Quantitative descriptions of white matter (WM) fiber shape and cortical folding patterns are important for neuroscience research. This paper presents a novel computational method for WM fiber shape pattern analysis, that is, WM fibers are clustered into five primitive shape patterns: closed ‘U’, ‘M’, curved line, open ‘U’ and straight line, based on the automatic clustering of their shape features. We study the correlation between the fiber shape patterns and the folding patterns of cortical regions that the fibers connect, and find that these patterns are highly correlated. Also, we investigate the shape patterns of the WM fibers that connect different brain regions such as cortical lobes and subcortical regions, and find that the shape patterns of the fibers connecting different brain regions are quite specific. These results might provide insights into the structural and functional architecture of the human brain.

Index Terms—fiber shape, cortical folding patterns, correlation matrix map

1. INTRODUCTION

Recently, there have been growing interests in the shape analysis of white matter fibers [1, 2, 3] obtained from Diffusion Tensor Imaging (DTI) and in folding pattern analysis of the cortex obtained from MRI [4, 5, 6]. However, joint study of the relationship between fiber shape and cortical folding patterns remains to be conducted. Neuroscience research demonstrates that the two biological processes of axogenesis and corticogenesis are highly correlated [7]. Therefore, we hypothesize that the shapes of WM fibers and cortical folding are strongly correlated. This paper aims to test this hypothesis and quantitatively analyze the relationship between fiber shape patterns and cortical folding patterns, which is based on our recently developed approach of parametric representation of folding patterns via polynomials [8].

Based on the shape features of fiber curves, the affinity propagation clustering algorithm [10] is adopted to cluster fibers into five primitive shape patterns: closed ‘U’, ‘M’, curved line, open ‘U’ and straight line. Our results show that this clustering is quite stable across individual brains, and the sensitivity and specificity of the fiber shape classification are quite high. Furthermore, we performed a joint study of the fiber shape patterns and the folding patterns of cortical regions [8] that the fibers connect. Our results show that fiber shape and cortical folding patterns are highly correlated. Also, we investigate the shape patterns of the WM fibers that connect different brain regions such as cortical lobes and subcortical regions, and find that the shape patterns of the fibers connecting different brain regions are quite specific. These results might provide insights into the architecture of the human brain, as well as its development.

2. METHOD

2.1. Preprocessing

Generally, the preprocessing for the fiber curves obtained via DTI tractography [12] consists of cubic spline interpolation, re-sampling, 2D projection and starting point normalization. First, the fiber curves were interpolated using cubic spline interpolation. Second, 100 points were sampled on each of the dense fibers using the constant-speed parameterization [15]. Third, we observed that many features detailed below are based on the notion of clockwise, which is difficult to define in 3D space. However, fibers shape can be mostly preserved if they were properly projected onto a 2D plane. By defining an energy function of a fiber curve that integrates all of its curvedness, the fiber was rotated until its energy on the x-y plane was maximized (Fig.1(a), (b)). To verify our energy preserving assumption, 3000 fibers were randomly selected and projected. On average, 94.81% of energy was preserved. Consequently, we discarded the other two projections on y-z and x-z planes. Finally, some shape features of the open fiber curve depend on the starting point selection. If the starting point cannot be uniquely determined in the two ends, features extracted from the same fiber can be quite different. To solve this problem, we rotated the projections of fibers and set both of their ends on x-axis, and prescribe two criteria for them (Fig.1(b)): 1) the starting points are closer to the gravity centers of the fibers, if not, exchanges the two ends. 2) Fibers are mostly oriented clockwise according to the starting point (green arrow in Fig.1(b)), if not, mirror them along the x-axis. Accordingly, the features will be invariant under rotation and mirroring transformation.

2.2. Feature extraction

We use ‘fiber’ instead of its 2D projection in terminology throughout this paper. For convenience, we present a fiber with N points as a set \( S = \{ s_i \}_{i=1,2,...,N} \). Shape features are extracted from the normalized fibers obtained in section 2.1 with the first and last 10 points of fibers pruned, and the extracted features are both translation- and scaling- invariant. In general, the features can be divided into five categories as described below.

Accumulative Integral of Angles of Tangents: The angles are formed by tangents of any two adjacent points. By calculating the angles from the starting point, we define the angle to be positive/negative if the latter tangent rotates clockwise/counterclockwise from the former one. The \( i \)-th value of the accumulative integral is indeed a summation of the first \( i \) angles. Fig.1(e) illustrates the accumulative integral curve. It is noticed that tangents are very sensitive to fluctuating noise on fibers. For a straight fiber that is contaminated and devious, if the angles have no discrimination in positiveness/negativeness, the accumulative...
integral will increase incorrectly. The accumulative integral, however, cannot yet be directly used as a feature because it is still a curve. Consequently, the Fourier transformation is implemented on the accumulative integral, and amplitudes of its first four frequency components are selected because the rest components attenuate significantly (Fig.1(f)).

\[
C_i = \sum_{m=1}^{M} \begin{bmatrix}
1 & x_{S_{i,m}} & y_{S_{i,m}} \\
1 & x_{S_i} & y_{S_i} \\
1 & x_{S_{i+1}} & y_{S_{i+1}}
\end{bmatrix}
\]

\(S_i\) is convex/concave if \(C_i\) is negative/positive, and the absolute value of \(C_i\) represents the degree of convexity/concavity (Fig.1(c)).

Based on the features described above, we construct a feature vector \(\vec{v}\) containing 12 elements for each fiber: (1) Centroids Ratio; (2) Average Bend and (3) Deviation across the fiber with 10 units segment length; (4) Bend of the entire fiber; (5)-(8) First 4 frequency components’ amplitudes of Fourier transformation of Accumulative Integral of Angles; (9) Average CBA and (10) Deviation; (11) Ratio between summation of convex degrees and summation of concave degrees; (12) Ratio between numbers of convex points and concave points.

2.3. Fiber shape clustering

In order to cluster the fibers by their feature vectors obtained in section 2.2, we adopt the affinity propagation clustering algorithm [10] and design the similarity of two fibers \(i\) and \(j\) as the Mahalanobis distance:

\[
S(i, j) = \sqrt{(\vec{v}_i - \vec{v}_j)^T W (\vec{v}_i - \vec{v}_j)}
\]

where \(W\) is a diagonal matrix in which the diagonal element \(W_{ii}\) is the weight of the \(i\)-th feature; \(\text{Cov}\) is the feature covariance [5].

Computing the similarity matrix for all fibers of one case will entail heavy computation when the affinity propagation algorithm is directly used. Fortunately, when fibers are randomly split into smaller sub-groups, the clustering results across the sub-groups (both the class number and centers) are highly stable. So we can integrate the results by simply regrouping the class centers.

3. EXPERIMENTAL RESULTS

Six cases were randomly selected from a data set obtained from John Hopkins [11], in which high resolution DT-MR images were acquired with the spatial resolution of 0.94×0.94×2.5mm. The fiber trajectories were reconstructed from 3-D diffusion tensor data via the streamline tractography implemented in DTI Studio [12], in which the algorithm stops when it reaches a point with fractional anisotropy (FA) less than 0.25 or when a change in direction is greater than 70°.

3.1. Fiber clustering results

Altogether, six fiber shape patterns, closed ‘U’, ‘M’, curved line, open ‘U’, straight line and unknown shape, are stably obtained when the parameter \(p\) of the affinity propagation algorithm is adjusted within the range (-150, -180). Fibers in last “unknown” class are often the ones that are poorly tracked and link two adjacent fibers at the same ends or the ones that twist into a helix shape, although the fiber trajectories were reconstructed with no tracking direction greater than70°. Interestingly, the proposed method can automatically isolate those fibers and group them into the same class. Because the percentage of those fibers is very small (about 0.000025%) in each case, we can disregard them in our analyses.

Fig.2 shows the clustering results. Fig. 2(a1) and (a2) are the overviews of classified fibers that are color-coded according to
indices in Fig.2(b). For visualization, the fibers are sparsified to 10% of the original number. We compare the five shapes across cases to investigate stability of their proportions. Small standard deviation bars of all types shown in Fig.2(b) demonstrate the reproducibility of the clustering results. Fig.2(c1-c5) detail each fiber shape by randomly selecting some bundles (highlighted in pink color).

To evaluate the accuracy of the method, 4000 fibers are randomly selected. All fibers are manually labeled into the five classes by experts, and the results are used as pseudo ground-truth to compute the sensitivity and specificity. The summary in Table 1 shows high accuracy of the method, given both high sensitivities and specificities across shapes.

### 3.2. Fiber shapes vs cortical folding patterns

In this section, we investigate the relationship between fiber shapes and cortical folding patterns. The triangulated cortical surface mesh is reconstructed based on DTI segmentation [13] for each case. We map the classified fiber ends onto the cortical surface by in-house tools. The mapping algorithm enables the fibers to grow or shrink within a small range until it reaches the surface. Then we label each vertex by voting the density of each fiber shape and code it with the color of the densest fiber shape. The same color coding as Fig.2(b) is used here. The polynomial-based algorithm in [8] is applied to classify the surface vertices into eight folding patterns: peak, ridge, saddle-ridge, inflection, flat, saddle-valley, valley and pit (shown in Fig.3(b)). Fig.3(c) illustrates the fiber shape map, in which the vertices with white color indicate that no fiber ends can reach them by either growing or shrinking.

![Fig.2. Fiber clustering results: (a1) and (a2) Overview of classified fibers; (b) Histograms and standard deviation bars of the five fiber types; (c1) closed ‘U’ shape; (c2) ‘M’ shape; (c3) curved line shape; (c4) open ‘U’ shape; (c5) straight line shape.](image)

![Fig.3. Surface map of cortical folding patterns and fiber shapes: (a) color bar for (b); (b) cortical folding patterns; (c) fiber shapes.](image)

![Fig.4. The correlation matrix map between folding patterns and fiber shapes. The color bar represents the folding patterns as the same order as in Fig.3(a); the gray bar illustrates the connection strength. Abbreviations: SL, straight line; OU, open ‘U’; CU, closed ‘U’.](image)

Furthermore, we investigate which folding patterns the two ends of a fiber connect to and how this connection correlates with fiber shapes. Accordingly, we design a matrix map (Fig.4) in which the rows and columns are arranged in the same folding pattern sequence as the one in Fig.3(a). The cells in the matrix represent the fiber shape information connecting the two corresponding folding patterns. For example, for peak-ridge pair (refer to the color bar in Fig.3(a)) that is located in the cell at the fifth column and the third row of the Fig.4, we pick out the fibers in which the two ends are mapped respectively onto peaks and ridges, and sort the fiber numbers according to shapes. We find that the closed ‘U’ shape fibers prevail in connecting peak and ridge. Finally, we mark this cell with closed ‘U’ (CU) and assign it the number of closed ‘U’ fibers, which is defined as connection strength and color coded with normalized gray bar (right side in Fig.4). After processing all the cells in the similar way, we flip the cells below the diagonal line and add their values to the counterparts, neglecting the connection order. The final matrix map in Fig.4 is the average result of 6 brains.

![Table 1. The statistical measurement of the performance of the classification method.](image)
distributed in the brain; while closed ‘U’ fibers surpass the others in connection strength. Moreover, greater connection strength values are mostly within the box highlighted with white dots, where the cells are coincidently the crossings of peaks, flats, ridges and saddle-ridges, which altogether roughly cover the gyral regions [8]. Hence, we can propose a hypothesis that when the brain develops, line shape fibers might play an important role in setting up the initial cortical gyrification framework.

3.3. Fiber shapes vs lobe localization

In this section we discuss the relationship of fiber shapes and their connections between brain lobes. The lobes are segmented by registering the cases to the template in HAMMER toolkit [14]. Similar to the analysis method in section 3.2, we design a matrix map in which the lower triangular half is omitted (see Fig.5). We sort the lobe sequence and cut the matrix into three parts framed by different colors: the green/blue frame contains cells connecting the lobes within the left/right hemisphere itself; the red frame encircles those connecting different hemispheres. Fig.5. The correlation matrix map between lobes and fiber shapes. Abbreviations: SL, straight line; OU, open ‘U’; CU, closed ‘U’; M,'M'; CL, curved line.

In either green or blue frame, i.e., in the same hemisphere, we can find that close ‘U’ shape is rarely spotted and almost all straight line shape and curved line shape fibers congregate in the subcortical region column and temporal lobe row. That is, almost every straight line fiber locates one end in either subcortical region or temporal lobe, and the other end in other lobes of exactly the same hemisphere. Besides, fibers connecting different lobes in the same hemisphere are more likely of open ‘U’ shape. In the red frame, i.e., between hemispheres, we can find that straight line shape is rarely spotted, but most close ‘U’ shape and ‘M’ shape appear in this area. If the two lobes are close to each other, close ‘U’ shape will connect them, or if they are further apart, they are connected by open ‘U’ shape. Moreover, fibers with one ends in the subcortical regions are ‘M’ shape when connecting to regions in the other hemisphere. In addition, in terms of the symmetry line (the diagonal line) through the red frame and the similarity between green frame and blue frame, we can conclude that fiber shapes are symmetrically distributed in the brain.

4. DISCUSSION AND CONCLUSION

We present a novel method to classify fibers based on inherent shape features in this paper. Its effectiveness and accuracy is demonstrated by the statistical measurements. Furthermore, the method is self-contained and independent of atlases or ROIs, so that the analysis result is reproducible. Above all, the method takes the brain as a complete entity and the results are correlated with the cortical representation, including folding patterns and lobe localization. With the assistance of correlation matrix maps we build up the look-up-tables which are rich in correlation information, and might provide clues for exploring the architecture and development mechanisms of the human brain. The method can be applied to current bundle-based fiber classification methods for further analysis, e.g., as a pre-processing step for bundle clustering. We also hope that the method can be potentially useful for studies of abnormal fiber patterns in many brain diseases. In the future, more effective shape features will be considered in the light of their contributions to the clustering performance.

5. REFERENCE