Review of Fast Density-Peaks Clustering and Its Application to Pediatric White Matter Tracts

Shichao Cheng[1], Yuzhuo Duan[1], Xin Fan[1], Dongyu Zhang[1], and Hua Cheng[2(✉)]

[1] School of Software, Dalian University of Technology, Dalian, China
xin.fan@ieee.org
[2] Medical Imaging Center, Beijing Childrens Hospital, Beijing, China
chhuaer@hotmail.com

Abstract. Clustering white matter (WM) tracts from diffusion tensor imaging (DTI) is primarily important for quantitative analysis on pediatric brain development. A recently developed algorithm, density peaks (DP) clustering, demonstrates great robustness to the complex structural variations of WM tracts without any prior templates. Nevertheless, the calculation of densities, the core step of DP, is time consuming especially when the number of WM fibers is huge. In this paper, we propose a fast algorithm that accelerates the density computation about 50 times over the original one. We convert the *global* calculation for the density as well as critical parameter in the process into *local* computations, and develop a binary tree structure to orderly store the neighbors for these local computations. Hence, the density computation turns out to direct access of the structure, rendering significantly computational saving. Experiments on synthetic point data and the JHU-DTI data set validate the efficiency and effectiveness our fast DP algorithm compared with existing clustering methods. Finally, we demonstrate the application of the proposed algorithm on the analysis of pediatric WM tract development.

Keywords: Fast clustering · White matter tracts · DTI

1 Introduction

Diffusion-tensor imaging (DTI), capturing the diffusion of water molecules in human brain, provides a non-invasive assessment of microstructural organization of white matter (WM) [1]. Fiber bundles detected by DTI typically form a set of tract bunddles that share common neurophysiological connectivity and neuropsychiatric function. For example, the inferior longitudinal fasciculus (ILF) connects ipsilateral temporal and occipital lobes, and integrates auditory and speech nuclei [2]. There exist numerous studies investigating the integrity evolution of WM pathways in human brain via these clustered WM tracts. One interesting aspect for these studies is to discover the development process of WM tracts [3,4]. However, classical methods typically analyze spatially averaged parameters transformed into a common template space by registration

M. Valdés Hernández and V. González-Castro (Eds.): MIUA 2017, CCIS 723, pp. 436–447, 2017.
DOI: 10.1007/978-3-319-60964-5_38

algorithms that smooth local alterations along tracts. In this paper, we develop a fast algorithm to automatically cluster fibers into WM bundles in the original space for tract-wise development analysis [5,6].

In the past decade, researchers resorted to classical clustering algorithms in the machine learning community, including K-means [7,8], Spectral Clustering (SC) [9], Gaussian mixture models (GMM) [10], Affinity Propagation (AP) [11,12], and DBSCAN [13], to address the issue of tract clustering. These clustering algorithms gained some success as they grouped fiber tracts having similar intrinsic characteristics. However, one major drawback of these methods lies in that they need to specify the number of clusters. A wrong configuration of this parameter can lead to inaccurate segmentation. Recently, prior knowledge of brain anatomy is incorporated into clustering to guide the tract segmentation process. Li *et al.* adopted a hybrid approach for automatic clustering white fibers, using a top-down atlas [14], while Jin *et al.* fuse labels from multi-atlas for tract clustering [15]. Considering variations among individuals especially pediatric ones undertaking rapid developments, it is an open issue how largely the bundle clustering specific to a subject can rely on a universal atlas. Additionally, little attention has been paid to fiber outliers, greatly affecting the classical algorithms.

Those problems are solved by a newly developed algorithm, density peaks (DP) clustering, is able to tackle the challenges for clustering WM tracts [16]. The algorithm discovers cluster centers and assigns cluster labels based on the density per data point. Users can intuitively determine the number of clusters from a graph characterizing the density distribution of a subject's fibers, so that need no prior atlas. The density graph also removes the outliers by specifying the border density for each cluster. The outputs of the algorithm show high consistency with manual parcellation [17,18]. Unfortunately, the generation of density graph has to numerate all fiber pairs (the square of the fiber count) in order to generate the decision graph, resulting in expensive computational load for over $100,000$ fibers.

In this paper, we propose a fast algorithm for the density-peaks clustering, and apply the algorithm to the analysis of WM tract development. We observe that the density of a data point largely relies on its neighboring points, and the farther points including outliers have few effect on density calculation. This observation enables us to use *local* points to compute the *global* density of a point over all the other points. The neighboring points (fibers) of any individual fiber are orderly stored in a binary tree. Hence, a one-pass traverse on the tree generates the densities for all fibers, and thus circumvents the repeated calculations on all point pairs. This computation process on the tree significantly reduces the expensive computational expenses on the density calculation in the original DP algorithm, yet yields an almost equivalent density graph. We validate our fast DP on the synthetic data and DTI images, and compare the performance with the classical K-means, GMM, Mean-shift, SC, AP, and DBSCAN algorithms on the JHU-DTI data set (http://lbam.med.jhmi.edu/). Finally, we apply the fast DP to automatically cluster several WM tracts, and investigate their development across 18 to 36 months. This automatic process produces coincide findings with the latest results discovered by an atlas-based approach [4].

2 DTI Preprocessing

We downloaded DTI images from JHU brain MRI laboratory[1], and randomly selected 12 subjects for algorithm validation. Clinical pediatric images in our brain development experiments were obtained from 3 Tesla(T) Philips scanner with 30 or 31 diffusion gradient directions, manually inspected by radiologists, and no structural abnormality was found. DtiStudio[2] [19] and FSL[3] [20] were mainly used for the preprocessing.

DTI data contains 35 MRI images and the first 5 images are minimally diffusion-weighted images. All the raw diffusion-weighted images (DWIs) were coregistered to the first minimally diffusion-weighted images to correct participant's motion effect and the linear portion of eddy current distortions. A 12-parameter affine transformation of automated image registration AIR[4] accompanied in Dtistudio was used for the registration [21]. Coregistered DWIs were sent to brain extraction tool (BET[5]) to strip brain's skull and outer tissues [22]. Option "Apply to 4D FMRI data" was chose and parameter "Fractional intensity threshold" was set as 0.5. In order to keep the conforming characteristics of all subjects' image, sampling function of DiffeoMap[6] was used to resample images to a specified resolution, image resolution $181 * 217 * 181\,\mathrm{mm}^3$ and physical resolution $1.00 * 1.00 * 1.00\,\mathrm{mm}^3$ was adopted. DTI mapping of DtiStudio then was used on the resampled DWIs, tensor data was generated. Images of DTI derived parameters including FA and ADC were also preserved in this step for further analysis.

For each subject, after its tensor data was generated, which was sent into DtiStudio to do fiber tracking based on Fiber Assignment by Continuous Tracking (FACT) algorithm and a brute-force reconstruction approach. We set the fiber tracking starting from all seed points with FA value larger than 0.4, and stopping when reaching one voxel with FA value smaller than 0.25 or direction change larger than 42°. These values were chosen according to practical experiment in which we can tract some immature fibers better. Then we draw a crossing area containing ILF, IFO and Fmajor named ROI fibers according to the protocol in [2] which can be seen in Fig. 1 **A**. The complete fiber tracts generated by fiber tracking include more than 100,000 fibers. We removed those fibers shorter than 35 mm as considered as "artifact" [23].

3 Fast Density-Peaks Algorithm

In this section, we briefly introduce the density peaks clustering [16] with the analysis on its bottleneck for efficiency, and then provide our new algorithm illustrated in Fig. 1 as well as its complexity analysis.

[1] http://lbam.med.jhmi.edu.
[2] https://www.mristudio.org.
[3] http://fsl.fmrib.ox.ac.uk.
[4] http://bishopw.loni.ucla.edu/air5.
[5] http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET.
[6] https://www.mristudio.org.

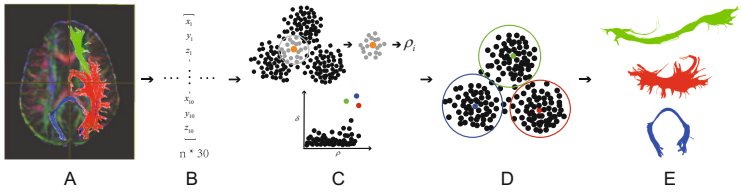

Fig. 1. Pipeline of fiber clustering. The crossing area named ROI fibers can be segmented into 3 clusters correctly including Fmajor (red), IFO (green) and ILF (red) by FDP algorithm, this figure shows procedures including **A**. tractography (to draw a crossing area contains Fmajor, ILF and IFO), **B**. fiber representation, **C**. calculate local density of each fiber to choose clustering center, **D**. fiber clustering and **E**. visualization results, respectively. (Color figure online)

3.1 Density Peaks Clustering

The DP clustering algorithm finds density peaks as cluster centers, and assigns labels for every point (fiber in our context) based on the similarity and density with the neighboring points [16]. Cluster centers are those points as following:

- the density of a cluster center should be higher than its neighbors;
- a cluster center ought to be far away from other points with higher density than itself.

Therefore, the calculation for the density of per point is the core of DP.

Considering a set of fibers $\{f_1, f_2, ..., f_n\}$, the distance between f_i and f_j is denoted as d_{ij} whose definition can be found in [17]. The density $\bar{\rho_i}$ of f_i is defined as:

$$\bar{\rho_i} = \sum_j \exp(-\frac{d_{ij}^2}{d_c^2}), \tag{1}$$

where d_c is the cutoff distance. Another critical attribute parameter δ_i is defined as the minimum distance from f_i to all other fibers whose density is higher than ρ_i:

$$\delta_i = \min_{j:\rho_j > \rho_i} (d_{ij}). \tag{2}$$

The clustering process using these two attributes is given as follows: (1) determine d_c; (2) compute $\bar{\rho}$ for each fiber according to Eq. 2; (2) compute δ for each fiber and keep the nearest fiber with higher density as its neighbor; (3) select cluster centers according to the decision graph, given by the values of $\bar{\rho}$ and δ; (4) assign each fiber to the same cluster of its neighbor according to fiber density descending order; (5) calculate border density for each cluster, those fibers with smaller density than border density are set as outliers.

Complexity analysis:

- The choice of d_c is so critical that we cannot specify its value in an *Ad Hoc* way but have to derive from the data set itself. Typically, the value is the tail

in the top 5% of all distances d_{ij} in the ascending order. We have to sort the sequence of all distance pairs at the complexity of $O(n^2 \log(n^2))$, where n is the number of fibers obtained by tractography.
– For the calculation of $\bar{\rho}$ for a fiber, say f_i, we need to sum all the distances between f_i and other fibers, $d_{ij}(j = 1, ..., n, i \neq j)$. Consequently, the total complexity for $\bar{\rho}$ turns out to be $O(n^2)$.

3.2 Fast Algorithm

The bottleneck to efficiently compute the density for one point (fiber) lies in the enumeration of all the others. Actually, the nearer a fiber is, the more contribution it makes as the contributions of other points to the density exponentially decay with the distances to the point of interest. As shown in the above subsection, $\exp(-\frac{d_{ij}^2}{d_c^2})$ is smaller than 0.02 even though $d_{ij} = 2d_c$, and furthermore, $\exp(-\frac{d_{ij}^2}{d_c^2})$ is so subtle as 10^{-11}. We only need to consider a small fraction of neighboring points (the K nearest neighbors) as the approximation of the density $\bar{\rho}$ for the fiber:

$$\rho_i = \sum_{j \in A} \exp(-\frac{d_{ij}^2}{d_c^2}), A = \{j | j \in \text{K nearest neighbors of fiber } i\}. \tag{3}$$

The approximation enables us to use *local* points to compute the *global* density of a fiber. Distances between every two fibers have been obtained and the neighboring fibers of any individual fiber are orderly stored in leaf nodes of a binary tree. Hence, a one-pass traverse on the tree searching nearer neighbors for all fibers without sorting all distances data, and circumvents the repeated calculations on all fiber pairs. This computation process on the tree significantly reduces the expensive computational cost on the density calculation in the original DP algorithm, but yields an almost equivalent decision graph.

We take a point i as an example and Fig. 2 to illustrate the process. Firstly, we calculate the mean distance over all distances between the fiber i and another, and keep the data points with the distances less than the mean shown as the orange dots in Fig. 2). Then, we update the mean value for current set(orange ones), and finds those points closer to fiber i than the updated mean. These mean updating and neighbor searching processes runs iteratively until K neighbors of fiber i are found as colored in green. The upper row of Fig. 2 demonstrates this iterative process. All the distances between fiber i and the others can be stored in a binary tree where the left subtree always store distances smaller than the updated mean. So we only transverse on left subtree reaching the most left leaf to get K nearest neighbors of any fiber within time complexity of $O(\log n)$. Then the total computation is significantly reduced by this strategy. Both the storage and computational complexity reduce from n to K for each fiber's density.

Subsequently, we use these K neighbors to calculate ρ. The value of ρ is smaller than $\bar{\rho}$ but make few effects on the distribution of density. It can not only preserve the density information, but also, actually more importantly, lowers

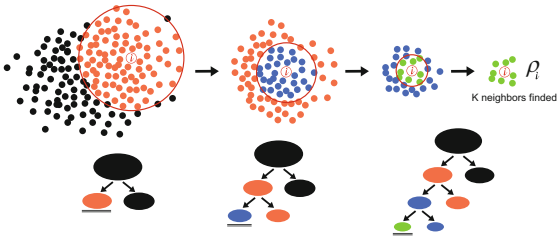

Fig. 2. Neighbor-searching process. Find the K nearest neighbors of one data point by iteratively updating mean distances via a binary tree. The red circle denotes the mean distance for the current set. The bottom row gives the binary tree for the searching process.

down the complexity for ρ calculation of n points from $O(n^2)$ to $O(Kn)$. We validate its effectiveness and efficiency in Sect. 4

The cutoff distance d_c can be calculated by this strategy instead of using all the other distances. We can obtain d_c as the top 5% of the K nearest neighbors given the tree instead of sorting all distances. We generate an equivalent d_c value, but the time cost for sorting can be reduced to $O(log(n^2))$.

Table 1 compares the Algorithm complexity of DP and our improved FDP in d_c and ρ calculation.

Table 1. Algorithm complexity comparison between DP and FDP. n represents the number of data point

	DP	FDP
ρ	O(n^2)	O(Kn)
d_c	O($n^2\ log(n^2)$)	O($log(n^2)$)

4 Clustering Results on Synthetic and Benchmark Sets

We validate the proposed FDP algorithm on several synthetic data sets widely used in the machine learning community to demonstrate the accuracy of FDP equivalent to DP, but extremely faster implementation. We also compared FDP with other classical clustering algorithms on clustering fibers on a benchmark set.

4.1 Comparisons Between FDP and DP

We apply FDP and DP on the sets including Aggregation, R15, Compound, and Jain, showing complex distributions[7]. Figure 3 shows the original data, and their corresponding decision graphs derived from DP and FDP, respectively. The decision graphs present quite close with unnoticeable differences, and accordingly their accuracies are almost the same as given in Table 2.

[7] http://cs.joensuu.fi/sipu/datasets.

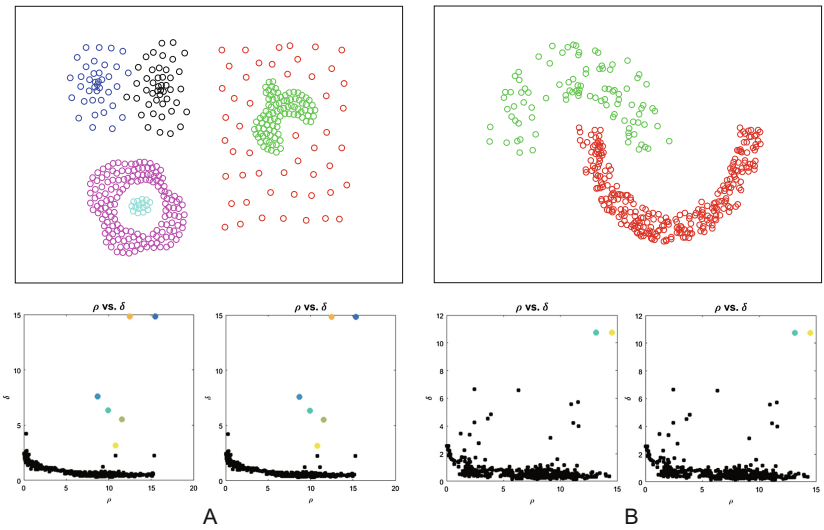

Fig. 3. Decision graphs comparison between DP and FDP in the 2 datasets. Part A and B upper panel are Compound, jain data, lower panel displays their decision graphs (DP in the left and FDP in the right side) respectively.

Time costs of FDP considerably decrease compared with DP. We apply DP and FDP on 3 subjects from the JHU dataset, and calculate the averaged time expenses listed in Table 3. All the time expenses for distance sorting, d_c and ρ calculating significantly decrease. For more than 10^3 fibers, DP takes 10 times computations to FDP on the distance sorting and d_c calculating. More significantly, the time for calculating ρ of FDP is only 1/100 of DP.

4.2 Comparisons with Classical Cluster Algorithms

We compare FDP with six classical clustering algorithms, including SC, GMM, K-means, Mean-shift, Affinity AP, and DBSCAN on the JHU-DTI dataset. For three fiber tracts of interest, we calculate the Dice ratio (DR) of the resultant bundles to manually labeled ones (ground truth) following the protocol in [2]. Figure 4 shows the box plots of the Dice ratios for clustering algorithms, where

Table 2. Accuracy comparison between DP and FDP. The 4th and 5th columns shows the results corresponding to A, B part of Fig. 3.

Accuracy(%) Data / Method	Aggregation	R15	Compound	Jain
DP	100	99.83	87.47	85.52
FDP	100	99.67	87.72	85.52

Table 3. Time expenses for DP and FDP.

Subject	Distance sorting (s)		d_c (s)		ρ (s)	
	DP	FDP	DP	FDP	DP	FDP
1	0.2040	0.0069	0.2068	0.0090	0.2118	0.0167
2	0.2972	0.0312	0.3070	0.0393	0.5344	0.0723
3	0.1259	0.0175	0.1312	0.0210	0.3141	0.0369

the dotted line, blue rectangle and red short line denote the data range, the top 25% to 75% bounds, and the mean of DR, respectively. The mean value close to 1 as well as narrow bounds indicate high consistency with the ground truth. For the tracts of IFO and ILF, FDP has obvious more consistent overlapping with the ground truth than the other four algorithms listed in Fig. 4. For Fmajor, FDP outputs higher mean DR value than the others, but works a bit less stable than the mean-shift. The visualization results in Fig. 5 also demonstrate the clustering outputs of FDP are the most consistent with the manually labeled ones. Detailed discussions are given below.

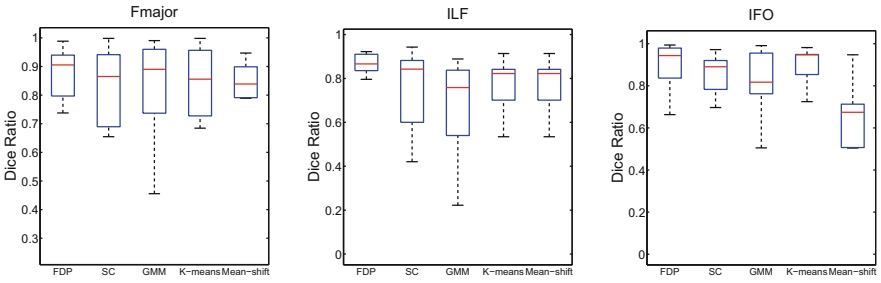

Fig. 4. Fiber's DR of different clustering algorithms. FDP stands for Fast density-peaks clustering.

SC applies the eigen decomposition to the similarity matrix of data points, and then clusters the eigen vectors derived from the decomposition. In Fig. 4, SC results have close mean values with those of k-means, but work less stable than k-means and ours. This can be explained by the crossing regions exist in the right part of Fmajor highlighted by yellow in the SC results of Fig. 5.

The solution to GMM is found by the expectation maximization (EM) algorithm, which may fall into local extremes. Also, correct initialization is quite critical to the solution. Thus, the mean values of the algorithm fall as the median of the compared algorithms, but it works quite unstable as shown by larger bounds in Fig. 5.

K-means outputs close performance to ours for IFO, but worse on Fmajor and ILF. The computational load of the classical k-means are quite high. There exists

accelerated version of k-means in the context of fiber clustering [8]. However, the number of clusters cannot be set by an intuitive way as FDP does. Additionally, FDP is able to locate the 'core' fiber that is important for statistical analysis on WM tracts, but k-means only gives a mean of the clusters.

The Mean-shift algorithm finds cluster centers having the largest density. It applies to the case when probability density function has only one extreme in a local region. As we can see from Fig. 4, the mean-shift is not suitable for fiber clustering especially for IFO and Fmajor in Fig. 5. Also, the crossing area (right side of Fmajor) cannot be well separated by the mean shift. Fibers in the crossing region also have relatively larger densities so that the algorithm may converge to the local optima giving wrong clustering centers.

AP is famous for its automatic clustering without the need of setting the cluster number *a priori*. Unfortunately, AP cannot automatically find a correct cluster number in our experiments. AP is quite sensitive to outliers so that identify much more clusters than three in our experimental setup. Also, over one hundred times of iterations are necessary for clustering one tract, spending extremely long time to generate a plausible solution.

DBSCAN shares the similar philosophy with FDP in the sense that both focus on density of data points. The parameter setting has a great influence on the output. These parameters include the number of neighbors and neighboring radius. We found in our experiments that slight changes on these parameters output significant variations on the clustering. Thus, we can hardly find a stable and plausible solution on these three tracts of interest.

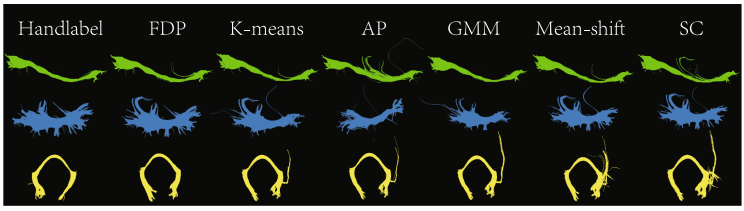

Fig. 5. Fiber clustering results visualization. It illustrates fiber clustering results using FDP, K-means, Affinity Propagation (AP), Gaussian Mixture Model (Gmm), Mean-shift and Spectral Clustering (SC) algorithms respectively. Take handlabel tracts as ground truth.

5 Tract-Based Analysis on Pediatric Brain Development

We analyze the development on the clustered tracts, Fmajor, ILF and IFO, given by FDP, and provide the preliminary results below.

5.1 Fiber Core Properties

Using FDP clustering method, there will be a core fiber tract for each fiber bundles shown in Fig. 6. We take this core fiber's characteristic to describe the

property of the clustered fiber bundles [24]. We sampled each fiber tracts into ten 3D points. As the points in the boundary is not very stable compared with points in the middle of fiber tracts. Diffusion properties are then calculated for the inner six points of each fiber tract. Every point's property was taken by interpolating the image value (FA, MD).

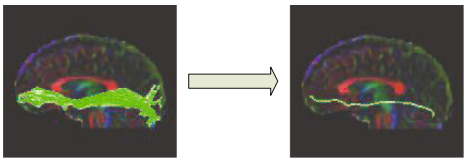

Fig. 6. Fiber bundel core. Inferior fronto-occipital fasciculus (IFO)'s core fiber tract. Statistics is performed along this fiber.

5.2 Tract-Wise Development Analysis

About 11 healthy subjects aged 18–36 months participate in this investigation. We calculate FA value of the core fiber tract for each fiber bundles. The concrete FA value and the correlation coefficient are different, but there is no doubt that they both have the similar increasing trend. P values are small than 0.05 in all groups, it indicates that there are significant correlations between ages and bundles' FA through center fiber. The results suggest that the progressive increase of FA values in each fiber bundles (including ILF, IFO and Fmajor) are observed for healthy Pediatrics with significant correlations ($P < 0.05$ for all scatter plots in Fig. 7) between FA and age. It maintains highly consistent with the conclusion of clinic and other researchers' study [4].

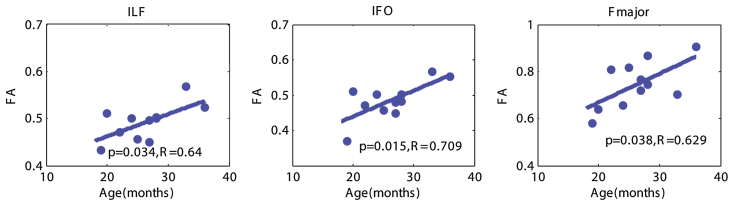

Fig. 7. Tract-based Analysis results according to core fibers.

6 Conclusion

In this paper, we propose a fast algorithm for clustering based on density peaks, applying the algorithm to segment fibers and analyzing the development of WM tract. We also propose using K nearest neighbors to calculate density. Experimental results show that Fast density-peaks clustering has a prefect performance

in accuracy and algorithm accelerating. In the future, we will evaluate our method using more DTI data, and explore functions of different fiber bundles in Pediatric's brain.

Acknowledgments. We would like to thank Dr. Hua Cheng from Beijing Childrens Hospital for providing research data. It's acknowledged that paediatric data used for the experiment has been granted by the children's legal guardians.

References

1. Basser, P.J., Mattiello, J., Lebihan, D.: Mr diffusion tensor spectroscopy and imaging. Biophys. J. **66**(1), 259–267 (1994)
2. Wakana, S., Caprihan, A., Panzenboeck, M.M., Fallon, J.H., Perry, M., Gollub, R.L., Hua, K., Zhang, J., Jiang, H., Dubey, P., Blitz, A., van Zijl, P., Mori, S.: Reproducibility of quantitative tractography methods applied to cerebral white matter. NeuroImage **36**, 630–644 (2007)
3. Lebel, C., Gee, M., Camicioli, R., Wieler, M., Martin, W., Beaulieu, C.: Diffusion tensor imaging of white matter tract evolution over the lifespan. Neuroimage **60**, 340–352 (2012)
4. Ouyang, M., Cheng, H., Mishra, V., Gong, G., Mosconi, M.W., Sweeney, J., Peng, Y., Huang, H.: Atypical age-dependent effects of autism on white matter microstructure in children of 2–7 years. Hum. Brain Mapp. **37**, 819–832 (2015)
5. Jin, Y., Huang, C., Daianu, M., Zhan, L., Dennis, E.L., Reid, R.I., Jack, C.R., Zhu, H., Thompson, P.M.: 3D tract-specific local and global analysis of white matter integrity in Alzheimer's disease. Human Brain Mapp. **38**(3), 1191–1207 (2016)
6. O'Donnell, L.J., Golby, A.J., Westin, C.-F.: Fiber clustering versus the parcellation-based connectome. NeuroImage **80**, 283–289 (2013)
7. O'Donnell, L., Westin, C.-F.: White matter tract clustering and correspondence in populations. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3749, pp. 140–147. Springer, Heidelberg (2005). doi:10.1007/11566465_18
8. Reichenbach, A., Goldau, M., Heine, C., Hlawitschka, M.: V–bundles: clustering fiber trajectories from diffusion MRI in linear time. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A. (eds.) MICCAI 2015. LNCS, vol. 9349, pp. 191–198. Springer, Cham (2015). doi:10.1007/978-3-319-24553-9_24
9. O'Donnell, L.J., Kubicki, M., Shenton, M.E., Dreusicke, M.H., Grimson, W.E.L., Westin, C.-F.: A method for clustering white matter fiber tracts. Am. J. Neuroradiol. **27**(5), 1032–1036 (2006)
10. Wang, Q., Yap, P.T., Wu, G., Shen, D.: Application of neuroanatomical features to tractography clustering. Hum. Brain Mapp. **34**, 2089–2302 (2013)
11. Zhang, T., Chen, H., Guo, L., Li, K., Li, L., Zhang, S., Shen, D., Xiaoping, H., Liu, T.: Characterization of u-shape streamline fibers: methods and applications. Med. Image Anal. **18**(5), 795–807 (2014)
12. Jin, Y., Cetingül, H.E.: Tractography-embedded white matter stream clustering. In: 2015 IEEE 12th International Symposium on Biomedical Imaging, pp. 432–435 (2015)
13. Mai, S.T., Goebl, S., Plant, C.: A similarity model and segmentation algorithm for white matter fiber tracts. In: 2012 IEEE 12th International Conference on Data Mining, pp. 1014–1019 (2012)

14. Li, H., Xue, Z., Guo, L., Liu, T., Hunter, J., Wong, S.T.C.: A hybrid approach to automatic clustering of white matter fibers. NeuroImage **49**(2), 1249–1258 (2010)
15. Jin, Y., Shi, Y., Zhan, L., Gutman, B.A., de Zubicaray, G.I., McMahon, K.L., Wright, M.J., Toga, A.W., Thompson, P.M.: Automatic clustering of white matter fibers in brain diffusion MRI with an application to genetics. NeuroImage **100**, 75–90 (2014)
16. Rodriguez, A., Laio, A.: Clustering by fast search and find of density peaks. Science **344**, 1492–1296 (2014)
17. Chen, P., Fan, X., Liu, R., Tang, X.: Fiber segmentation using a density-peaks clustering algorithm. Biomed. Imaging **20**, 633–637 (2015)
18. Kamali, T., Stashuk, D.: Automated segmentation of white matter fiber bundles using diffusion tensor imaging data and a new density based clustering algorithm. Artif. Intell. Med. **73**, 14–22 (2016)
19. Jiang, H., van Zijl, P.C.M., Kim, J., Pearlson, G.D., Mori, S.: Dtistudio: resource program for diffusion tensor computation and fiber bundle tracking. Comput. Methods Programs Biomed. **81**, 106–116 (2006)
20. Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M.: Fsl. NeuroImage **62**, 782–790 (2012)
21. Woods, R.P., Grafton, S.T., Watson, J.D.G., Sicotte, N.L., Mazziotta, J.C.: Automated image registration: Ii. intersubject validation of linear and nonlinear models. J. Comput. Assist. Tomogr. **22**, 153–165 (1998)
22. Smith, S.M.: Fast robust automated brain extraction. Hum. Brain Mapp. **17**, 143–155 (2002)
23. Mayer, A., Zimmerman-Moreno, G., Shadmi, R., Batikoff, A., Greenspan, H.: A supervised framework for the registration and segmentation of white matter fiber tracts. IEEE Trans. Med. Imaging **30**, 131–145 (2011)
24. Johnson, R.T., Yeatman, J.D., Wandell, B.A., Buonocore, M.H., Amaral, D.G., Nordahl, C.W.: Diffusion properties of major white matter tracts in young, typically developing children. NeuroImage **88**, 143–154 (2014)