle parameters such as number of tractography fibers produced or tract volume. It just assumes that the bundle core has roughly the same shape across hemispheres. From visual inspection of the data (see for example Figure 1 and Figure 11) we judge this assumption to be valid in the cingulum bundle, but due to known shape differences in the language areas (Toga and Thompson, 2003) the assumption may be overly simplistic for evaluating the arcuate fasciculus (AF). Moving forward, one possible way to address this issue is the creation of hemisphere-specific coordinate systems, followed by an additional registration step to align the vectors of data across hemispheres.

In this paper we have developed the TBM method for analysis of DTI-derived scalar invariants like FA that measure the microstructural morphometry of the tract. However the identical TBM machinery also allows study of macrostructural tract morphology such as widths, volumes, areas, and curvatures (note that the size scale of shape variation need not correspond to the size

scale of FA variability). For example we propose estimating the cross-sectional area of the bundle at each arc length by first projecting all points at that arc length into the plane locally perpendicular to the mean fiber, then calculating the area within the convex hull (polygon) containing the points. This area and related quantities can be measured along the tract for each subject and can be compared in the group. Further investigation into the whole-tract morphological measures may help determine the validity of interhemispheric coordinate systems, by quantifying for example whether the sharpest bend in the AF is aligned across hemispheres. However, because all morphological measures could be calculated in the original or group-registered coordinate systems, some investigation is needed to determine which is appropriate. For example, it is likely that widths should be calculated after registering to correct for head size, but it is not clear that affine transformations should be applied before measuring curvatures.

Regarding the statistical methods employed, in order to investigate the scale of significant regions in the FA data, we formed no hypothesis about the magnitude of FA differences or the expected cluster size. However, the neighboring tract-based measurements are clearly not independent, and detection of subtle elevations along the entire tract (as seen for example in λ_2 in Figure 17) could be improved by statistical analysis based on suprathreshold cluster size (Nichols and Holmes, 2002).

An important strength of the TBM method relative to other methods for white matter analysis in groups (Smith et al., 2006; Yushkevich et al., 2008) is that we use a subject-specific segmentation of tractography, so we are more confident that we are comparing the same anatomical region across subjects. However TBM can benefit from improvements in tractography, for example by using fibers from two-tensor (Qazi et al., 2008; Tuch et al., 2002) or diffusion spectrum tractography (Wedeen et al., 2008). This may increase robustness regarding structures such as AF that are difficult to trace, and allow inclusion of more terminal regions of tracts that are not currently present in all subjects and thus are excluded. Perhaps more interestingly, a higher-order model could provide tract-specific parameters such as the tract's fraction in the voxel (Peled et al., 2006), reducing the impact of crossing tracts on the analysis.

A possible issue with the method is the interaction between tensor anisotropy used for tractography restriction to white matter, and the scalar measures extracted from the tensors along the tracts. This is important to consider when interpreting results, however as we have not analyzed the very ends of the tracts, the importance of this issue is likely reduced.

Finally, one should note that tractography can only estimate the location of a fiber tract of interest. This is due to the greatly different size scales of axons (micron diameter) and image voxels (mm diameter). For the same reason, the scalar invariants measured in this work are affected by multiple tracts that may be present in a voxel. Thus the asymmetries measured in AF and CB may be caused by anatomical differences in those tracts, as well as by differences in nearby or crossing tracts.

5 Conclusion

We have introduced an automatic method that we call tract-based morphometry, or TBM, for measurement and analysis of diffusion MRI data along fiber bundles. Overall, the TBM approach brings analysis of DTI data into the clinically and neuroanatomically relevant framework of the tract anatomy.

Acknowledgments

This work has been supported by NIH grants U54EB005149, R01MH074794, P41RR13218, K08NS048063 and U41RR019703. Thank you to Susumu Mori at JHU for the diffusion MRI data (R01AG20012 / P41RR15241). Thanks to Lilla Zollei for the congealing registration code.

References

- Ashburner J, Friston K. Voxel-based morphometry-the methods. *NeuroImage* 2000;11(6):805–821. [PubMed: 10860804]
- Basser P, Pajevic S, Pierpaoli C, Duda J, Aldroubi A. In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine* 2000;44:625–632. [PubMed: 11025519]
- Batchelor PG, Calamante F, Tournier J-D, Atkinson D, Hill DLG, Connelly A. Quantification of the shape of fiber tracts. *Magnetic Resonance in Medicine* 2006;55(4)
- Buchel C, Raedler T, Sommer M, Sach M, Weiller C, Koch M. White Matter Asymmetry in the Human Brain: A Diffusion Tensor MRI Study. *Cereb. Cortex* 2004;14(9):945–951. [PubMed: 15115737]
- Corouge I, Fletcher P, Joshi S, Gouttard S, Gerig G. Fiber tractoriented statistics for quantitative diffusion tensor MRI analysis. *Medical Image Analysis* 2006;10(5):786–798. [PubMed: 16926104]
- Gong G, Jiang T, Zhu C, Zang Y, Wang F, Xie S, Xiao J, Guo X. Asymmetry analysis of cingulum based on scale-invariant parameterization by diffusion tensor imaging. *Human Brain Mapping* 2005;24(2): 92–98. [PubMed: 15455461]
- Heiervang E, Behrens T, Mackay C, Robson M, Johansen-Berg H. Between session reproducibility and between subject variability of diffusion MR and tractography measures. *NeuroImage* 2006;33(3):867–877. [PubMed: 17000119]
- Johansen-Berg H, Behrens T. Just pretty pictures? What diffusion tractography can add in clinical neuroscience. *Curr Opin Neurol* 2006;19(4):379–385. [PubMed: 16914977]
- Jones DK, Catani M, Pierpaoli C, Reeves SJ, Shergill SS, O'Sullivan M, Golesworthy P, McGuire P, Horsfield MA, Simmons A, Williams SC, Howard RJ. Age effects on diffusion tensor magnetic resonance imaging tractography measures of frontal cortex connections in schizophrenia. *Human Brain Mapping* 2006;27:230–238. [PubMed: 16082656]
- Kubicki M, Westin C-F, Maier SE, Mamata H, Frumin M, Ernst-Hirschfeld H, Kikinis R, Jolesz FA, McCarley RW, Shenton ME. Cingulate Fasciculus Integrity Disruption in Schizophrenia: A Magnetic Resonance Diffusion Tensor Imaging Study. *Biological Psychiatry* 2003;54:1171–1180.
- Kuhn HW. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly* 1955;2:83–97.
- Lin F, Yu C, Jiang T, Li K, Li X, Qin W, Sun H, Chan P. Quantitative analysis along the pyramidal tract by length-normalized parameterization based on diffusion tensor tractography: Application to patients with relapsing neuromyelitis optica. *NeuroImage* 2006;33(1):154–160. [PubMed: 16919971]
- Maddah M, Grimson W, Warfield S. Statistical modeling and EM clustering of white matter fiber tracts. *International Symposium on Biomedical Imaging (ISBI)* 2006:53–56.
- Maddah M, Grimson WEL, Warfield SK, Wells WM. A unified framework for clustering and quantitative analysis of white matter fiber tracts. *Medical Image Analysis* 2008;12(2):191–202. [PubMed: 18180197]
- Nichols T, Holmes A. Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping* 2002;15:1–25. [PubMed: 11747097]
- Niethammer, M.; Bouix, S.; Westin, C-F.; Shenton, ME. Fiber bundle estimation and parameterization; *Proc. Medical Image Conference and Computer Assisted Intervention (MICCAI)*; 2006 Oct. p. 252–259.
- O'Donnell L, Westin C-F. Automatic Tractography Segmentation Using a High-Dimensional White Matter Atlas. *IEEE Trans. Medical Imaging* 2007;26(11):1562–1575.
- O'Donnell, LJ.; Westin, C-F.; Golby, AJ. Tract-Based Morphometry; *Proc. Medical Image Conference and Computer Assisted Intervention (MICCAI)*; 2007. p. 161–168.

- Pagani E, Filippi M, Rocca M, Horsfield M. A method for obtaining tract-specific diffusion tensor MRI measurements in the presence of disease: application to patients with clinically isolated syndromes suggestive of multiple sclerosis. *NeuroImage* 2005;26(1):258–265. [PubMed: 15862226]
- Park HJ, Westin C-F, Kubicki M, Maier SE, Niznikiewicz M, Baer A, Frumin M, Kikinis R, Jolesz FA, McCarley RW, Shenton ME. White matter hemisphere asymmetries in healthy subjects and in schizophrenia: A diffusion tensor MRI study. *NeuroImage* 2004;24:213–223. [PubMed: 15325368]
- Peled S, Friman O, Jolesz F, Westin C-F. Geometrically constrained two-tensor model for crossing tracts in DWI. *Magnetic Resonance Imaging* 2006;24:1263–1270. [PubMed: 17071347]
- Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Chiro GD. Diffusion tensor MR imaging of the human brain. *Radiology* 1996;201:637. [PubMed: 8939209]
- Powell HR, Parker GJ, Alexander DC, Symms MR, Boulby PA, Wheeler-Kingshott CA, Barker GJ, Noppeney U, Koepp MJ, Duncan JS. Hemispheric asymmetries in language-related pathways: A combined functional MRI and tractography study. *NeuroImage* 2006;32(1):388–399. [PubMed: 16632380]
- Qazi AA, Radmanesh A, O'Donnell L, Kindlmann G, Peled S, Whalen S, Westin C-F, Golby AJ. Resolving crossings in the corticospinal tract by two-tensor streamline tractography: Method and clinical assessment using fMRI. *NeuroImage*. 2008In Press
- Reich D, Smith S, Jones C, Zackowski K, van Zijl P, Calabresi P, Mori S. Quantitative Characterization of the Corticospinal Tract at 3T. *AJNR Am J Neuroradiol* 2006;27(10):2168–2178. [PubMed: 17110689]
- Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, Behrens T. Tract-Based Spatial Statistics: Voxelwise Analysis of Multi-Subject Diffusion Data. *NeuroImage* 2006;31:1487–1505. [PubMed: 16624579]
- Toga A, Thompson P. Mapping brain asymmetry. *Nat Rev Neurosci* 2003;4(1):37–48. [PubMed: 12511860]
- Tuch D, Reese T, Wiegell M, Makris N, Belliveau J, Wedeen V. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn. Reson. Med* 2002;48:577–582. [PubMed: 12353272]
- Vernooij M, Smits M, Wielopolski P, Houston G, Krestin G, van der Lugt A. Fiber density asymmetry of the arcuate fasciculus in relation to functional hemispheric language lateralization in both right- and left-handed healthy subjects: A combined fMRI and DTI study. *NeuroImage* 2007;35(3):1064–1076. [PubMed: 17320414]
- Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, Hua K, Zhang J, Jiang H, Dubey P, Blitz A, van Zijl P, Mori S. Reproducibility of quantitative tractography methods applied to cerebral white matter. *NeuroImage* 2007;36(3):630–644. [PubMed: 17481925]
- Wedeen V, Wang R, Schmahmann J, Benner T, Tseng W, Dai G, Pandya D, Hagmann P, D'Arceuil H, de Crespigny A. Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *NeuroImage* 2008;41(4):1267–1277. [PubMed: 18495497]
- Xue R, van Zijl PC, Crain BJ, Solaiyappan M, Mori S. In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging. *Magnetic Resonance in Medicine* 1999;42(6):1123–1127. [PubMed: 10571934]
- Yushkevich PA, Zhang H, Simon TJ, Gee JC. Structure-specific statistical mapping of white matter tracts. *NeuroImage* 2008;41(2):448–461. [PubMed: 18407524]
- Zollei, L.; Learned-Miller, E.; Grimson, WEL.; Wells, MW, III. Computer Vision for Biomedical Image Applications. *ICCV* 2005; 2005. Efficient population registration of 3D data; p. 291-301.

Arcuate Fasciculus (N = 25 subjects)

Cingulum Bundle (N = 35 subjects)

Fig. 1.

Input multisubject fiber bundles in regions of the arcuate fasciculus (top) and cingulum (bottom). A random sample of 25 fibers is shown from each subject. Shades of gray indicate subject numbers. The bundles were automatically segmented from whole brain tractography data.

Fig. 2.

Illustration of the difference between matching perpendicular to the prototype in the distance map method (left) and allowing near-perpendicular matches in the optimal point match method (right). The bold trajectory (p) is the prototype fiber. The dashed lines, representing arc length assignments, connect points on the prototype to their matches on other fibers. Here the optimal point method better handles curvature of the prototype and fibers.

Fig. 3. Symmetric bilateral prototype fibers generated from multisubject arcuate fasciculus fiber bundles using three methods: longest (red), fiber-density-weighted longest (blue), and embedding (light green). View is from left and slightly posterior.

Fig. 4. Symmetric bilateral prototype fibers generated from multisubject cingulum fiber bundles using three methods: longest (red), fiber-density-weighted longest (blue), and embedding (light green). Views are from left (on left) and above (on right).

Fig. 5.

Arc length parameterizations (Section 2.4) in color in AF, in the entire bundle and in a zoomed region (dashed square). A sample of fibers from all subjects is shown. In gray, the leftmost DM image gives an example of regions that were excluded during matching (due to the Euclidean distance threshold and to restricted matching past the prototype endpoint). Note the PL method does not respect the curved tract geometry, and the DM method produces unexpected results at lower size scales.

Fig. 6.

Distortion measure $Length_{MS}$ (Section 2.4.6), the length of the input fiber that was mapped to each output arc length, in AF. A sample of fibers from all subjects is shown. To avoid any bias due to distortion, $Length_{MS}$ should vary as little as possible across all subjects and fibers, for each size scale. For example at the 2mm size scale, ideally the color should be a solid dark blue, indicating $Length_{MS}$ of 2mm everywhere (see colorbar). Generally the OP method produces less distortion: less variation is seen in the $Length_{MS}$ color.

Fig. 7.

Arc length parameterizations (Section 2.4) in color in CB, in the entire bundle and in a zoomed region (dashed square). A sample of fibers from all subjects is shown. In gray, the leftmost DM image gives an example of regions that were excluded during matching. Note that PL is less spatially consistent, and DM is adversely affected by prototype curvature at lower size scales. As fibers leave the structure and before they are truncated, OP is more likely to increment the arc length than DM, leading to subtle differences at 6mm.

Fig. 8.

$Length_{MS}$ (Section 2.4.6), the length of the input fiber that was mapped to each output arc length. To avoid any bias due to distortion, the color should vary as little as possible across all subjects and fibers, for each size scale. Generally the OP method produces less distortion: less variation is seen in the $Length_{MS}$ color. used with a discretization of 4mm, the Euclidean distance threshold from the prototype was set to 20mm for AF and 10mm for CB, and the OP (optimal point matching) method was employed for arc length parameterization.

Single Subject Measurements in Left AF

Group Measurements in Left AF

Fig. 9.

Example FA measurements vs. arc length from a single subject and the group, for all methods, in the left arcuate fasciculus at 1 mm discretization. In (a–c), each curve gives the measurements from one fiber, while in (d) each curve is the average measurement across one subject's fibers.

Single Subject Measurements in Left CB

(a) PL

(b) DM

(c) OP

Group Measurements in Left CB

(d) OP (N=35 subjects)

Fig. 10.

Example FA measurements vs. arc length from a single subject and the group, for all methods, in the left cingulum bundle at 1mm discretization. In (a–c), each curve gives the measurements from one fiber, while in (d) each curve is the average measurement across one subject's fibers.

Fig. 11.

Output mean fibers are highly similar regardless of method or spatial scale. Mean fibers for AF (top) and CB (bottom) were generated using OP, DM, and PL methods, for all discretizations of the fiber-density weighted longest prototype. The left column shows all mean fibers including regions where some subjects contributed no data; the right column shows only the region common across all subjects. The mean fibers in this common region are very similar across methods. The common region results are not very sensitive to the shape of the prototype as the PL method (red-orange fibers) used the prototype only to define the arc length origin.

Fig. 12.

Within-subject variability of FA, compared across methods (cutting plane or PL, distance map or DM, and optimal point match or OP) and size scales (of prototype fiber, shown along x-axis). The quantity plotted is the coefficient of variation computed for each subject (as the variance across all subject fibers of pointwise subject FA divided by pointwise subject mean FA), then averaged across subjects and along the fiber.

Fig. 13.

Within-group variability of FA, compared across methods (cutting plane or PL, distance map or DM, and optimal point match or OP) and size scales (of prototype fiber, shown along x-axis). Each data point represents the average along the fiber of the pointwise group coefficient of variation (FA variance in the group divided by group mean FA).

Multiscale Asymmetry (AF)

Fig. 14.

Spatial resolution vs. statistical significance tradeoff: repetition of inter-hemispheric FA analysis at several scales in AF. For each arc length coordinate, each subject's mean FA was computed for the left and right bundles. The (group) mean and standard error of these per-subject means is shown vs. arc length in mm (left column). The multiple comparison corrected p-value for significant difference is overlaid on a sample of fibers from the group (right column).

Multiscale Asymmetry (CB)

Fig. 15.

Spatial resolution vs. statistical significance tradeoff: repetition of inter-hemispheric FA analysis at several scales in CB. For each arc length coordinate, each subject's mean FA was computed for the left and right bundles. The (group) mean and standard error of these per-subject means is shown vs. arc length in mm (left column). The multiple comparison corrected p-value for significant difference is overlaid on a sample of fibers from the group (right column)..

Arcuate Fasciculus Asymmetry (N = 25)

Fig. 16.

Left and right hemisphere measurements in the group, and their significant differences, in the arcuate fasciculus. For each arc length coordinate, each subject's mean MD, λ_1 , λ_2 , and λ_3 values were computed for the left and right bundles. The (group) mean and standard error of these per-subject means is shown vs. arc length in mm (left column). The multiple comparison corrected p-value for significant difference is overlaid on a sample of fibers from the group (right column).

Cingulum Bundle Asymmetry (N = 35)

Fig. 17.

Left and right hemisphere measurements in the group, and their significant differences, in the cingulum bundle. For each arc length coordinate, each subject's mean MD, λ_1 , λ_2 , and λ_3 values were computed for the left and right bundles. The (group) mean and standard error of these per-subject means is shown vs. arc length in mm (left column). The multiple comparison corrected p-value for significant difference is overlaid on a sample of fibers from the group (right column).

Table 1

Whole bundle quantitative results in both hemispheres for the arcuate fasciculus (AF L and AF R) and cingulum bundle (CB L and CB R).

	AF L	AF R	CB L	CB R
FA	.400 ± .022**	.379 ± .026**	.412 ± .024**	.391 ± .026**
MD	1.07 ± .026	1.07 ± .023	1.07 ± .026**	1.07 ± .030**
λ_1	1.56 ± .044*	1.54 ± .050*	1.60 ± .055**	1.56 ± .055**
λ_2	.951 ± .029	.950 ± .029	.916 ± .026**	.932 ± .032**
λ_3	.685 ± .035**	.713 ± .032**	.695 ± .032**	.711 ± .034**

Values marked with a * are significantly different across hemispheres at the 0.05 level, while ** indicates significance at the 0.005 level. For MD, λ_1 ,

λ_2 and λ_3 , units are $\times 10^{-3} \frac{mm^2}{s}$.