Grid-Based Spectral Fiber Clustering

Jan Klein[a] and Philip Bittihn[b] and Peter Ledochowitsch[b] and Horst K. Hahn[a]
and Olaf Konrad[a] and Jan Rexilius[a] and Heinz-Otto Peitgen[a]

[a]MeVis Research, Center for Medical Image Computing, Bremen, Germany
[b]Goettingen University, Department of Physics, Goettingen, Germany

ABSTRACT

We introduce novel data structures and algorithms for clustering white matter fiber tracts to improve accuracy and robustness of existing techniques. Our novel fiber grid combined with a new randomized soft-division algorithm allows for defining the fiber similarity more precisely and efficiently than a feature space. A fine-tuning of several parameters to a particular fiber set - as it is often required if using a feature space - becomes obsolete. The idea is to utilize a 3D grid where each fiber point is assigned to cells with a certain weight. From this grid, an affinity matrix representing the fiber similarity can be calculated very efficiently in time $O(n)$ in the average case, where n denotes the number of fibers. This is superior to feature space methods which need $O(n^2)$ time. Our novel eigenvalue regression is capable of determining a reasonable number of clusters as it accounts for inter-cluster connectivity. It performs a linear regression of the eigenvalues of the affinity matrix to find the point of maximum curvature in a list of descending order. This allows for identifying inner clusters within coarse structures, which automatically and drastically reduces the a-priori knowledge required for achieving plausible clustering results. Our extended multiple eigenvector clustering exhibits a drastically improved robustness compared to the well-known elongated clustering, which also includes an automatic detection of the number of clusters. We present several examples of artificial and real fiber sets clustered by our approach to support the clinical suitability and robustness of the proposed techniques.

Keywords: Visualization, Fiber Clustering, Fiber Tracking, Diffusion Imaging Techniques

1. INTRODUCTION

Over the last few years, diffusion imaging techniques like diffusion tensor imaging (DTI), diffusion spectrum imaging (DSI) or Q-ball imaging (QBI) received increasing attention, especially in the neurosurgical community with the motivation to identify major white matter tracts afflicted by an individual pathology or tracts at risk for a given surgical approach.[1,2]

An explicit geometrical reconstruction of major white matter tracts has become available by fiber tracking (FT) based on the reconstructed tensor field.[3,4] Rather than requiring manual segmentation on every image slice, FT uses the directional information of the diffusion tensor to trace diffusion paths in 3D starting from a seed region.[4] For visualizing the fibers, the idea of streamtubes along with different color coding attributes (e.g., direction, uncertainty[5,6]) has been discussed.[7–10]

Fiber clustering offers an opportunity for improving the perception of the fiber bundles and connectivity in the human brain by grouping *anatomically similar* or *related* fibers. Due to the large number of points constituting each single fiber, the problem of efficiently computing the similarity between fibers is usually solved by using a feature space (FS).[11,12] FS methods map the high dimensional data to a low dimensional FS from which the affinity matrix is calculate by applying a distance function. Thus, each fiber is more or less well-characterized by a small set of calculated features like co-variance, main direction or the center of gravity.

Given the affinity matrix, a clustering algorithm can be applied. A recently proposed approach[12] recursively determines a minimal normalized cut to partition a fiber similarity graph into two subsets. This is done until the value of the cut is larger than a predefined threshold on which the number of clusters depends. Also the

Further author information: (Send correspondence to Jan Klein)
Jan Klein: E-mail: klein@mevis.de, Telephone: 49 421 218 8902

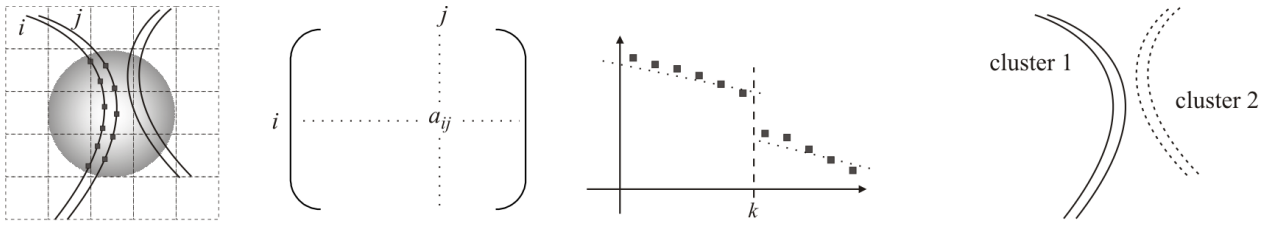

Figure 1. Overview of data processing for spectral fiber clustering.

unsupervised fuzzy c-means clustering[13] or the k-nearest neighbor approach[14] utilizing a segment-based similarity index have been proposed for fiber clustering. O'Donnell et al. use a k-way normalized cuts clustering algorithm based on a modified Hausdorff distance for computing the similarity between paths.[15]

Jonasson et al.[16] consider the problem of clustering fiber tracts obtained from high angular resolution diffusion images where more than one fiber orientation within a single voxel is possible. Instead of a FS, they use a voxel grid to determine a co-occurrence matrix for clustering. El Kouby et al.[17] developed a method where the connectivity information of ROIs defined by a Talairach grid is used for clustering. Similar to FS methods, this reduces the precision of the fiber similarity.

Other approaches propose a B-spline representation of the fiber tracts for a pairwise comparison[18] or a soft clustering based on pseudo coloring.[19] A method for evaluating fiber clustering algorithms was proposed by Moberts et al.[20] Outside the area of medical imaging, clustering algorithms are used, e.g., for analyzing gene expression data[21] or processing of point-sampled geometry.[22,23]

In this paper, we introduce a novel framework that allows for fiber clustering with an automatic determination of the number of clusters. For that purpose, we propose to use spectral clustering on top of an affinity matrix derived from the fiber grid (an outline of our framework is given in Figure 1).

We show that our new grid-based approach described in Section 2 defines fiber similarity more precisely than FS methods. In Section 3 we show how the cluster number is determined by our new eigenvalue regression method. Moreover, our clustering algorithms, optimizations and adaptions of two image segmentation methods[24,25] are briefly summarized. In Section 4 results are presented which show that our algorithms are more robust than Elongated k-means[25] (automatically determining the number of clusters). We conclude with a discussion and show possible directions for future work.

2. FIBER SIMILARITY

In contrast to a feature space (FS) — which neglects most of the metrical information about the fibers for the sake of a low computational time — the idea of the fiber grid (FG) is to gather more accurate spatial information in order to calculate a fiber affinity matrix. An obvious approach to account for the metrical information would be a point-by-point distance determination between all fibers which is in $O(n^2)$ and thus is too expensive for treating large fiber sets. By approaching the distance problem in a different way, the FG not only manages to preserve the benefits of accurate spacial information but is also in $O(n)$ which is even faster than the FS method.

Instead of comparing the fibers pointwise, the space can be divided into 3D cells of equal size. For simplicity, we chose a cubic grid with a side length of d. Then, each fiber point is assigned to the cell containing it, recording a label for the fiber it belongs to and a weight, depending on the assignment method (described later in this section). Using a suitable mapping of the fiber points to the cell indices, the assignment of one point takes constant time. Therefore, all points can be assigned to their corresponding cells in $O(n)$. Let $w_i(x,y,z)$ denote the sum of all weights stored in cell (x,y,z) arising from fiber i. The affinity matrix A with entries a_{ij} can now easily be calculated by increasing a_{ij} if fiber i and j contain data points that are assigned to the same cell:

$$a_{ij} \quad := \quad \sum_{(x,y,z)} w_i(x,y,z) \cdot w_j(x,y,z) \qquad i \neq j$$

$$
\begin{aligned}
g(x) &:= \arctan\left(\sigma\left(x - \frac{d+1}{2}\right)\right) \\
a &:= \frac{w_e - w_m}{g(0) + g(1) - 2g(0.5)} \\
b &:= \frac{2g(0.5)\cdot(w_m - w_e)}{g(0) + g(1) - 2g(0.5)} + w_m \\
f(x) &:= a\,(g(x) + g(1-x)) + b
\end{aligned}
$$

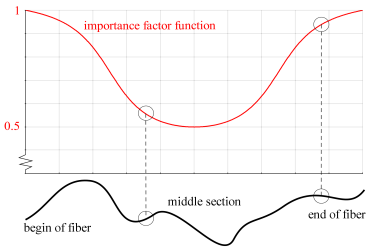

Figure 2. *Importance factor* function and its mapping to fiber position with $w_e = 1$ $w_m = 0.5$, $d = 0.5$ and $\sigma = 10$

The matrix A is normalized by multiplying it with $1/\max_{i,j\in\{1\ldots n\}} a_{ij}$ and setting its diagonal entries to 1. Note that the multiplication of the weights allows for a better description of the fiber similarity than a simple addition. Only if using the multiplication, it can be distinguished between balanced and unbalanced cells with respect to the number of fiber points from two fiber i and j.

Considering the limited spatial fiber density, one can assume that a single cell does not contain more than a fixed number of fibers in the average case. As a consequence, the time to calculate the fiber affinity resulting from a single cell is independent of the total number of fibers, thus in $O(1)$. As the number of cells which contain any data points is proportional to the number of fibers, the overall calculation of the affinity matrix is in $O(n)$.

A crucial detail of this fiber affinity measure is the way in which fiber points are assigned to grid cells:

1. **Hard Division** Each point is assigned to the cell containing it with a constant weight. While being very fast, this method has the disadvantage that two close points of fiber i and j lying on opposite sides of cell boundaries do not contribute to a_{ij} of the affinity matrix.

2. **Soft Division** Each point is additionally assigned to the surrounding cells. The corresponding weights are reduced by a factor which takes the distance from the central cell and the number of cells at the same distance into account. Thus, each shell of cells around the center cell gets the same weight. This allows for a more reliable clustering of fiber bundles with diameters exceeding the grid parameter d, as even two fibers which never share neighboring cells can now both have an affinity to a third fiber in between. For time efficiency purposes, each data point can be assigned randomly to only *one* cell if a sufficient point density of the fibers is given.

In large bundles of fibers, distinguishing features are not evenly distributed over each fiber, but can rather be found at their starting and ending sections. Thus, we introduce an *importance factor* to each fiber point that modifies the weight to be assigned to a cell by a function of the point's position on the fiber. A function which has proven very useful in our experiments because of its good adaptivity through some natural parameters is shown in Figure 2. The four parameters are w_e (weight of the end points), w_m (weight of the middle points), d (width of the middle section) and σ (sharpness of the transition between end weights and middle weights). The position x of the fiber point is taken to range from 0 to 1.

3. SPECTRAL FIBER CLUSTERING

As spectral clustering algorithms are more suitable for identifying non-convex clusters than conventional clustering algorithms like k-means, they are rapidly becoming popular in fields like image segmentation or speech recognition. The primary input of any spectral clustering algorithm is an affinity matrix which is usually normalized to account for different intracluster connectivity (denoted as $\hat{A}$). The eigenvectors of $\hat{A}$ are used to find a semi-optimal cut for partitioning a graph with edges weighted by the corresponding $\hat{A}_{ij}$. Most spectral clustering algorithms exploit only one eigenvector (e.g., the second-largest*) in order to cut the graph into two subgraphs,

* "largest" eigenvector denotes the eigenvector corresponding to the largest eigenvalue.

1: Compute affinity matrix A using the FG method.
2: Normalize A according to $\hat{A} = D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$ with $D = diag(\sum_{j=1}^{d} A_{ij})$
3: Calculate eigenvalues of $\hat{A}$. Sort and index (starting with one).
4: Form two groups: G_1 contains the two largest eigenvalues, G_2 the smaller eigenvalues down to a reasonable minimum.
5: **while** $|G_2| \geq 2$ **do**
6: Linearly fit eigenvalues of the two groups to their indices (separately)
7: Store the total error of both fits with index of $(\min_{x \in G_1}(x))$
8: Remove largest eigenvalue from G_2 and add it to G_1
9: Search for the smallest total error. Its index is the number of clusters.

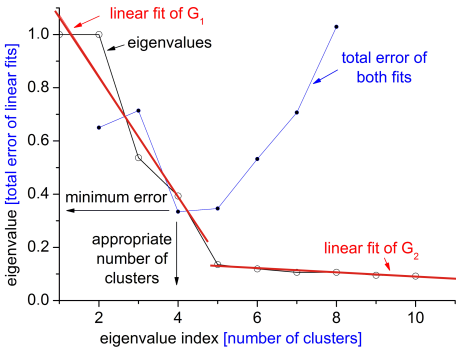

Figure 3. The algorithm detects the number of clusters automatically using our *eigenvalue regression*. The red straight lines in the right plot indicate the optimal linear fits for an example clustering. Moreover, the total error of both fits is shown.

and are performed recursively to yield k clusters.[12] However, it has been shown that a better clustering result can be obtained by using more eigenvectors and by cutting the graph directly into k clusters.[26]

In this section, we propose our novel *eigenvalue regression* to determine the number of clusters automatically. This approach is combined with our *multiple eigenvector clustering* (MEC), an adaption and optimization of[24] for our purpose of fiber clustering. It works accurately with the affinity matrix from the FG as an input.

For comparing our new approach, we also implemented the *elongated clustering*[25] (EC), a general spectral clustering algorithm that also determines the number of clusters automatically. We have optimized this algorithm for our fiber clustering problem. To the best of our knowledge, this is the only spectral clustering algorithm which uses multiple eigenvectors *and* allows for an automatic determination of the number of clusters. This method is described at the end of this section.

3.1. Multiple Eigenvector Clustering using Eigenvalue Regression

3.1.1. Eigenvalue Regression

From graph theory it is well-known that clusters in graphs correspond to eigenvectors of the affinity matrix with large eigenvalues. Depending on the intercluster connectivity, the normalized affinity matrix has more or less eigenvalues close to 1. The relevant eigenvalues for nearly separated clusters are bordered by an *eigengap*, the largest difference between two consecutive eigenvalues in a list of descending order.
Thus, for sparse graphs, the number of clusters can easily be found by searching for the largest gap between two eigenvalues. However, in practice, intercluster connectivity is not negligible and thus there is often no unambiguous eigengap corresponding to a comprehensible number of clusters. If a rigid eigengap method is used to determine the number of clusters, only the large and separated clusters can be found (by any clustering algorithm), therefore, inner clusters within coarse structures cannot be identified.
Our eigenvalue regression works as an eigengap detector for sparse graphs but can also detect a general change of the gradient of the eigenvalues. We plot the eigenvalues against their indices in the list of descending order and split this plot at such an index that the resulting halves can be linearly fitted with smallest error. Figure 3 illustrate our idea based on optimal linear fits.

3.1.2. Multiple Eigenvector Clustering

The algorithm proposed by Ng et al.,[24] which takes the affinity matrix A as input, constitutes the basis of our clustering algorithm. It has to be normalized to account for different intracluster connectivity. We denote the normalized affinity matrix as $\hat{A}$. In contrast to Ng et al.,[24] we perform a complete singular value decomposition (translation of the Algol procedure SVD[27]) on $\hat{A}$ in order to get all eigenvectors and eigenvalues. If it seems reasonable to assume a maximum number k_{max} of clusters for a given fiber set, an algorithm like the

1: Execute steps 1, 2 and 3 of the eigenvalue regression (Figure 3).
2: Estimate the number of clusters k by eigenvalue regression.
3: Assemble the k "largest" eigenvectors as columns of a matrix X and normalize its rows to have unit length $\rightarrow \hat{X}$.
4: Treat each row of $\hat{X}$ as a point in $\mathbb{R}^k$ and cluster the points into k clusters (for this purpose we use complete linkage).
5: **for all** i **do**
6: Assign fiber i to the cluster of row i of $\hat{X}$.

Algorithm 1: Multiple eigenvector clustering using eigenvalue regression

Arnoldi/Lanczos iteration can be used to calculate only the k_{max} largest eigenvalues and corresponding eigenvectors so that computation time can be reduced. The final number k of clusters is determined by our eigenvalue regression as described above. Afterwards, a matrix $\hat{X}$ is built in two steps. The k largest eigenvectors are assembled as columns of a matrix and its rows are normalized afterwards. Finally, the rows of $\hat{X}$ are treated as k-dimensional vectors. Performing a clustering algorithm on these vectors yields the desired result. We achieved best results using a hierarchical complete-linkage clustering algorithm. The indices of rows of $\hat{X}$ which are put into one cluster correspond to the indices of fibers in A which belong to the same cluster. The reasoning behind this algorithm as well as the advantage compared to a direct application of a clustering algorithm on the affinity matrix is described in.[24] We can summarize our optimizations for the purpose of fiber clustering (Algorithm 1) as follows:

- The affinity matrix A is provided by our FG method.
- The number of clusters is determined automatically by eigenvalue regression.
- Clustering of the rows of $\hat{X}$ is performed using complete-linkage[28] which is faster than k-means.

3.2. Elongated Clustering

Sanguinetti et al.[25] showed that a matrix consisting of the eigenvectors of the normalized affinity matrix implicitly contains the key information for an automatic detection of the number of clusters. The rows of this matrix are clustered using the Elongated k-means algorithm which accounts for the elongated structure of the clusters. Elongated k-means requires the a priori knowledge of two parameters: λ and ϵ which strongly influence the number of clusters.[25] We optimized the algorithm in order to compare its abilities to MEC with eigenvalue regression (Algorithm 2):

- The affinity matrix is calculated using the FG method.
- At each call of Elongated k-means it is initialized by preclustering with a complete linkage algorithm.
- In Elongated k-means: check for returning results to avoid periodicity.

4. RESULTS

All algorithms and data structures have been implemented in C++. As our fiber grid leads to the same clustering results in a considerable range of values for the grid parameter d, this parameter was always determined automatically by setting it to an experimentally determined fraction of the length of the fiber set's bounding box (we set it to a 1/15 of the smallest side of the bounding box). We have tested our approach with artificial and real fiber sets. For computing the real data, DTI images from a glioma patient were acquired on a 1.5T Siemens Sonata (image resolution $1.875 \times 1.875 \times 1.9$ mm^3, 60 slices, 6 gradient directions) and were filtered to an isotropic voxel size of 1.0mm^3 using a cubic B-spline filter. From that data, fiber sets were computed by

1: Execute steps 1, 2 of eigenvalue regression (Figure 3). Calculate eigenvectors of $\hat{A}$.
2: Set the temporary number of clusters, q, to $q = 2$.
3: Arrange the q largest eigenvectors as columns of a matrix X_q.
4: Set appropriate (empirically) values for sharpness (λ) and origin proximity (ϵ) for Elongated k-Means' distance function.
5: **repeat**
6: Cluster the rows of X using hierarchical clustering.
7: Determine the row with the largest euclidian norm in each cluster j.
8: Initialize the centers of the q clusters with the corresponding rows from step 7.
9: Add the $q + 1$ eigenvector to X_q getting X_{q+1} and initialize the $q + 1$th center in the origin.
10: Perform Elongated k-means with $q + 1$ centers on X_{q+1}.
11: **until** No row of X_{q+1} has been assigned to the $q + 1$th cluster ($q + 1$th cluster empty)
12: **for all** i **do**
13: Assign the fiber with index i to the cluster with index j iff row i of $\hat{X}$ has been assigned to cluster j.

Algorithm 2: Elongated clustering (EC) is an iterative spectral clustering algorithm with a built-in detection of the number of clusters using Elongated k-means

our deflection-based fiber tracking algorithm.[29] For generating the artificial fiber sets shown in Figure 4, we developed a tool that allows for interactively drawing fibers on arbitrary 3D surfaces.

Figure 6 and 7 show the corpus callosum clustered automatically by our MEC. Figure 5 shows a fiber set including the fasciculus arcuatus, a small bundle that is difficult to recognize, and that connects two important language regions, Wernicke's and Broca's area. The FG seems to provide the most suitable information for MEC to perform fully automatic clustering. Figure 5(iv) and 5(vi) show a major disadvantage of EC, namely the high sensitivity to one of its algorithmic parameters, ϵ. Small changes of ϵ often lead to extremely different clustering results, or to no clustering at all. In contrast to MEC, the EC can produce clustering results of equal quality for the FG and the FS, but the detected number of clusters is comparable to the number found by a rigid eigengap determination, thus, identifying only coarse structures. The artificial fiber set in Figure 4 was constructed to test whether spatially separated structures are recognized as different clusters and to what extent the different approaches are able to find finer substructures. As Figure 4(iii) and 4(iv) show, EC with FG or with FS is only able to detect coarse structures, where the FG leads to slightly better results compared to the FS. The eigenvalue regression determines the conceived number of seven clusters correctly from fiber grid data (Figure 4(i)), but cannot handle the affinity matrix produced by the FS method. Even though for the purpose of comparison the number of clusters was manually set to seven, MEC is not capable of forming the intended clusters from FS data (Figure 4(ii)). Only MEC combined with the FG delivers the desired clustering (Figure 4(i)).

Note that we have also implemented several other (non-spectral) clustering algorithms like hierarchical clustering, partitioning clustering, and self-organizing maps to work with our novel fiber grid. In all cases, better results are achieved than using a FS. However, all these algorithms are inferior to spectral clustering with respect to robustness, efficiency, and automation.

5. CONCLUSIONS AND FUTURE WORK

Feeding similarity information obtained by our fiber grid to clustering algorithms that are easy-to-use provides extremely robust and fast clustering of white matter fiber tracts. The time-consuming manual tuning of FS parameters is obsolete so that the a-priori knowledge required for achieving plausible clustering results is reduced drastically. Our novel eigenvalue regression is capable of determining a reasonable number of clusters as it

Figure 4. Artificial fiber set consisting of seven fiber bundles. (i) The clusters have only been correctly determined by multiple eigenvector clustering (MEC) based on our fiber grid. (ii) Not all clusters are correctly determined by MEC using a feature space. (iii) Only coarse clusters are found by elongated clustering (EC) using our novel fiber grid. (iv) EC using a feature space.

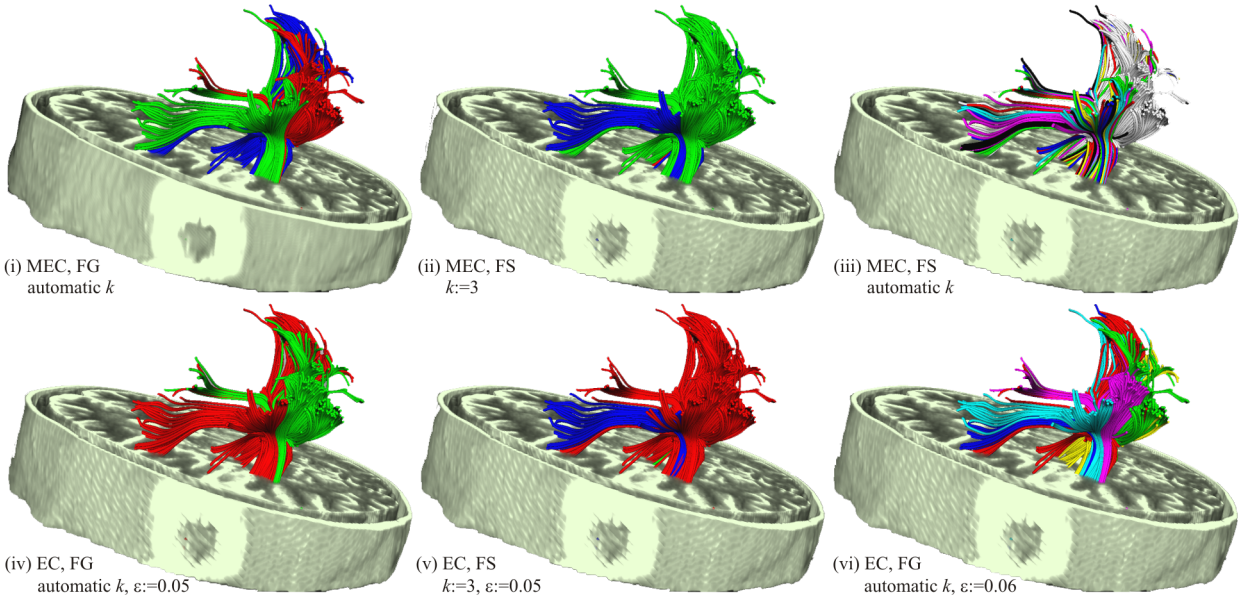

Figure 5. Wernicke's and Broca's area. Most reasonable results (i and vi) are achieved by our fiber grid (FG). Multiple eigenvector clustering (MEC, top row) is much more robust than elongated clustering (EC, bottom row), which is extremely sensitive to changes in algorithmic parameters (iv and vi). If using a feature space (FS), the number k of clusters was incorrectly determined (iii) and, thus, manually set to k:=3 (ii and v).

accounts for inter-cluster connectivity. Thus, inner clusters within coarse structures can be identified using our extended multiple eigenvector clustering. Compared to elongated clustering which also includes an automatic detection of the number of clusters, the MEC leads to more accurate and robust results.

We believe that this work opens up a number of avenues for future research. The computation time of our clustering algorithms could be further reduced by more sophisticated calculations, e.g., by restricting the number of eigenvalues and vectors that have to be computed. Furthermore, it should be examined if differently shaped arrangements of cells with respect to the FG allow for reducing the number of cells so that the affinity matrix can be determined more quickly.

The automatic determination of the number of clusters allows for several new applications. In the clinical context, e.g., where multimodal imaging is used, the integration of automatic fiber clustering into fMRI data visualization could be improved because anatomical structures which belong together can be well-separated from other surrounding fibers by our method. Thus, time consuming manual filtering steps could be avoided.

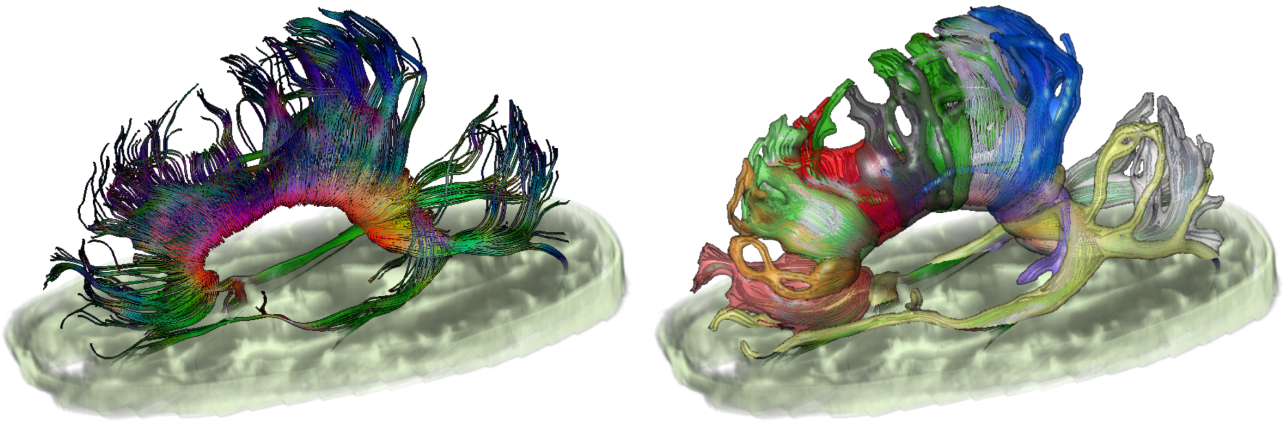

Figure 6. Fibers of the corpus callosum are automatically clustered by our spectral fiber clustering (MEC). We use transparent isosurfaces (with an adjustable distance to the fibers) for visualizing the clustering results.

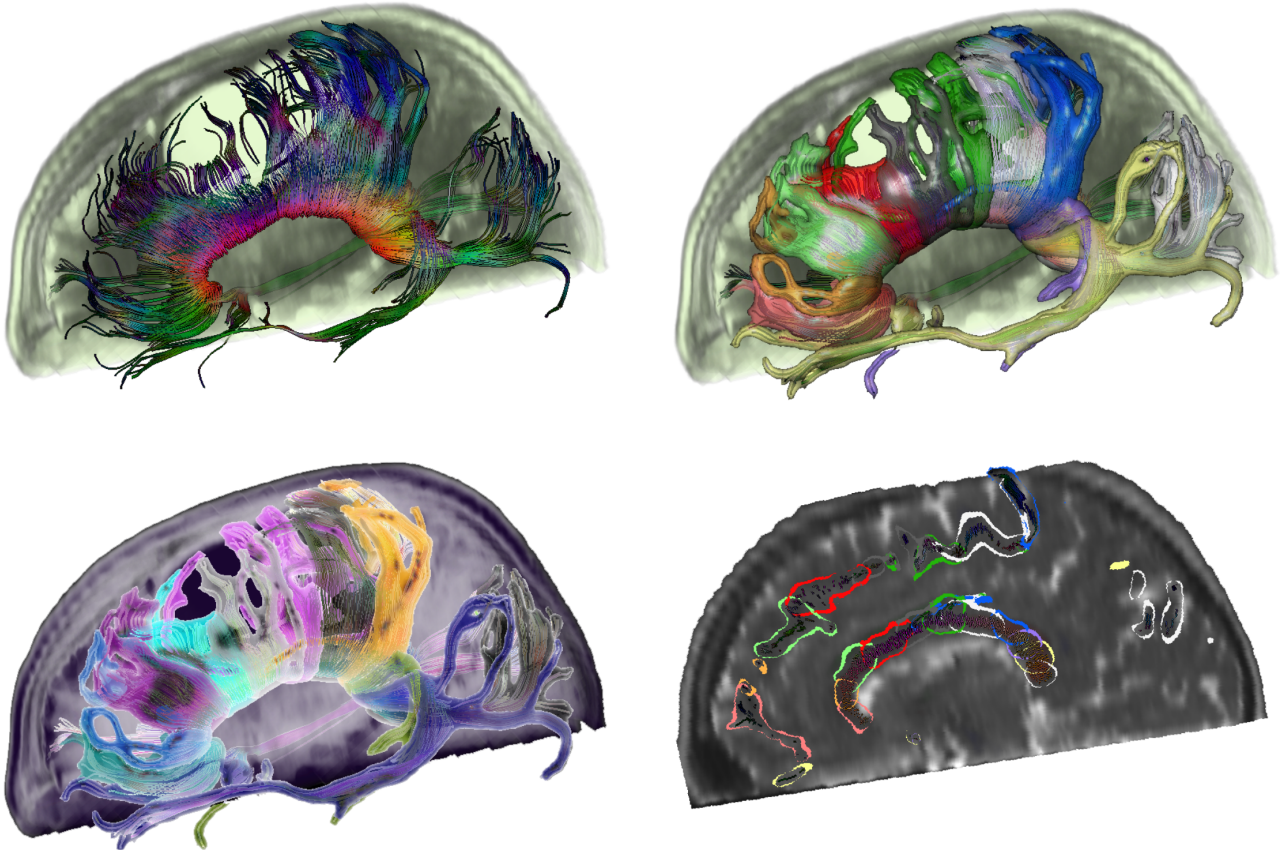

Figure 7. Upper row: sagittal view of clustering results, bottom left: if using inverse colors, boundaries between neighboring clusters are more emphasized, bottom right: clustering results projected onto anatomical image allow for easy correspondence to 3D.

REFERENCES

1. K. Yamada, O. Kizu, S. Mori, H. Ito, H. Nakamura, S. Yuen, T. Kubota, O. Tanaka, W. Akada, H. Sasajima, K. Mineura, and T. Nishimura, "Brain fiber tracking with clinically feasible diffusion-tensor MR imaging: Initial experience," *Radiology* **227**(1), pp. 295–301, 2003.
2. M. Kinoshita, K. Yamada, N. Hashimoto, A. Kato, S. Izumoto, T. Baba, M. Maruno, T. Nishimura, and T. Yoshimine, "Fiber-tracking does not accurately estimate size of fiber bundle in pathological condition: initial neurosurgical experience using neuronavigation and subcortical white matter stimulation," *NeuroImage* **25**(2), pp. 424–429, 2005.
3. P. Basser, "Fiber-tractography via diffusion tensor MRI (DT-MRI)," in *Proceedings of the 6th Annual Meeting of ISMRM*, p. 1226, 1998.
4. S. Mori, B. Crain, V. Chacko, and P. van Zijl, "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Ann Neurol.* **45**(2), pp. 265–269, 1999.
5. J. Klein, H. Hahn, J. Rexilius, P. Erhard, M. Althaus, D. Leibfritz, and H.-O. Peitgen, "Efficient visualization of fiber tracking uncertainty based on complex gaussian noise," in *Proc. 14th ISMRM Scientific Meeting & Exhibition (ISMRM 2006)*, p. 2753, 2006.
6. H. K. Hahn, J. Klein, C. Nimsky, J. Rexilius, and H.-O. Peitgen, "Uncertainty in diffusion tensor based fibre tracking," *Acta Neurochirurgica Supplementum* **98**, pp. 33–41, 2006.
7. D. Jones, A. Travis, G. Eden, C. Pierpaoli, and P. Basser, "PASTA: Pointwise assessment of streamline tractography attributes," *Magn. Reson. Med.* **53**, pp. 1462–1467, 2005.
8. S. Zhang, C. Curry, D. Morris, and D. Laidlaw, "Streamtubes and streamsurfaces for visualizing diffusion tensor MRI volume images," in *IEEE Visualization Work in Progress*, 2000.
9. D. Merhof, M. Sonntag, F. Enders, C. Nimsky, and G. Greiner, "Hybrid visualization for white matter tracts using triangle strips and point sprites," *IEEE Transactions on Visualization and Computer Graphics* **12**(5), pp. 1181–1188, 2006.
10. J. Klein, F. Ritter, H. Hahn, J. Rexilius, and H.-O. Peitgen, "Brain structure visualization using spectral fiber clustering," in *SIGGRAPH 2006, Research Poster, ISBN 1-59593-366-2*, 2006.
11. F. Enders, N. Sauber, D. Merhof, P. Hastreiter, C. Nimsky, and M. Stamminger, "Visualization of white matter tracts with wrapped streamlines," in *IEEE Visualization*, pp. 51–58, 2005.
12. A. Brun, H. Knutsson, H. J. Park, M. E. Shenton, and C.-F. Westin, "Clustering fiber tracts using normalized cuts," in *MICCAI'04*, pp. 368–375, 2004.
13. J. Shimony, A. Snyder, N. Lori, and T. Conturo, "Automated fuzzy clustering of neuronal pathways in diffusion tensor tracking," in *Soc. Mag. Reson. Med*, 2002.
14. Z. Ding, J. C. Gore, and A. W. Anderson, "Classification and quantification of neuronal fiber pathways using diffusion tensor MRI," *Magn. Reson. Med.* **49**, pp. 716–721, 2003.
15. L. O'Donnell, K. M, M. E. Shenton, M. Dreusicke, W. E. L. Grimson, and C.-F. Westin, "A method for clustering white matter fiber tracts," *AJNR* **27**(5), pp. 1032–1036, 2006.
16. L. Jonasson, P. Hagmann, J.-P. Thiran, and V. J. Wedeen, "Fiber tracts of high angular resolution diffusion mri are easily segmented with spectral clustering," in *Proceeding of ISMRM*, p. 1310, 2005.
17. V. E. Kouby, Y. Cointepas, C. Poupon, D. Rivière, N. Golestani, J.-B. Poline, D. L. Bihan, and J.-F. Mangin, "MR diffusion-based inference of a fiber bundle model from a population of subjects," in *MICCAI'05*, pp. 196–204, 2005.
18. M. Maddah, A. Mewes, S. Haker, W. E. L. Grimson, and S. Warfield, "Automated atlas-based clustering of white matter fiber tracts from DTMRI," in *MICCAI'05*, pp. 188–195, 2005.
19. A. Brun, H.-J. Park, H. Knutsson, and C.-F. Westin, "Coloring of DT-MRI fiber traces using laplacian eigenmaps," in *EUROCAST'03*, pp. 564–572, 2003.
20. B. Moberts, A. Vilanova, and J. van Wijk, "Evaluation of fiber clustering methods for diffusion tensor imaging," in *IEEE Visualization*, pp. 65–72, 2005.
21. K. Yeung, D. Haynor, and W. Ruzzo, "Validating clustering for gene expression data," *Bioinformatics* **17**(4), pp. 309–318, 2001.
22. M. Pauly and M. Gross, "Spectral processing of point-sampled geometry," in *SIGGRAPH'01*, pp. 379–386, 2001.

23. M. Pauly, M. Gross, and L. P. Kobbelt, "Efficient simplification of point-sampled surfaces," in *IEEE Visualization*, pp. 163–170, 2002.
24. A. Ng, I. Jordan, and R. Weiss, "On spectral clustering: Analysis and an algorithm," *Advances in Neural Information Processing Systems* **14**, pp. 849–856, 2002.
25. G. Sanguinetti, J. Laidler, and N. Lawrence, "Automatic determination of the number of clusters using spectral algorithms," in *MLSP 2005*, pp. 55–60, 2005.
26. C. Alpert, A. Kahng, and S. Yao, "Spectral partitioning: The more eigenvectors, the better," *Discrete Applied Math* **90**, pp. 3–26, 1999.
27. Golub and Reinsch, "Algol procedure SVD," *Num. Math.* **14**, pp. 403–420, 1970.
28. D. Krznaric and C. Levcopoulos, "The first subquadratic algorithm for complete linkage clustering," in *ISAAC '95*, pp. 392–401, 1995.
29. M. Schlueter, O. Konrad, H. K. Hahn, B. Stieltjes, J. Rexilius, and H.-O. Peitgen, "White matter lesion phantom for diffusion tensor data and its application to the assessment of fiber tracking," *Medical Imaging: Image Processing* **5746**, pp. 835–844, 2005.