NeuroImage 97 (2014) 284–295

Contents lists available at ScienceDirect

NeuroImage

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

Automatic clustering and population analysis of white matter tracts using maximum density paths

Gautam Prasad [a,b], Shantanu H. Joshi [c], Neda Jahanshad [a,b], Julio Villalon-Reina [a,b], Iman Aganj [g], Christophe Lenglet [h], Guillermo Sapiro [i,j], Katie L. McMahon [k], Greig I. de Zubicaray [l], Nicholas G. Martin [m], Margaret J. Wright [l,m], Arthur W. Toga [a,b,d,e], Paul M. Thompson [a,b,c,d,e,f,*]

[a] Imaging Genetics Center, Institute for Neuroimaging & Informatics, University of Southern California, Los Angeles, CA, USA
[b] Laboratory of Neuro Imaging, Institute for Neuroimaging & Informatics, University of Southern California, Los Angeles, CA, USA
[c] Department of Neurology, University of California Los Angeles, CA, USA
[d] Dept. of Neurology, Psychiatry, Engineering, Radiology, University of Southern California, Los Angeles, CA, USA
[e] Dept. of Ophthalmology, University of Southern California, Los Angeles, CA, USA
[f] Department of Pediatrics, University of Southern California, Los Angeles, CA, USA
[g] Martinos Center for Biomedical Imaging, Radiology Department, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
[h] Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota, Minneapolis, MN, USA
[i] Dept. of Electrical and Computer Engineering, Computer Science, Duke University, NC, USA
[j] Dept. of Biomedical Engineering, Duke University, NC, USA
[k] Center for Advanced Imaging, University of Queensland, Brisbane, Australia
[l] School of Psychology, University of Queensland, Brisbane, Australia
[m] QIMR Berghofer Medical Research Institute, Herston, Australia

ARTICLE INFO

Article history:
Accepted 8 April 2014
Available online 18 April 2014

Keywords:
HARDI
Tractography
MRI
Brain
Clustering
Atlas
Dijkstra
Shortest path
Geodesic distance
Hough
Connectivity
Maximum density path
Curve registration
Longest path

ABSTRACT

We introduce a framework for population analysis of white matter tracts based on diffusion-weighted images of the brain. The framework enables extraction of fibers from high angular resolution diffusion images (HARDI); clustering of the fibers based partly on prior knowledge from an atlas; representation of the fiber bundles compactly using a path following points of highest density (maximum density path; MDP); and registration of these paths together using geodesic curve matching to find local correspondences across a population. We demonstrate our method on 4-Tesla HARDI scans from 565 young adults to compute localized statistics across 50 white matter tracts based on fractional anisotropy (FA). Experimental results show increased sensitivity in the determination of genetic influences on principal fiber tracts compared to the tract-based spatial statistics (TBSS) method. Our results show that the MDP representation reveals important parts of the white matter structure and considerably reduces the dimensionality over comparable fiber matching approaches.

Introduction

Diffusion weighted imaging (DWI) measures the directional diffusion of water through the brain in vivo. By following the dominant directions of diffusion across the brain, whole-brain tractography algorithms can reconstruct the brain's major white matter pathways, extracting a vast number of fibers that are amenable to statistical analysis. We can then study these white matter regions in individuals and populations to better understand disease effects (Daianu et al., 2013; Jahanshad et al., 2012b; Takahashi et al., 2002), changes in brain microstructure and connectivity with age (Abe et al., 2002; Dennis et al., 2012), hemispheric differences (Jahanshad et al., 2010), sex differences (Peled et al., 1998), and genetic influences (Jahanshad et al., 2013a; Kochunov et al., 2010).

High angular resolution diffusion imaging (HARDI) enables a more accurate representation of fiber directions compared to the

* Corresponding authors at: Professor of Neurology, Psychiatry, Radiology, Engineering, Pediatrics, & Ophthalmology Director, USC Imaging Genetics Center, Associate Director, USC Institute for Neuroimaging and Informatics 2001 N. Soto Ave, Keck/USC School of Medicine, University of Southern California, Los Angeles, CA, USA. Fax: (323) 442-7247.
E-mail address: pthomp@usc.edu (P.M. Thompson).

http://dx.doi.org/10.1016/j.neuroimage.2014.04.033
1053-8119/

more standard single-tensor model (Basser and Pierpaoli, 1996). The single-tensor model does not account for fiber crossing or mixing, but the orientation distribution function (ODF) (Tuch, 2004) can be derived from HARDI images to discriminate multiple fibers with different orientations passing through a voxel (Leow et al., 2009; Zhan et al., 2010).

The large number of fibers generated by the tractography algorithms first needs to be clustered according to known anatomical pathways before comparing them across subjects. A wealth of clustering methods has been applied to tractography results including fuzzy clustering (Shimony et al., 2002), normalized cuts (Brun et al., 2004), *k*-means (O'Donnell and Westin, 2005), spectral clustering (O'Donnell et al., 2006), Dirichlet distributions (Maddah et al., 2008), hierarchical clustering (Visser et al., 2011), a Gaussian process framework (Wassermann et al., 2010b), and median filtering (Prasad et al., 2011a). Some of these methods readily benefit from prior anatomical information provided by an atlas of likely locations of the tracts in the brain (Yendiki et al., 2011), suggesting when to split or combine clusters to conform to known anatomy. In one approach (Jin et al., 2011a,b, 2013), several labeled atlases are deformed onto a fiber set extracted from a new subject, and a fiber matching and voting process are used to help decide the anatomical bundles to which the fibers belong.

Following clustering, several methods can be used for fiber bundle matching. (Colby et al., 2011) use a parametric curve-based method to resample fibers in a bundle based on shared seed points and then compute correspondences from the resampling to create a representative path for an individual or group. A similar re-sampling approach is used in a method (Yeatman et al., 2012) that filters fiber bundles to match a probabilistic atlas. (Corouge et al., 2006) analyze fiber bundles by resampling and then aligning them across subjects using Procrustes analysis (Goodall, 1991) to generate a mean shape. (Roberts et al., 2005) apply a density measure derived from tractography results. Their measure (fiber density index; FDi) quantifies the average number of detected fiber paths passing through voxels in a ROI. (Wassermann et al., 2010a) use Gaussian processes to create voxel-wise probability maps of white matter structure. The fiber locations in high density regions of the image space are used by O'Donnell et al. (2009) as a template to align other fibers and compute correspondences. Yushkevich et al. (2008) analyze white matter tracts using deformable geometric medial models that allow for integration of nearby tensor-based features to reduce the dimensionality and improve registration. (Patel et al., 2010) use a fast-marching algorithm to encapsulate white matter tracts in voxel based boundaries, which are then matched using variational techniques.

In contrast to the parameterized methods mentioned above, white matter analysis can also be performed using a voxel-based approach. A popular method known as tract based spatial statistics (TBSS) (Smith et al., 2006), uses a skeletonized representation of white matter and uses nonlinear registration for matching the skeletons. Although it is a very popular approach, TBSS does not explicitly represent tracts that would be recognized by anatomists, and therefore is not guaranteed to produce a consistent labeling of tracts from one brain to another (Schwarz et al., 2013). Although voxel-based methods can also be used to analyze DWI, they are often sensitive to the image registration (Tustison et al., 2012). Most existing white matter analysis techniques focus on nonlinear registration of fractional anisotropy (FA) images as in TBSS (Smith et al., 2006) and voxel-based morphometry (VBM), which can be applied to DWI-derived maps such as FA (Jones et al., 2005). Other approaches that focus on diffusion tensor correspondences are usually based on a global image registration, but a high-dimensional registration of tensor fields may also be used, as can tensor-based statistics (Chiang et al., 2008; Lee et al., 2009; Lepore et al., 2008; Yeo et al., 2009; Zhang et al., 2006). Given the richness of information provided by tractography, it seems advantageous to directly study the fiber tract bundles rather than simply analyzing voxel-based representation.

Approach

Our work adopts a parameterized approach by refining the representation of white matter structure into compact and localized paths, represented as 3D curves. These paths represent the most influential regions in tractography and are used as compact dimensional representations of the fiber bundle. Our method uses an additional local registration of specific white matter regions to fix biases (Tustison et al., 2012) in voxel-based analysis and many of the problems of registration algorithms (Klein et al., 2009) that work on the entire image. Additionally, our approach may offer increased statistical power as it finds shape homologies across different white matter tracts.

Termed the maximum density path (MDP) approach, it incorporates information from tractography-derived fibers by selecting a subset of fiber bundles from a white matter atlas in the same space. We generate a density image from the fiber bundles and use it to create a graph with voxel locations as nodes and fiber density measures as edges. We implement a widely used graph search algorithm to find the MDP between two pre-specified regions of interest (ROI) in the atlas. The MDPs represent fiber bundles that characterize a tract using points of highest density. These compact descriptions of a tract's scale, location, and high-level geometric information are computed for all subjects in a population. We find correspondences across the paths by bringing them into the same space using geodesic curve registration. Finally, the average MDP for a given population is computed using a nonlinear iterative method. As an example, we use our method to determine genetic influences on white matter tracts based on a large cohort of over 565 twin subjects scanned using HARDI at 4-Tesla. We compare the results to those obtained by the more standard TBSS method.

MDPs have been used as one tool for pilot studies of sex differences and a variety of diseases (Nir et al., 2012; Prasad et al., 2011b). In the current study we explicate the technical details of the method, validate its repeatability, compare it to the widely used TBSS, and use MDPs to study heritability along with genetic associations. The main contributions of this work are as follows:

- Fiber tract bundles are represented by compact reduced dimensional representations known as maximum density paths (MDPs).
- MDPs are represented by vector valued functions and are analyzed in an intrinsic and invariant manner.
- Shape matching between MDPs is achieved using geodesic curve registration that not only yields smooth deformations between MDPs, but also provides shape distances between them.
- Group analysis of MDPs is conveniently performed using an intrinsic statistical framework that enables the computation of shape averages and their first order variations.
- Fiber bundle analysis via MDPs is used to identify highly heritable regions in the white matter tissues in twin subjects and is also used to show genetic associations.

Materials and methods

This section describes important steps starting with the extraction of fibers using HARDI tractography, clustering of fibers using a white matter ROI atlas, representation and matching of fiber bundles using MDPs, and finally, statistical analysis of MDPs in a population. The schematic pipeline outlining the extraction and representation of MDPs is shown in Fig. 1, whereas the workflow for statistical group analysis is shown in Fig. 2.

HARDI tractography using the Hough transform

We use a global tractography algorithm (Aganj et al., 2011) to extract fibers from HARDI images.

The algorithm uses extensive information provided by HARDI at each voxel, parametrized by the orientation distribution function (ODF).

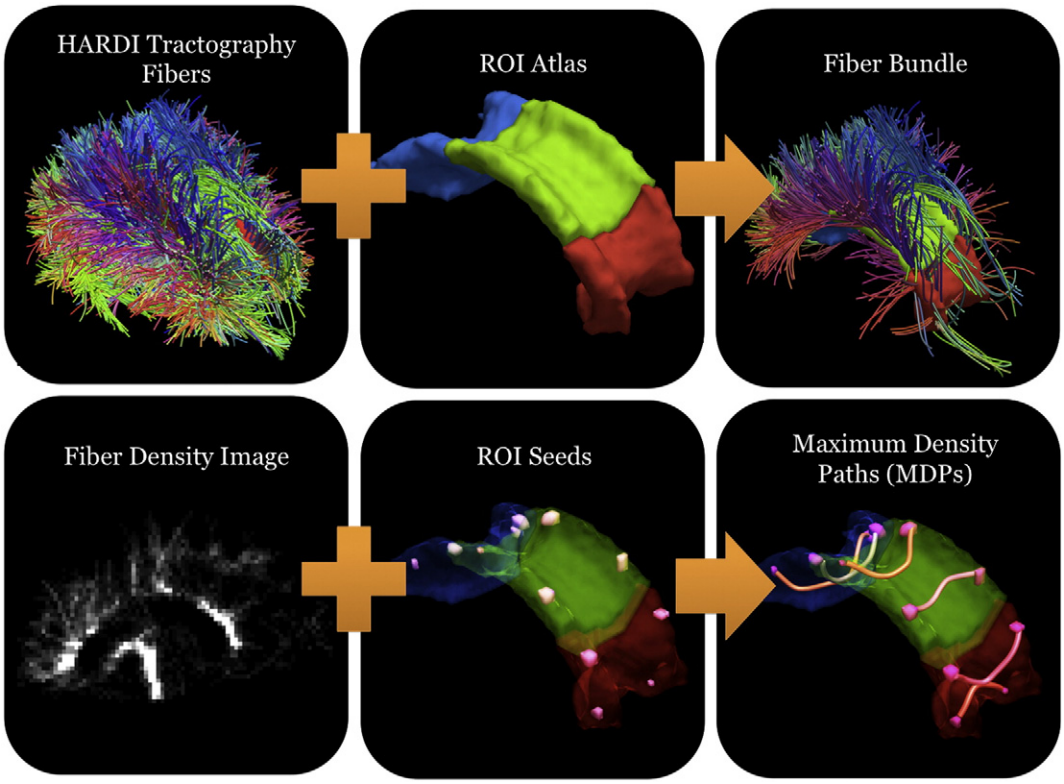

Fig. 1. Schematic of the pipeline for extraction, clustering, and representation of maximum density paths (MDPs) for a single subject. The first panel shows the fibers from our HARDI global tractrography method. The co-registered region of interest (ROI) atlas is used to select a fiber bundle representing a particular white matter tract. The resulting fiber bundle is then converted to a volumetric density image, which is transformed into a graph. Selected seed points in the image form the nodes of the graph, that are used to compute maximum density paths. The maximum density path compactly summarizes a given white matter structure and enables specific matching of these regions across subjects using curve registration.

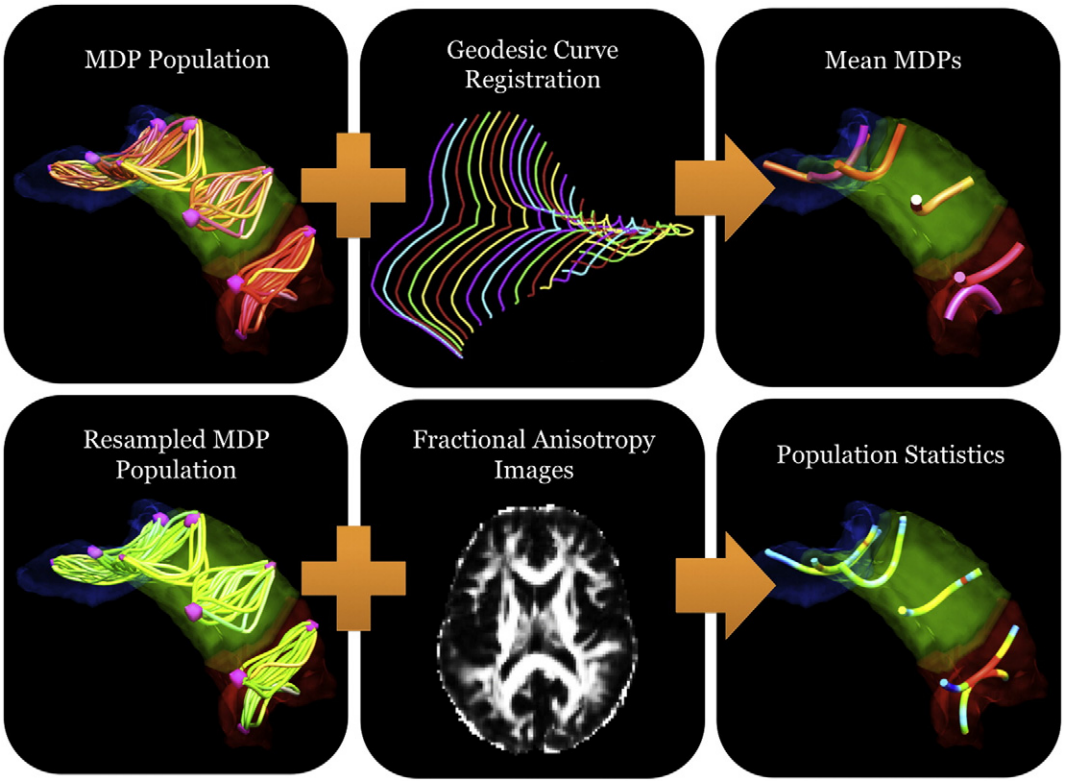

Fig. 2. Schematic of the pipeline for performing statistical analysis of populations of MDPs. The first panel represents MDPs for a population of subjects for a given white matter tract. A nonlinear iterative method that uses geodesic curve registrations is used to compute an average MDP representing the mean shape of the population. A correspondence is established for all subject MDPs via the average MDP. Fractional anisotropy (FA) from each subject is resampled for the corresponding points and compared across the population.

Our tractography method selects fibers in the diffusion image space by generating scores for all possible curves at a seed point. These curves are parameterized using 2nd order polynomials. An additional parameter controls the maximum expected curve length and is set to a value representing the largest dimension of the volume. In practice, the number of curves evaluated at each seed point is around one million based on resolution and computational resources.

First, seed points are generated using a prior probability based on FA from the single-tensor model of diffusion (Basser and Pierpaoli, 1996), defined as

$$\sqrt{\frac{1}{2}}\frac{\sqrt{(\lambda_1-\lambda_2)^2+(\lambda_2-\lambda_3)^2+(\lambda_3-\lambda_1)^2}}{\sqrt{\lambda_1^2+\lambda_2^2+\lambda_3^2}}, \tag{1}$$

where λ_1, λ_2, and λ_3 are the eigenvalues of the diffusion tensor. These seed points are used to generate curves that receive a score estimating the probability of their existence, computed from the voxels the curve passes through.

The ODFs at each voxel from our HARDI images were computed using a normalized and dimensionless estimator derived from Q-ball imaging (QBI) (Aganj et al., 2010). This method uses the Jacobian factor r^2 for the constant solid angle (CSA) ODF as

$$\frac{1}{4\pi}+\frac{1}{16\pi^2}\mathrm{FRT}\left\{\nabla_b^2 ln\left(-ln\frac{S(\hat{u})}{S_0}\right)\right\}. \tag{2}$$

In this equation, $S(\hat{u})$ is the diffusion signal, S_0 is the baseline image, FRT is the Funk–Radon transform, and ∇_b^2 is the Laplace–Beltrami operator. This estimate outperforms the original definition (Tuch, 2004) with superior resolution for detecting multiple fiber orientations (Aganj et al., 2010; Descoteaux and Bore, 2012; Fritzsche et al., 2010; Ghosh et al., 2013).

The Hough transform tractography chooses fibers from all possible curves generated in the image at a certain space and parameter resolution. These curves are parametrized by their arc length, s, ranging in value from L_- to L_+ and approximated using simple polynomials. The scores for each possible curve, s, derive from the ODF and FA values

$$\int_{L_-}^{L_+}\left(\log\left[\mathrm{ODF}_{\vec{x}(s)}(\hat{t}(s))\mathrm{FA}\left(\vec{x}(s)\right)\right]+\lambda\right)ds, \tag{3}$$

where $\mathrm{ODF}_{\vec{x}(s)}(\hat{t}(s))$ is the value from the ODF at the 3D location $\vec{x}(s)$ and direction specified by the unit tangent vector $\hat{t}(s)$. The method scores as many fibers as possible arising from a seed point and uses the voting process provided by the Hough transform to select the best fitting curve. These filtered curves comprise the final set of fibers produced by the method for a single subject, which we refer to as F.

The method is probabilistic in its selection of a fiber at a certain seed point but does not generate volumetric data giving a distribution of fibers in the white matter. It chooses seed points (voxel locations) randomly throughout the white matter tissue with a probability proportional to their fractional anisotropy. Once a seed point is chosen, the algorithm scores all possible fibers that pass through this point. The number of fibers is restricted by the parameterization and range of the variables involved, but is close to one million candidate fibers. For each fiber this score represents the probability of that fiber existing and is constructed by integrating the orientation distribution function over the span of the fiber combined with the probability of the corresponding seed point. The method then uses the Hough transform to select the fiber with the highest probability or highest score as the final fiber for that seed point.

The tractography algorithm was run on each subject image and generated around 10,000 fibers (Fig. 1 shows a representative example with our data), which represents a good balance between computational efficiency and sampling enough of the image space (Prasad et al., 2013c). We subsequently clustered these fibers using a white matter atlas.

Fiber clustering with white matter ROI atlas

Fibers extracted using the Hough transform-based tractography method are clustered using a ROI atlas to incorporate prior anatomical information. We use the Johns Hopkins University (JHU) atlas (Wakana et al., 2004), which delineates 50 white matter regions of interest (ROI). This ROI atlas is first registered to our subject space using an affine transform provided by FMRIB's Linear Image Registration Tool (FLIRT) (Jenkinson and Smith, 2001). This is then followed by a nonlinear transform from the Automatic Registration Toolbox (ART) (Ardekani et al., 2005; Klein et al., 2009) to refine the registration further.

We then cluster the fiber bundles by measuring the intersection of fibers with the ROI atlas as follows. A fiber intersection score is computed by counting the number of ROI voxels that intersect with the fiber tract. This score is used to select fibers that belong to an ROI and thus a white matter tract. Spuriously intersecting fibers are eliminated by applying an experimentally determined threshold that is dependent on the number and the type of fibers obtained from the tractography algorithm. Formally, if F is the set of fibers for a subject and r is a specific white matter ROI label in the atlas, then the subset of selected fibers in a bundle is given by,

$$B=\left\{f|f\in F,\int_f A(s,r)ds>t\right\}, \tag{4}$$

where t is the intersection threshold and A is an indicator function defined to be

$$A(s,r)=\begin{cases}1 & \text{if } s\in \text{region } r\\ 0 & \text{otherwise.}\end{cases} \tag{5}$$

Bundle representation using the maximum density path

The fiber bundle B representing a white matter tract is reduced to a compact representation also referred to as the maximum density path as follows. We first compute a density volume of our fiber bundle to characterize our search space, and denote it as

$$I_d\left(\vec{x}\right)=\sum_{b\in B}Q\left(b,\vec{x}\right), \tag{6}$$

where $\vec{x}$ represents a 3D voxel location and Q is the indicator function

$$Q\left(b,\vec{x}\right)=\begin{cases}1 & \text{if } b \text{ intersects } \vec{x}\\ 0 & \text{otherwise.}\end{cases} \tag{7}$$

This value specifies the number of fibers that intersects each voxel. We then construct a graph that represents the voxel-wise fiber density in our fiber bundle. The above voxels are used as nodes in a graph, $G=\{N,E\}$ (a set of nodes and undirected edges connecting them) with those of positive value connected to their 26 neighbors by undirected edges. In our formulation, the weight of an edge between nodes i and j is calculated as the sum of the voxel densities it connected as

$$I_d\left(\vec{x}(i)\right)+I_d\left(\vec{x}(j)\right), \tag{8}$$

with $I_d\left(\vec{x}(k)\right)$ as the density for the voxel location $\vec{x}(k)$ corresponding to node k. These edge weights are then modified by subtracting each from the maximum initial edge cost, e_m, such that edges in high density regions have weights close to zero. These edge weights are designed to allow the shortest path algorithm to go through edges in high density regions. We use Dijkstra's algorithm (Dijkstra, 1959) to compute the path through this graph following the nodes with highest density. Dijkstra's algorithm is a graph search method that finds the shortest

path from a source node to every other node. However, the number of nodes in the graph may be large and when the algorithm is used for a single destination node, it may be stopped once that path is found. To find the shortest path to represent a white matter region, we require the graph to have start and end nodes to constrain the path to a specific region of the graph. These nodes are specified by an expert in the ROI atlas. The ROI points for the start and end locations may not always correspond to the positive density values derived from our bundle. Thus, we find the closest voxels in the density volume as the corresponding start and end nodes in subsequent computation with Euclidean distance used as the metric.

In our implementation, we used Dijkstra's algorithm (Dijkstra, 1959) to find only the shortest path between the start and the end nodes selected in the graph. If Dijkstra's algorithm is unable to find the path between the start and end nodes our method automatically identifies this situation and takes steps to remedy the graph and finds the shortest path. The algorithm will be incapable of finding a connection between the two nodes if the structure of the graph is such that there are no edges from the subgraph containing the start node with the subgraph containing the end node. This can be caused by scanner artifacts or suboptimal solutions due to the fiber tractrography algorithm. In this situation, we add extra edges and nodes to the graph so all voxels within our ROI are fully connected with their neighbors. The edges are weighted by the largest edge cost in the current graph. This allows gaps between the start and end nodes to be filled in and use as few edges as possible in regions with unknown data. The nodes in the path correspond to a set of voxel locations in our image space. We smooth the path so it is better conditioned for subsequent processing. We convolve the 3D coordinates of our path with a Gaussian kernel to achieve this, though fitting these points to a spline would also have sufficed. A summary of these steps is presented in Algorithm 1. We represent the maximum density path

Algorithm 1 Maximum Density Path Method

```
Generate fibers from tractography
for r = 1 to M (number of atlas regions) do
    for f = 1 to N (number of fibers) do
        Find intersection of fiber f and region r
        If intersection measure ∫_f A(s,r)ds > t add to bundle B
    end for
    Convert fiber bundle, B, to density image
    Generate graph G{N, E} with density image voxels as nodes and edge weights as e_m − (I_d(i) + I_d(j))
    Find the closest (in Euclidean distance) start, a, and end, b, nodes from region r in the atlas
    Use Dijkstra's algorithm to find the shortest path, p, between nodes a and b
    Convolve p with a Gaussian kernel to generate the maximum density path (MDP), β
end for
```

by the coordinate function of the parameterized curve, and denote it to be β such that β:[0, 1] → $\mathbb{R}^3$. Fig. 3 shows an example of a maximum density path computed for a fiber bundle. Additionally, we show the density and the FA images that correspond to the fibers. For comparison, we also show a representative mean fiber by applying Procrustes analysis to align the fibers in the bundle to their mean. We then compute a new mean (shown in blue) of the fibers after alignment. This example shows that even if the bundle includes a few spurious fibers, it can drastically change the appearance of the mean fiber derived from Procrustes analysis, while the MDP remains stable.

Fig. 5 shows an example of MDPs for a population of subjects overlaid on each other. Some of these paths are short because the corresponding regions of interest in the white matter atlas are small. This means the seed points specified in the atlas are not very far apart and even if the fibers are much larger they are summarized by the structure within the white matter region and points with the highest density. An alternative could be to use a probabilistic white matter atlas and threshold the regions so they encompass a larger fraction of the fiber lengths in the white matter region.

Shape analysis of maximum density paths

This section outlines the method for shape representation and analysis for maximum density paths. The maximum density paths denoting tracts are modeled as continuous open curves in $\mathbb{R}^3$ but they can also be considered as points in an infinite-dimensional space of curves. This space is induced by a suitable Riemannian metric defined on its tangent space. Shape matching between MDPs is enabled by measuring shortest length paths, also known as geodesics connecting two shapes in the shape space. The geodesic not only measures the length of the path and quantifies the geometric distance between two shapes, but also represents an optimal geometric deformation that highlights the anatomical differences between the shapes. Additionally geodesics are an important ingredient for constructing intrinsic population averages for shapes — an essential step in statistical analysis of shapes.

Representation of MDPs

We represent the shape of the coordinate function of the MDP using a vector-valued function (Joshi et al., 2007a,b; Srivastava et al., 2011) as

$$q(s) = \frac{\dot{\beta}(s)}{\sqrt{\|\dot{\beta}(s)\|}} \in \mathbb{R}^3. \tag{9}$$

Here, $s \in [0, 1]$, $\|\cdot\| \equiv \sqrt{(\cdot,\cdot)_{\mathbb{R}^3}}$, and $(\cdot,\cdot)_{\mathbb{R}3}$ is the standard Euclidean inner product in $\mathbb{R}^3$. Our goal is to achieve an elastic shape matching between MDPs. We would also prefer that the shape matching is invariant with respect to the orientation and scale of the MDPs. Owing to the derivative, the function q is invariant to the translation of the MDP coordinate curve. To impose scale-invariance, we normalize the q functions by its magnitude. Thus we denote $\mathcal{Q} \equiv \left\{ q \middle| q(s) : [0,1] \to \mathbb{R}^3 \middle| \int_0^1 (q(s), q(s))_{\mathbb{R}^3} ds = 1 \right\}$ as the space of all unit-length curves. On account of scale invariance, the space $\mathcal{Q}$ becomes an infinite-dimensional unit-sphere of functions, and represents all open elastic curves invariant to translation and uniform scaling. The elasticity of the representation is due to the presence of the square-root in the denominator that allows the q function to have arbitrary speeds. To define a metric on the space $\mathcal{Q}$, we first define its tangent space which is the set of all tangent vectors orthogonal to q. Formally, the tangent space of $\mathcal{Q}$ is given by $Tq(\mathcal{Q}) = \{w = (w_1, w_2, \cdots, w_n) | w(s): [0, 1] \to \mathbb{R}^3, \forall_s \in [0, 1]$ such that $\int_0^1 (w(s), q(s))\mathbb{R}^3 \, ds = 0\}$, where $n = 3$. Here each w_i represents a tangent vector in the tangent space of $\mathcal{Q}$. Now, the metric on the tangent space $Tq(\mathcal{Q})$ is defined as follows. Given a curve $q \in \mathcal{Q}$, and the first order perturbations of q given by $u, v \in Tq(\mathcal{Q})$, respectively, the inner product between the tangent vectors u, v to $\mathcal{Q}$ at q is defined as,

$$\langle u, v \rangle = \int_0^1 (u(s), v(s))_{\mathbb{R}^3} ds. \tag{10}$$

Now given two shapes q_1 and q_2, the translation and scale-invariant shape distance between them is simply found by measuring the length of the geodesic, or the great circle connecting

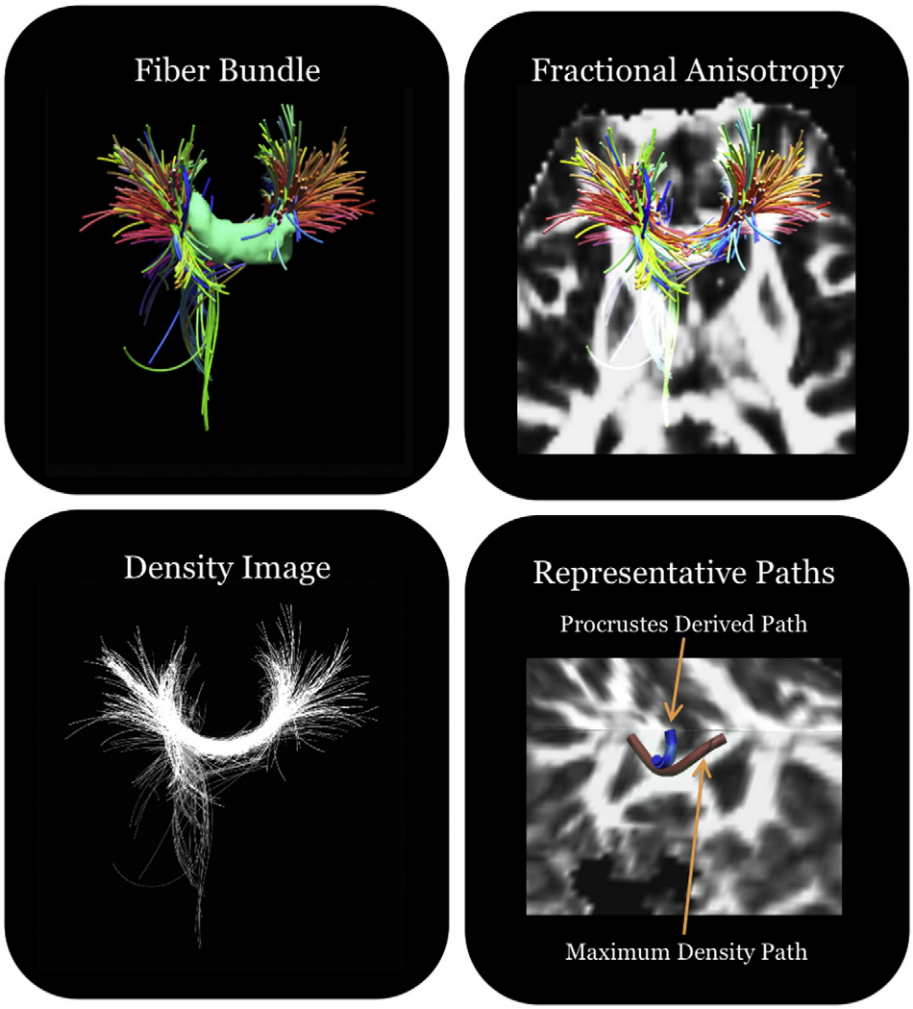

Fig. 3. This figure shows the advantage of using the maximum density path method to represent fiber bundles. The first panel shows the tractography fibers that intersect a white matter tract. The selected fibers all intersect the region of interest to the same degree, but the selection process includes spurious fibers. In the context of fractional anisotropy, we can see that the spurious fibers may be part of another adjacent tract. Our method represents the fiber bundle as a density image and searches for a representative path based on seed points from a registered atlas. The final slide shows the resulting maximum density path compared to a path found by taking the mean of the fibers and using Procrustes analysis to align and recompute a representative mean. If only the fiber shape is used, the resulting representative path from Procrustes analysis may not effectively traverse the region of interest due to spuriously included fibers. By leveraging the distribution of fibers the MDPs seek to build a representation of the white matter that is reflective of the underlying feature geometry of the most important regions and help matching across subjects for subsequent population analysis.

them on the sphere. Thus given a tangent vector $f \in Tq_1(\mathcal{Q})$ in the direction of q_2 given by $f = q_2 - < q_1, q_2 > q_1$, the geodesic (Joshi et al., 2007a) on $\mathcal{Q}$ between the two points $q_1, q_2 \in \mathcal{Q}$ along f, for an infinitesimal time t is given by

$$\chi_t(q_1; f) = cos\left(tcos^{-1}\langle q_1, q_2\rangle\right)q_1 + sin\left(tcos^{-1}\langle q_1, q_2\rangle\right)f. \tag{11}$$

Then the geodesic distance between the two shapes q_1 and q_2 is given by

$$d(q_1, q_2) = \int_0^1 \sqrt{\langle \dot{\chi}_t, \dot{\chi}_t \rangle} dt. \tag{12}$$

The geodesic distance (Joshi et al., 2007a) given in Eq. (12) is only invariant to translation and scale. To make it invariant to rotations, we consider the shortest distance

$$d_r(q_1, q_2) = \underset{O_3 \in SO(3)}{\arg\min}\, d(q_1, O_3 q_2). \tag{13}$$

Eq. (13) can be minimized either using gradient descent over the tangent space of $SO(3)$ or by using singular value decomposition (Rohlf and Slice, 1990). In this paper, we find the rotation invariant distance as

$$d_r(q_1, q_2) = d\left(q_1, \hat{O}_3 q_2\right), \tag{14}$$

where $\hat{O}_3 = ADB^T = \int_0^{2\pi} q_1(s) q_2(s)^T ds$, A and B are left and right unitary matrices, and D is a matrix given by $D = \begin{bmatrix} 1 & 0 \\ 0 & |A||B| \end{bmatrix}$. Finally, since we are representing MDPs by parameterized curves, we would like the shape matching to be invariant to reparameterizations. Following (Joshi et al., 2007a), we denote the reparameterization of a MDP curve using a group action by a diffeomorphism γ, given by $q \cdot \gamma = \sqrt{\dot{\gamma}} q(\gamma)$. Then the optimal reparameterization $\hat{\gamma}$ is approximated by the minimizer

$$\hat{\gamma} = arg\ \underset{\gamma}{min}\left(\int_0^{2\pi}\left[\left\|q_1 - \widetilde{q}_2 \cdot \gamma\right\|^2 + \left\|\widetilde{q}_2 - q_1 \cdot \gamma^{-1}\right\|^2\right] ds\right), \tag{15}$$

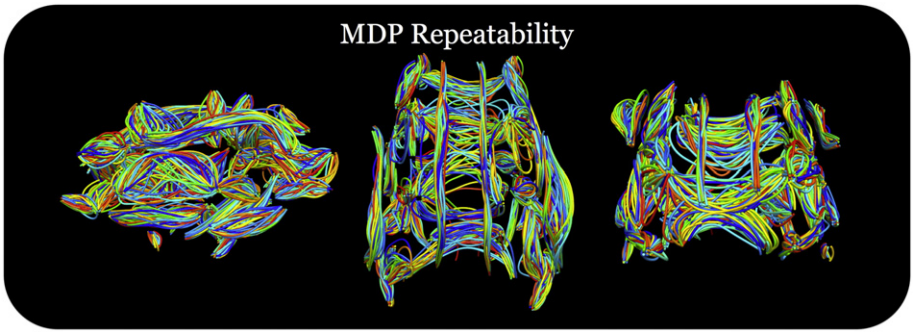

Fig. 4. To study the variability of the maximum density paths (MDPs), we used 23 subjects with repeat scans. We used the MDP algorithm to represent 67 white matter regions from an ROI atlas and find correspondences between the two acquisitions for a single subject. We used paired-sample *t*-tests to compare the fractional anisotropy (FA) values across corresponding points. We found that there were no significant differences in the MDPs when correcting for multiple comparisons using the false discovery rate (FDR) at the 0.05 level. This helps provide support that the statistical analysis tools using MDPs will be able to investigate patterns of white matter structure and may be less affected by noisy or highly variable representation inherent to the algorithm.

where $\tilde{q}_2 = \hat{O}_3 q_2$. In this paper, we use dynamic programming to obtain the solution to Eq. (15). The fully elastic, pose, scale, and reparameterization invariant distance between MDPs is given by

$$d_e(q_1, q_2) = d_r(q_1, q_2 \cdot \hat{\gamma}) = d\left(q_1, \left(\hat{O}_3 q_2\right) \cdot \hat{\gamma}\right). \tag{16}$$

The optimal geodesic path can also be denoted by a one-parameter flow Ψ and the tangent vector $\dot{\alpha}_t$, such that

$$\Psi_t : \Psi_0(q_1, \dot{\alpha}_t) = q_1, \Psi_1(q_1, \dot{\alpha}_t) = \left(\hat{O}_3 q_2\right) \cdot \hat{\gamma}. \tag{17}$$

The optimal tangent vector can then be written from Eq. (17) as

$$\hat{\dot{\alpha}}_t = \Psi_1^{-1}\left(q_1, \left(\hat{O}_3 q_2\right) \cdot \hat{\gamma}\right). \tag{18}$$

Statistical analysis of MDPs across a population

To evaluate group-level effects of sex, age, disease or even genetic influences on the MDP representations of white matter tracts, we need a suitable mechanism for performing statistical analysis. As MDPs are represented by parameterized functions defined on a shape space, one natural approach is to use the inherent non-linearity of the shape space, and define appropriate statistical measures under the Riemannian metric in Eq. (10). This approach is also called an intrinsic statistical analysis and leads to the definition of the Karcher mean (also known as the Fréchet mean) (Joshi et al., 2013; Le, 1995; Srivastava et al., 2005) in the shape space of all MDPs. Given a collection of MDP shapes $\{q_i\}$, $i = 1, \cdots, N$, the Karcher mean is defined as

$$\mu = \arg\min_{q_\mu} \sum_{i=1}^{N} d_e\left(q_\mu, q_i\right)^2. \tag{19}$$

The computation of the Karcher mean involves computing geodesics at each step iteratively and proceeds as follows. For the first iteration, an extrinsic mean (Euclidean average) is computed and projected on the shape space. This is assumed to be the current estimate of the Karcher mean. For the subsequent iterations, geodesics are computed between all the individual shapes in the population to the current estimate. The tangent vectors $\left(\hat{\dot{\alpha}}_t^i, i = 1, \ldots, N\right)$ are then computed as a result of minimizing Eq. (16) and averaged together. A geodesic flow is then constructed using Eq. (17) to yield a new estimate of the mean shape. This procedure is repeated until the geodesic variance given by the sum of the squared geodesic distances to the mean shape is minimized, and the mean converges. The Karcher mean completely relies on the geometry of the shape space and is useful in computing intrinsic statistical estimates such as covariances of MDPs. Additionally, the geodesics produce correspondences, making it easier to compare white matter measures projected on the MDPs across a population. This is also useful

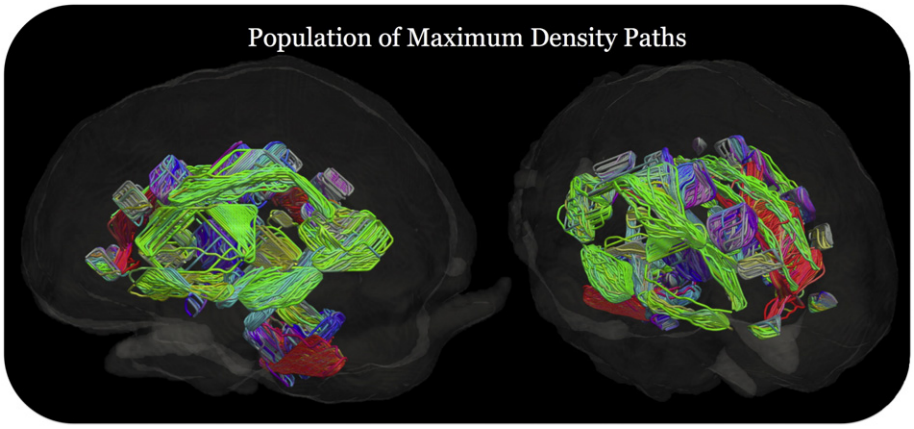

Fig. 5. We show here the maximum density paths (MDPs) computed for 238 subjects in 67 regions from the white matter atlas before they are matched together using curve registration. This sample was used for our heritability analysis. The color represents the direction or orientation of the middle segment connecting two points at the middle of each MDP. If the orientation is anterior to posterior it is colored green, if it is from left to right it is red, and if it is superior to inferior it is colored blue. In each of the 67 areas, we register the 238 MDPs to a mean MDP to find correspondences using geodesic curve registration. We then sample FA along corresponding points in each subject for our subsequent statistical and genetic analysis. A few of these paths are relatively short because the corresponding region of interest in the white matter atlas was small and thus the MDP will cover a shorter range even though the tractography fibers it represent may be much larger.

for studying differences in disease, sex, aging, or even heritability in a population.

Results

Experiments

We show experimental results on a dataset of $N = 565$ young adults, including healthy twins and their siblings. The participants were scanned with a 4-Tesla Bruker Medspec MRI scanner, collecting 3D 105-gradient high angular resolution diffusion images (HARDI) and standard structural T1-weighted magnetic resolution images (MRI). The images consisted of 55 slices, 2-mm thick, with a 1.79×1.79 mm^2 in-plane resolution. For each person, we collected 94 diffusion-weighted images ($b = 1159$ s/mm^2) using a uniform distribution of gradient directions on the hemisphere. We also collected 11 b_0 (non-diffusion encoding) images and corrected all images for eddy current distortions and motion with FSL (www.fmrib.ox.ax.uk/fsl). Our cohort consisted of 367 women and 198 men, ranging from 20 to 29 years of age. Study participants gave informed consent; institutional ethics committees at the Queensland Institute of Medical Research, the University of Queensland, the Wesley Hospital, and at UCLA approved the study.

For each T1-weighted image, we manually removed non-brain tissue and registered it to the Colin27 (Holmes et al., 1998) high resolution brain template using a 9-parameter transformation. These skull-stripped and registered T1-weighted images each has a corresponding average b_0 diffusion weighted image (DWI), combining all the eleven images. These average images were masked using BET (Smith, 2002) and this space was used to generate FA maps. Additionally, we used the FA images to compute a geometrically-centered study-specific mean template (mean deformation template; MDT). We registered the JHU ROI atlas to the MDT for the MDP analysis. Our ROI atlas contained 50 different white matter regions, which were seeded to extract 67 different MDPs. Our results based on all 565 subjects are shown in Fig. 6.

Repeatability of MDPs

We examined the reliability of the MDP construction procedure by analyzing subjects with repeat scans. Twenty-three subjects in the total population used in our analysis had repeat scans, which were used to test the stability of MDP construction across the two acquisitions. We used the MDP algorithm to find corresponding points along each MDP and used paired-sample t-tests to study if the FA values in these white matter tract representations were significantly different. Fig. 4 shows the collection of MDPs with 46 curves in each white matter region from the ROI atlas. Each of the 23 pairs is colored randomly with the two MDPs in a single pair having matching coloring. We corrected for multiple comparisons using the false discover rate (FDR) (Benjamini and Hochberg, 1995) at the 0.05 level and none of the values were significantly different between scans. This provides an indication of the stability of MDP representation and may help support a more meaningful interpretation of the subsequent statistical analyses.

Genetic effects on white matter morphology using MDPs

The twin cohort in the data is made up of monozygotic (MZ) and dizygotic (DZ) pairs, allowing us to estimate the relative contributions of additive genetic factors (A), shared environment (C), and unshared or unique environment (E) to the measures derived from the scans — in our case, FA along the MDPs. This standard "A/C/E" model describes the FA at each point on the MDP as a combination of latent variables, $\mathrm{FA} = aA + cC + eE$. In this formulation the total variance is $\mathrm{var(FA)} = a^2 + c^2 + e^2$ with $\mathrm{var}(A) = a^2$, $\mathrm{var}(C) = c^2$, and $\mathrm{var}(E) = e^2$. We are able to estimate the unobserved factor loadings as there is a difference in the theoretical covariance of FA for a MZ twin pair, $a^2 + c^2$, and for a DZ twin pair, $(1/2)a^2 + c^2$, which we solved using a maximum likelihood fitting (Neale and Cardon, 1992) that estimated the parameters of the model. These methods are detailed in (Chiang et al., 2009).

Several studies (Chiang et al., 2011; Jahanshad et al., 2013b; Patel et al., 2010; Thompson et al., 2001) have shown evidence for heritability of the white matter structure in the brain. Here, we use heritability as a metric to understand how well MDPs were able to model and capture the underlying morphology of the white matter structure in our data. If the representation is able to effectively pick up heritability effects then our hypothesis is that the MDP matching across subjects reflects the underlying anatomical homology, and that the MDP model is better able to describe white matter brain structure.

We fitted the "A/C/E" model to the FA values on the skeleton that fell within the ROI atlas. In our experiments we compared the full "A/C/E" model to the simpler formulation with two variables using minus two times the log-likelihood ratio, which approximately follows a χ^2 distribution, meaning that $P > 0.05$ indicates a good fit. We found that the shared environment term (C component) did not have a significant fit for either method, so we used a simplified "A/E" model instead. This model selection procedure and selecting the "A/E" model instead of the more complicated "A/C/E" is widely used (Geschwind et al., 2002) and common with real data (Baaré et al., 2001). In the "A/E" model, a^2

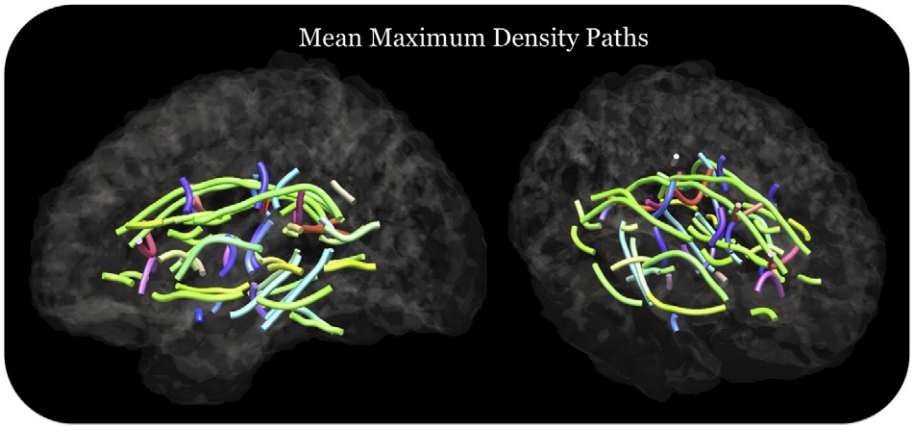

Fig. 6. Average maximum density paths (MDPs) for 67 fiber bundles in 565 twin images. The color represents the direction or orientation of the middle segment connecting two points at the middle of each MDP. If the orientation is anterior to posterior it is colored green, if it is from left to right it is red, and if it is superior to inferior it is colored blue. The paths are derived from tractography fibers, clustered into white matter tracts, and then represented as paths that follow points in these white matter bundles of maximum density of fibers. The 67 different paths come from regions in our white matter atlas that have been annotated with seed points, which become the endpoints of the paths. The mean MDPs provide a template for curve registration to find correspondences between the individual subjects and allow us to compactly represent population statistics across the entire white matter.

represents the proportion of the variance due to additive genetic factors and the other parameter in the model, e^2, represents the proportion of variance that is due to environmental factors including measurement error. Thus we model the components of FA variance as simply $\text{var(FA)} = a^2 + e^2$ and since we are interested in the relative proportion of variance captured by each component we can normalize this equation by var(FA) to interpret their relationship as $1 = \tilde{a}^2 + \tilde{e}^2$ or $1 - \tilde{a}^2 = \tilde{e}^2$. The goal here is to use the model that best fits the data to understand the genetic and environmental contributions to the variance. Maximizing heritable estimates in this case may imply minimizing measurement error and therefore may represent a stronger metric for measuring white matter microstructure. In addition, these highly heritable regions present good candidates for genetic associations and could be useful for cutting down on the dimensionality of the image for these types of analyses.

We also compared our method of analyzing white matter bundles using MDPs and analyzing the white matter skeletons from TBSS (Fig. 7) in our subjects. We used the genetic contribution due to FA using both methods to compare heritability. We found that the density and FA images smoothed with an isotropic Gaussian filter with full-width at half maximum (FWHM) = 3 mm produced higher a^2 values (Chiang et al., 2012). We restricted the analysis to 238 (48 monozygotic and 71 dizygotic) of the 565 twins because of issues with the nonlinear registration from TBSS in the omitted subjects.

We computed genetic associations with mixed-model regression (Aulchenko et al., 2007; Jahanshad et al., 2012a) along the MDPs using the genes NTRK1, CLU, and COMT. We found NTRK1 passed a local false discovery rate (FDR) threshold (for a single MDP) in 20 regions across our white matter atlas represented as MDPs. In addition, we found CLU passed local FDR at the anterior limb of internal capsule right, posterior limb of internal capsule right, and anterior corona radiata left, and COMT passed at the sagittal stratum right (including inferior longitudinal fasciculus and inferior fronto-occipital fasciculus). When we used a global FDR, by combining the 67 MDPs into one large MDP, NTRK1 was the only SNP that survived in 600 of the total 1897 points in the global MDP. The results from NTRK1 and CLU agree with earlier studies of this dataset using voxel-based maps (Braskie et al., 2011, 2012).

Discussion and conclusion

We have presented a method for extracting, representing, and analyzing the geometry of white matter bundles using maximum density paths. Our method enables population analysis of diffusion-weighted images without relying exclusively on global registration of the images into the same space. Image registration is performed only once to transform the ROI white matter atlas to the subject space for the purpose of initializing the seed points for clustering fibers from tractography. Density image volumes are computed from the fiber bundles, and MDPs are constructed using Dijkstra's algorithm by imposing a graph structure on the images. The shapes of MDPs are then brought into correspondence through geodesic curve registration, allowing us to focus specifically on the white matter region we are interested in without involving the rest of the image. Further, our method introduces a way to perform localized statistical analysis of white matter tracts. The MDPs, use the start and end points from major

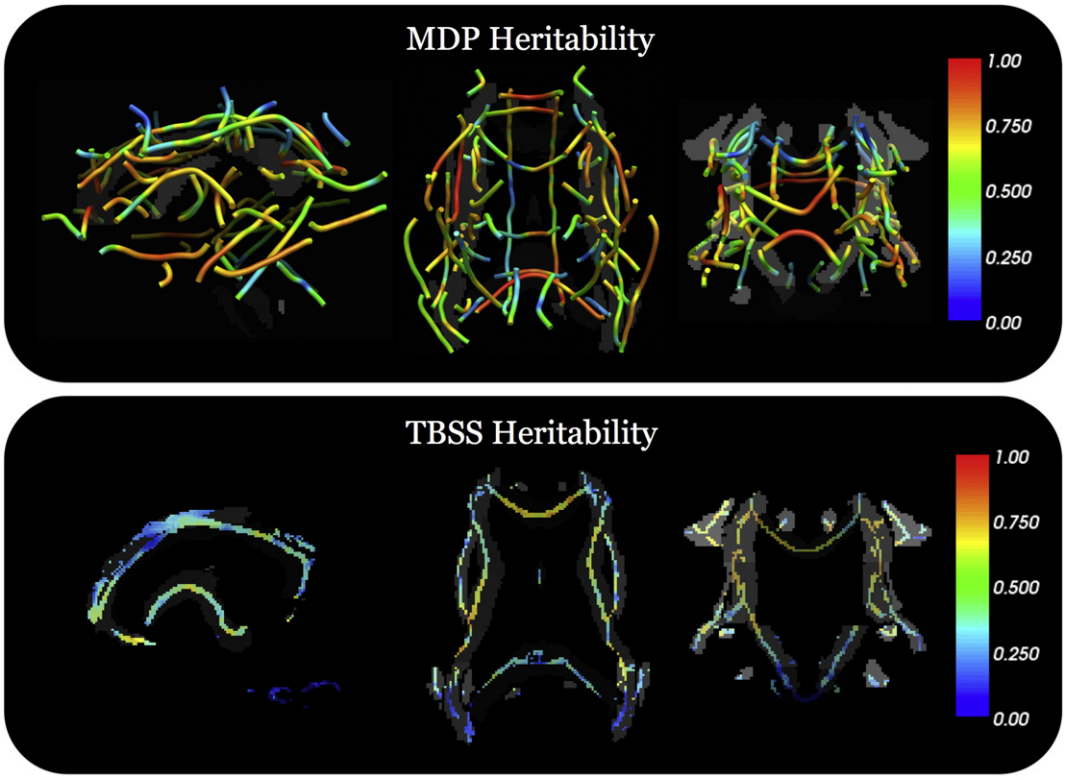

Fig. 7. Our twin data contained monozygotic (MZ) twins that share 100% of their genetic material and dizygotic twins (DZ) that share 50% of their genetic material. This structure in our data allows us to use structural equation models, particularly the A/E model, to estimate the amount of variance in a phenotype (in our case the white matter structure) that is due to genetic effects (heritability), or to unique environment factors and measurement error. In this figure we show the genetic contribution to white matter structure using our maximum density (MDP) path representation method compared with the white matter skeleton from tract-based spatial statistics (TBSS). A high value (red) means a large fraction of the variance in white matter structure is determined by genetic effects while a low value (blue) means the variance in structure was accounted for more by the environment. Since this is the proportion of variance accounted for by heritability, an analogous figure of the environmental contributions would involve simply reversing the color bar. Previous studies have shown high heritability of white matter tissue and we used the fraction of genetic determination as the metric to evaluate how well our MDP representation summarized the white matter structure in our data. The MDP method may have a better ability to pick up on the heritability because of the curve registration that is computed for each white matter region individually, which improves coherence of homologous points across subjects. We used a subset of 238 subjects for this analysis (48 MZ pairs, 71 DZ pairs). The top panel shows the MDPs from 67 white matter regions with slices of the atlas overlaid. The TBSS results show orthogonal slices of the TBSS skeleton overlaid on the white matter atlas.

white matter pathways, and provide a compact representation so that correspondences can be easily computed. The correspondences can be directly visualized on the MDPs to reveal which part of a white matter tract depends on an external parameter. These tools provide the foundation for any study of white matter tracts or any type of population analysis using diffusion weighted images. The complete procedure is available as an end-to-end computational pipeline for white matter tract-based analysis of diffusion images.

In the examples presented here, we used a global Hough transform method for tractography, but the MDP representation is general enough to be used with any type of tractography method. Our method relies on density images from tractography, which could be computed using streamlines (Basser et al., 2000), a deflection based algorithm (Lazar et al., 2003), or any of the recent deterministic and probabilistic methods (Zhan et al., 2013). In our case, we chose a tractography method that could benefit from the information rich HARDI data, but depending on the resources and data available, researchers may prefer to use fibers from diffusion tensor imaging or diffusion spectrum imaging (Wedeen et al., 2005) based algorithms. The graph-based representation for the fiber density volume enabled us to conveniently incorporate the density information in the structure, and further led to an efficient solution provided by Dijkstra's algorithm. However we could have formulated the problem using snakes (Kass et al., 1988) or splines (Park and Lee, 2007) as well. Alternatively, other density representations such as those using surfaces (Zijdenbos et al., 1994) or using the volumetric segmentations (Kubicki et al., 2005) directly would have introduced a host of issues with registration and subsequent analysis of correspondences. We chose to use an ROI atlas to select fibers for analysis and representation with the MDP though alternative approaches may work without relying on the registration of the atlas into the image space. Future work could incorporate automatic clustering of tractography fibers using approaches such as a hierarchical Dirichlet processes mixture model (Wang et al., 2011), a voxel based approach (Guevara et al., 2011), a spectral approach (Guevara et al., 2011), or even shape clustering (Joshi and Srivastava, 2003; Joshi et al., 2004). Hierarchical approaches may enable a user to specify the resolution of MDPs in the brain tissue. As an alternative to FA, any other type of statistics on the density paths could be used instead. We could compute mean diffusivity (Le Bihan et al., 2001), generalized FA (Barmpoutis et al., 2009), or the tensor distribution function and interpolate them along each MDP. Our white matter analysis framework could even be scored by their capacity (Prasad et al., 2013b) and used as measures of connectivity to complement (Prasad et al., 2013a) and optimize our representation of brain connectivity networks (Prasad et al., 2014).

Preliminary studies have used MDPs to study sex differences (Prasad et al., 2011b), Alzheimer's disease (Nir et al., 2012, 2014), 22q11.2 deletion syndrome (Villalon-Reina et al., 2012), and depression (Sacchet et al., 2014). Our results in the current study showed promise in our new representation and agreed with voxelwise analyses of the entire white matter tissue in diffusion images (Chiang et al., 2011; Jahanshad et al., 2013b; Patel et al., 2010; Thompson et al., 2001) that showed a pattern of highly heritable regions. In this work, we evaluated MDPs by their ability to detect the effects of heritability in a cohort of monozygotic and dizygotic twins. Heritability analysis of FA in 50 regions of interest delineated in an ROI atlas, suggested promise of the method for detecting other factors that affect tracts, such as disease and risk genes. When comparing the genetic contributions (Fig. 7) to brain structure detected by our MDP method versus the TBSS method, we showed that MDPs can represent and display the structure using only one-tenth of the points in a TBSS skeleton. This reduction of the structural image data, which contains millions of voxels, may prove useful for genome-wise association studies (Cichon et al., 2009; Stein et al., 2010), as an alternative to voxel-based morphometry, or instead of group comparisons of statistics from segmentations. These data reduction steps may reduce the computational expense of a genome-wide search, and may also increase statistical power. In summary, we find that using tractography and creating MDPs give a similar skeletonized, yet more neuroanatomically accurate estimate of white matter microstructure than does TBSS as we found through improved heritability measures in the same sample.

Acknowledgments

This study was supported by NIH R01 grants MH097268, NS080655, AG040060, and EB008432. Additional support was provided by the National Health and Medical Research Council (NHMRC 486682, 1009064), Australia. Genotyping was supported by NHMRC (389875).

Author disclosure statement

The authors have no competing financial interests.

References

Abe, O., Aoki, S., Hayashi, N., Yamada, H., Kunimatsu, A., Mori, H., Yoshikawa, T., Okubo, T., Ohtomo, K., 2002. Normal aging in the central nervous system: quantitative MR diffusion-tensor analysis. Neurobiol. Aging 23, 433–441.

Aganj, I., Lenglet, C., Sapiro, G., Yacoub, E., Ugurbil, K., Harel, N., 2010. Reconstruction of the orientation distribution function in single- and multiple-shell Q-ball imaging within constant solid angle. Magn. Reson. Med. 64, 554–566.

Aganj, I., Lenglet, C., Jahanshad, N., Yacoub, E., Harel, N., Thompson, P.M., Sapiro, G., 2011. A Hough transform global probabilistic approach to multiple-subject diffusion MRI tractography. Med. Image Anal. 15, 414–425.

Ardekani, B., Guckemus, S., Bachman, A., Hoptman, M., Wojtaszek, M., Nierenberg, J., et al., 2005. Quantitative comparison of algorithms for inter-subject registration of 3D volumetric brain MRI scans. J. Neurosci. Methods 142, 67–76.

Aulchenko, Y., De Koning, D., Haley, C., 2007. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. Genetics 177, 577–585.

Baaré, W., Pol, H., Boomsma, D., Posthuma, D., de Geus, E., Schnack, H., van Haren, N., van Oel, C., Kahn, R., 2001. Quantitative genetic modeling of variation in human brain morphology. Cereb. Cortex 11, 816–824.

Barmpoutis, A., Hwang, M., Howland, D., Forder, J., Vemuri, B., 2009. Regularized positive-definite fourth order tensor field estimation from DW-MRI. NeuroImage 45, S153–S162.

Basser, P., Pierpaoli, C., 1996. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. J. Magn. Reson. Ser. B 111, 209–219.

Basser, P., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A., 2000. In vivo fiber tractography using DT-MRI data. Magn. Reson. Med. 44, 625–632.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 57, 289–300.

Braskie, M., Jahanshad, N., Stein, J., Barysheva, M., McMahon, K., de Zubicaray, G., Martin, N., Wright, M., Ringman, J., Toga, A., et al., 2011. Common Alzheimer's disease risk variant within the clu gene affects white matter microstructure in young adults. J. Neurosci. 31, 6764–6770.

Braskie, M., Jahanshad, N., Stein, J., Barysheva, M., Johnson, K., McMahon, K., de Zubicaray, G., Martin, N., Wright, M., Ringman, J., et al., 2012. Relationship of a variant in the NTRK1 gene to white matter microstructure in young adults. J. Neurosci. 32, 5964–5972.

Brun, A., Knutsson, H., Park, H., Shenton, M., Westin, C., 2004. Clustering fiber traces using normalized cuts. Med. Image Comput. Comput. Assist. Interv. 7, 368–375.

Chiang, M., Leow, A., Dutton, R., Barysheva, M., Rose, S., McMahon, K., de Zubicaray, G., Toga, A., Thompson, P., 2008. Fluid registration of diffusion tensor images using information theory. IEEE Trans. Med. Imaging 27, 442–456.

Chiang, M., Barysheva, M., Shattuck, D., Lee, A., Madsen, S., Avedissian, C., Klunder, A., Toga, A., McMahon, K., De Zubicaray, G., et al., 2009. Genetics of brain fiber architecture and intellectual performance. J. Neurosci. 29, 2212–2224.

Chiang, M., McMahon, K., de Zubicaray, G., Martin, N., Hickie, I., Toga, A., Wright, M., Thompson, P., 2011. Genetics of white matter development: a DTI study of 705 twins and their siblings aged 12 to 29. NeuroImage 54, 2308–2317.

Chiang, M., Barysheva, M., McMahon, K., de Zubicaray, G., Johnson, K., Montgomery, G., Martin, N., Toga, A., Wright, M., Shapshak, P., et al., 2012. Gene network effects on brain microstructure and intellectual performance identified in 472 twins. J. Neurosci. 32, 8732–8745.

Cichon, S., Craddock, N., Daly, M., Faraone, S., Gejman, P., Kelsoe, J., Lehner, T., Levinson, D., Moran, A., Sklar, P., et al., 2009. Genomewide association studies: history, rationale, and prospects for psychiatric disorders. Am. J. Psychiatry 166, 540–556.

Colby, J., Soderberg, L., Lebel, C., Dinov, I., Thompson, P., Sowell, E., 2011. Along-tract statistics allow for enhanced tractography analysis. NeuroImage 59 (4), 3227–3242.

Corouge, I., Fletcher, P., Joshi, S., Gouttard, S., Gerig, G., 2006. Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. Med. Image Anal. 10, 786–798.

Daianu, M., Jahanshad, N., Nir, T., Toga, A., Jack, C., Weiner, M., Thompson, P., 2013. Breakdown of brain connectivity between normal aging and Alzheimer's disease: a structural k-core network analysis. Brain Connect. 3.

Dennis, E., Jahanshad, N., McMahon, K., de Zubicaray, G., Martin, N., Hickie, I., Toga, A., Wright, M., Thompson, P., 2012. Development of brain structural connectivity

between ages 12 and 30: A 4-Tesla diffusion imaging study in 439 adolescents and adults. NeuroImage 64, 671–684.
Descoteaux, M., Bore, A., 2012. Testing classical single-shell HARDI techniques. Biomedical Imaging: From Nano to Macro, 2012 IEEE International Symposium on (HARDI Reconstruction Workshop). IEEE, p. 5.
Dijkstra, E., 1959. A note on two problems in connexion with graphs. Numer. Math. 1, 269–271.
Fritzsche, K.H., Laun, F.B., Meinzer, H.P., Stieltjes, B., 2010. Opportunities and pitfalls in the quantification of fiber integrity: what can we gain from Q-ball imaging? NeuroImage 51, 242–251.
Geschwind, D., Miller, B., DeCarli, C., Carmelli, D., 2002. Heritability of lobar brain volumes in twins supports genetic models of cerebral laterality and handedness. Proc. Natl. Acad. Sci. 99, 3176–3181.
Ghosh, A., Tsigaridas, E., Mourrain, B., Deriche, R., 2013. A polynomial approach for extracting the extrema of a spherical function and its application in diffusion MRI. Med. Image Anal. 17, 503–514.
Goodall, C., 1991. Procrustes methods in the statistical analysis of shape. J. R. Stat. Soc. Ser. B Methodol. 53, 285–339.
Guevara, P., Poupon, C., Rivière, D., Cointepas, Y., Descoteaux, M., Thirion, B., Mangin, J., 2011. Robust clustering of massive tractography datasets. NeuroImage 54, 1975–1993.
Holmes, C., Hoge, R., Collins, L., Woods, R., Toga, A., Evans, A., 1998. Enhancement of MR images using registration for signal averaging. J. Comput. Assist. Tomogr. 22, 324–333.
Jahanshad, N., Lee, A., Barysheva, M., McMahon, K., de Zubicaray, G., Martin, N., Wright, M., Toga, A., Thompson, P., 2010. Genetic influences on brain asymmetry: a DTI study of 374 twins and siblings. NeuroImage 52, 455–469.
Jahanshad, N., Kohannim, O., Hibar, D., Stein, J., McMahon, K., de Zubicaray, G., Medland, S., Montgomery, G., Whitfield, J., Martin, N., et al., 2012a. Brain structure in healthy adults is related to serum transferrin and the H63D polymorphism in the HFE gene. Proc. Natl. Acad. Sci. 109, E851–E859.
Jahanshad, N., Valcour, V., Nir, T., Kohannim, O., Busovaca, E., Nicolas, K., Thompson, P., 2012b. Disrupted brain networks in the aging HIV + population. Brain Connect. 2, 335–344.
Jahanshad, N., Rajagopalan, P., Hua, X., Hibar, D., Nir, T., Toga, A., Jack, C., Saykin, A., Green, R., Weiner, M., et al., 2013a. Genome-wide scan of healthy human connectome discovers SPON1 gene variant influencing dementia severity. Proc. Natl. Acad. Sci. 110, 4768–4773.
Jahanshad, N., Kochunov, P., Sprooten, E., Mandl, R., Nichols, T., Almassy, L., Blangero, J., Brouwer, R., Curran, J., de Zubicaray, G., et al., 2013b. Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: a pilot project of the ENIGMA–DTI working group. NeuroImage 81, 455–469.
Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5, 143–156.
Jin, Y., Shi, Y., Jahanshad, N., Aganj, I., Sapiro, G., Toga, A., Thompson, P., 2011a. 3D elastic registration improves HARDI-derived fiber alignment and automated tract clustering. Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on. IEEE, pp. 822–826.
Jin, Y., Shi, Y., Joshi, S.H., Jahanshad, N., Zhan, L., De Zubicaray, G.I., McMahon, K.L., Martin, N.G., Wright, M.J., Toga, A.W., et al., 2011b. Heritability of white matter fiber tract shapes: a hardi study of 198 twins. Multimodal Brain Image Analysis. Springer, pp. 35–43.
Jin, Y., Shi, Y., Zhan, L., de Zubicaray, G., McMahon, K., Martin, N., Wright, M., Thompson, P., 2013. Automatic HARDI white matter labeling by fusion of multiple tract atlases and its application to genetics. Biomedical Imaging: From Nano to Macro, 2013 IEEE International Symposium on. IEEE, pp. 512–515.
Jones, D., Symms, M., Cercignani, M., Howard, R., et al., 2005. The effect of filter size on VBM analyses of DT-MRI data. NeuroImage 26, 546–554.
Joshi, S., Srivastava, A., 2003. A geometric approach to shape clustering and learning. Statistical Signal Processing, 2003 IEEE Workshop on. IEEE, pp. 302–305.
Joshi, S., Srivastava, A., Mio, W., Liu, X., 2004. Hierarchical organization of shapes for efficient retrieval. Computer Vision-ECCV 2004. Springer, pp. 570–581.
Joshi, S., Klassen, E., Srivastava, A., Jermyn, I., 2007a. A novel representation for Riemannian analysis of elastic curves in Rn. 2007 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–7.
Joshi, S., Klassen, E., Srivastava, A., Jermyn, I., 2007b. Removing shape-preserving transformations in square-root elastic (SRE) framework for shape analysis of curves. Energy Minimization Methods in Computer Vision and Pattern Recognition, pp. 387–398.
Joshi, S.H., Narr, K.L., Philips, O.R., Nuechterlein, K.H., Asarnow, R.F., Toga, A.W., Woods, R.P., 2013. Statistical shape analysis of the corpus callosum in schizophrenia. NeuroImage 64, 547–559.
Kass, M., Witkin, A., Terzopoulos, D., 1988. Snakes: active contour models. Int. J. Comput. Vis. 1, 321–331.
Klein, A., Andersson, J., Ardekani, B., Ashburner, J., Avants, B., Chiang, M., Christensen, G., Collins, D., Gee, J., Hellier, P., et al., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage 46, 786.
Kochunov, P., Glahn, D., Lancaster, J., Winkler, A., Smith, S., Thompson, P., Almasy, L., Duggirala, R., Fox, P., Blangero, J., 2010. Genetics of microstructure of cerebral white matter using diffusion tensor imaging. NeuroImage 53, 1109–1116.
Kubicki, M., Park, H., Westin, C., Nestor, P., Mulkern, R., Maier, S., Niznikiewicz, M., Connor, E., Levitt, J., Frumin, M., et al., 2005. DTI and MTR abnormalities in schizophrenia: analysis of white matter integrity. NeuroImage 26, 1109–1118.
Lazar, M., Weinstein, D., Tsuruda, J., Hasan, K., Arfanakis, K., Meyerand, M., Badie, B., Rowley, H., Haughton, V., Field, A., et al., 2003. White matter tractography using diffusion tensor deflection. Hum. Brain Mapp. 18, 306–321.
Le, H., 1995. Mean size-and-shapes and mean shapes: a geometric point of view. Adv. Appl. Probab. 27, 44–55.
Le Bihan, D., Mangin, J., Poupon, C., Clark, C., Pappata, S., Molko, N., Chabriat, H., 2001. Diffusion tensor imaging: concepts and applications. J. Magn. Reson. Imaging 13, 534–546.
Lee, A., Lepore, N., Brun, C., Barysheva, M., Chou, Y., Chiang, M., Madsen, S., McMahon, K., de Zubicaray, G., Wright, M., Toga, A., Thompson, P., 2009. The multivariate A/C/E model and the genetics of fiber architecture. Biomedical Imaging: From Nano to Macro, 2009 IEEE International Symposium on. IEEE, pp. 125–128.
Leow, A., Zhu, S., Zhan, L., McMahon, K., de Zubicaray, G., Meredith, M., Wright, M., Toga, A., Thompson, P., 2009. The tensor distribution function. Magn. Reson. Med. 61, 205–214.
Lepore, N., Brun, C., Chou, Y., Chiang, M., Dutton, R., Hayashi, K., Lu, A., Lopez, O., Aizenstein, H., Toga, A., Becker, J., Thompson, P., 2008. Generalized tensor-based morphometry of HIV/AIDS using multivariate statistics on deformation tensors. IEEE Trans. Med. Imaging 27, 129–141.
Maddah, M., Zollei, L., Grimson, W., Wells, W., 2008. Modeling of anatomical information in clustering of white matter fiber trajectories using Dirichlet distribution. 2008 IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE, pp. 1–7.
Neale, M., Cardon, L., 1992. Methodology for Genetic Studies of Twins and Families. Springer.
Nir, T., Prasad, G., Joshi, S., Villalon, J., Jahanshad, N., Toga, A., Jack, C., Weiner, M., Thompson, P., 2012. Predicting future brain atrophy from DTI-based maximum density path analysis in mild cognitive impairment and Alzheimer's disease. Medical Image Computing and Computer-Assisted Intervention Workshop on Novel Imaging Biomarkers for Alzheimer's Disease and Related Disorders pp. 178–189.
Nir, T., Villalon-Reina, J., Prasad, G., Jahanshad, N., Joshi, S., Toga, A., Bernstein, M., Jack Jr., C.R., Weiner, M., Thompson, P., 2014. DTI-based maximum density path analysis and classification of Alzheimer's disease. Neurobiol. Aging.
O'Donnell, L., Westin, C., 2005. White matter tract clustering and correspondence in populations. Med. Image Comput. Comput. Assist. Interv. 8, 140–147.
O'Donnell, L., Kubicki, M., Shenton, M., Dreusicke, M., Grimson, W., Westin, C., 2006. A method for clustering white matter fiber tracts. Am. J. Neuroradiol. 27, 1032–1036.
O'Donnell, L., Westin, C., Golby, A., 2009. Tract-based morphometry for white matter group analysis. NeuroImage 45, 832–844.
Park, H., Lee, J., 2007. B-spline curve fitting based on adaptive curve refinement using dominant points. Comput. Aided Des. 39, 439–451.
Patel, V., Chiang, M., Thompson, P., McMahon, K., de Zubicaray, G., Martin, N., Wright, M., Toga, A., 2010. Scalar connectivity measures from fast-marching tractography reveal heritability of white matter architecture. Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on. IEEE, pp. 1109–1112.
Peled, S., Gudbjartsson, H., Westin, C., Kikinis, R., Jolesz, F., 1998. Magnetic resonance imaging shows orientation and asymmetry of white matter fiber tracts. Brain Res. 780, 27–33.
Prasad, G., Jahanshad, N., Aganj, I., Lenglet, C., Sapiro, G., Toga, A., Thompson, P., 2011a. Atlas-based fiber clustering for multi-subject analysis of high angular resolution diffusion imaging tractography. Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on. IEEE, pp. 276–280.
Prasad, G., Joshi, S., Jahanshad, N., Villalon, J., Aganj, I., Lenglet, C., Sapiro, G., McMahon, K., de Zubicaray, G., Martin, N., Wright, M., Toga, A., Thompson, P., 2011b. White matter tract analysis in 454 adults using maximum density paths. Medical Image Computing and Computer-Assisted Intervention Workshop on Computational Diffusion MRI pp. 1–12.
Prasad, G., Burkart, J., Joshi, S., Thompson, P., 2013a. A dynamical clustering model of brain connectivity inspired by the n-body problem. Medical Image Computing and Computer-Assisted Intervention Workshop on Multimodal Brain Image Analysis pp. 129–137.
Prasad, G., Joshi, S., Nir, T., Toga, A., Thompson, P., 2013b. Flow-based network measures of brain connectivity in Alzheimer'S disease. 2013 IEEE 10th International Symposium on. Biomedical Imaging (ISBI). IEEE, pp. 258–261.
Prasad, G., Nir, T., Toga, A., Thompson, P., 2013c. Tractography density and network measures in Alzheimer's disease. 2013 IEEE 10th International Symposium on. Biomedical Imaging (ISBI). IEEE, pp. 692–695.
Prasad, G., Joshi, S., Thompson, P., 2014. Optimizing brain connectivity networks for disease classification using EPIC. 2014 IEEE 11th International Symposium on. Biomedical Imaging (ISBI). IEEE, pp. 1–4.
Roberts, T., Liu, F., Kassner, A., Mori, S., Guha, A., 2005. Fiber density index correlates with reduced fractional anisotropy in white matter of patients with glioblastoma. Am. J. Neuroradiol. 26, 2183–2186.
Rohlf, F., Slice, D., 1990. Extensions of the procrustes method for the optimal superimposition of landmarks. Syst. Biol. 39, 40–59.
Sacchet, M., Prasad, G., Foland-Ross, L., Joshi, S., Hamilton, J., Thompson, P., Gotlib, I., 2014. Characterizing white matter connectivity in major depressive disorder: automated fiber quantification and maximum density paths. Biomedical Imaging: From Nano to Macro, 2014 IEEE International Symposium on, IEEE, pp. 1–4.
Schwarz, C.G., Reid, R.I., Gunter, J.L., Senjem, M.L., Przybelski, S.A., Zuka, S.M., Whitwell, J.L., Vemuri, P., Josephs, K.A., Kantarci, K., Thompson, P.M., Petersen, R.C., Jack Jr., C.R., A.D.N. I., 2013. Improved DTI registration allows voxel-based analysis that outperforms tract-based spatial statistics. NeuroImage 94, 65–78.
Shimony, J., Snyder, A., Lori, N., Conturo, T., 2002. Automated fuzzy clustering of neuronal pathways in diffusion tensor tracking. Proc. Intl. Soc. Mag. Reson. Med, p. 1.
Smith, S., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155.
Smith, S., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T., Mackay, C., Watkins, K., Ciccarelli, O., Cader, M., Matthews, P., et al., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. NeuroImage 31, 1487–1505.
Srivastava, A., Joshi, S., Mio, W., Liu, X., 2005. Statistical shape analysis: clustering, learning, and testing. IEEE Trans. Pattern Anal. Mach. Intell. 27, 590–602.
Srivastava, A., Klassen, E., Joshi, S.H., Jermyn, I., 2011. Shape analysis of elastic curves in Euclidean spaces. IEEE Trans. Pattern Anal. Mach. Intell. 33, 1415–1428.

Stein, J., Hua, X., Lee, S., Ho, A., Leow, A., Toga, A., Saykin, A., Shen, L., Foroud, T., Pankratz, N., et al., 2010. Voxelwise genome-wide association study (vGWAS). NeuroImage 53, 1160–1174.

Takahashi, S., Yonezawa, H., Takahashi, J., Kudo, M., Inoue, T., Tohgi, H., 2002. Selective reduction of diffusion anisotropy in white matter of Alzheimer's disease brains measured by 3.0 Tesla magnetic resonance imaging. Neurosci. Lett. 332, 45–48.

Thompson, P., Cannon, T., Narr, K., Van Erp, T., Poutanen, V., Huttunen, M., Lönnqvist, J., Standertskjöld-Nordenstam, C., Kaprio, J., Khaledy, M., et al., 2001. Genetic influences on brain structure. Nat. Neurosci. 4, 1253–1258.

Tuch, D., 2004. Q-ball imaging. Magn. Reson. Med. 52, 1358–1372.

Tustison, N., Avants, B., Cook, P., Kim, J., Whyte, J., Gee, J., Stone, J., 2012. Logical circularity in voxel-based analysis: normalization strategy may induce statistical bias. Hum. Brain Mapp. 1–15.

Villalon-Reina, J., Prasad, G., Joshi, S., Jalbrzikowski, M., Toga, A., Bearden, C., Tompson, P., 2012. Statistical analysis of maximum density path deformation fields in white matter tracts. Medical Image Computing and Computer-Assisted Intervention Workshop on Novel Imaging Biomarkers for Alzheimer's Disease and Related Disorders, pp. 198–209.

Visser, E., Nijhuis, E., Buitelaar, J., Zwiers, M., 2011. Partition-based mass clustering of tractography streamlines. NeuroImage 54, 303–312.

Wakana, S., Jiang, H., Nagae-Poetscher, L., van Zijl, P., Mori, S., 2004. Fiber tract–based atlas of human white matter anatomy. Radiology 230, 77–87.

Wang, X., Grimson, W., Westin, C., 2011. Tractography segmentation using a hierarchical Dirichlet processes mixture model. NeuroImage 54, 290–302.

Wassermann, D., Bloy, L., Kanterakis, E., Verma, R., Deriche, R., 2010a. Unsupervised white matter fiber clustering and tract probability map generation: applications of a Gaussian process framework for white matter fibers. NeuroImage 51, 228–241.

Wassermann, D., Kanterakis, E., Gur, R., Deriche, R., Verma, R., 2010b. Diffusion-based population statistics using tract probability maps. Med. Image Comput. Comput. Assist. Interv. 13, 631–639.

Wedeen, V., Hagmann, P., Tseng, W., Reese, T., Weisskoff, R., 2005. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn. Reson. Med. 54, 1377–1386.

Yeatman, J., Dougherty, R., Myall, N., Wandell, B., Feldman, H., 2012. Tract profiles of white matter properties: automating fiber-tract quantification. PLoS One 7, 1–15.

Yendiki, A., Panneck, P., Srinivasan, P., Stevens, A., Zöllei, L., Augustinack, J., Wang, R., Salat, D., Ehrlich, S., Behrens, T., et al., 2011. Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. Front. Neuroinforma. 5.

Yeo, B., Vercauteren, T., Fillard, P., Peyrat, J., Pennec, X., Golland, P., Ayache, N., Clatz, O., 2009. DT-REFinD: diffusion tensor registration with exact finite-strain differential. IEEE Trans. Med. Imaging 28, 1914–1928.

Yushkevich, P., Zhang, H., Simon, T., Gee, J., 2008. Structure-specific statistical mapping of white matter tracts. NeuroImage 41, 448–461.

Zhan, L., Leow, A., Jahanshad, N., Chiang, M., Barysheva, M., Lee, A., Toga, A., McMahon, K., de Zubicaray, G., Wright, M., et al., 2010. How does angular resolution affect diffusion imaging measures? NeuroImage 49, 1357–1371.

Zhan, L., Jahanshad, N., Jin, Y., Toga, A., McMahon, K., Martin, N., Wright, M., de Zubicaray, G., Thompson, P., 2013. Brain network efficiency and topology depend on the fiber tracking method: 11 tractography algorithms compared in 536 subjects. IEEE International Symposium on Biomedical Imaging, ISBI 2013. IEEE, pp. 1134–1137.

Zhang, H., Yushkevich, P.A., Alexander, D.C., Gee, J.C., 2006. Deformable registration of diffusion tensor MR images with explicit orientation optimization. Medical Image Analysis 10 (5), 764–785.

Zijdenbos, A., Dawant, B., Margolin, R., Palmer, A., 1994. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Trans. Med. Imaging 13, 716–724.