

A slicing-based coherence measure for clusters of DTI integral curves

Abstract. We present a slicing-based coherence measure for clusters of DTI integral curves. For a given cluster, we probe samples from the cluster by “slicing” it with a plane at regularly spaced locations parameterized by curve arclengths. Then we compute a stability measure based on the spatial relations between the projections of the curve points in individual slices and their change across the slices. We demonstrate its use in refining agglomerative hierarchical clustering results of DTI curves that correspond to neural pathways. Expert evaluation shows that refinement based on our measure can lead to improvement of clustering that is not directly recovered using standard methods.

1 Introduction

Diffusion-Tensor Magnetic Resonance Imaging (DTI) measures the rate of self-diffusion due to the Brownian motion of water molecules in tissues [1]. Integral curves showing paths of fastest diffusion are among the most common information derived from DTI volumes, enabling the exploration of fibrous structures such as brain white matter and muscles non-invasively *in-vivo*. They are generated by tracking the principal eigenvector of the underlying diffusion tensor field in both directions [2] and often visualized as streamlines or variations of streamlines (streamtubes and hyperstreamlines) in 3D [3]. However, these 3D models are generally visually dense making it difficult for experts to ascertain anatomical and functional structures clearly. Therefore, there is a considerable interest in developing effective clustering methods.

In this context, we introduce a measure of *coherence* of a (hypothesized) cluster of curves. The cluster coherence measure we propose relies on evaluating the *stability* of further subdivision of the cluster. To this end, we use the configurations at what we call “slices”—cross-sections of the cluster. Each slice is effectively an embedding of the curves into points in two-dimensional space. These points can be clustered using any off-the-shelf clustering algorithm (we have experimented with Gaussian mixture clustering and spectral clustering). Each slice therefore provides a *vote* for each pair of curves being together or separate in the overall clustering. Furthermore, assuming reasonable smoothness of the curves, we can assess the temporal coherence of these votes: two adjacent slices carry more weight voting the same way than if their votes are opposite.

We demonstrate our measure’s use to improve an agglomerative hierarchical clustering algorithm that has been shown to be working relatively well in clustering integral curves corresponding neurofibers [4]. When our slice-based method detects that a stable split exists in the cluster, it provides a specific partitioning, that can be used as part of a clustering algorithm. Expert evaluation shows

that this mechanism may be superior to the standard hierarchical clustering approach. While our primary motivation in designing the slicing-based coherence measure is the task of refining an initial clustering assignment, it can be used for validating clusterings, quantifying connectivity or parameterizing clusters.

In the next section we discuss related work briefly, focusing on clustering methods. We then describe the details of the slice-based coherence measure in Section 3. Its application in improving hierarchical clustering results is described in Section 4, followed by discussion and conclusion in Section 5.

2 Related work

There have been several clustering methods proposed for DTI curves. All of them are adaptations of some of the well-known clustering methods including fuzzy c-means (a variation of k-means) [5], agglomerative hierarchical clustering [6,7], and spectral clustering [8,9]. An evaluation of the most popular fiber clustering algorithms can be found in [4]. While we are not aware of any coherence measure specific to DTI integral curves (or 3D curves of any origin, for that matter), we found stability (or confidence) argument for cluster analysis in statistics to be common [10,11,12] and often used to compare and validate clustering methods. In this context, our notion of stability can be seen as an “internal” geometric index primarily aiming to quantify the quality of hypothesized individual clusters.

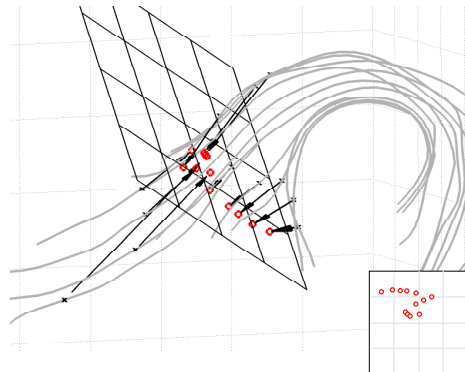


Fig. 1. Illustration: the cluster is sliced at arclength 0.2. Crosses (black): points on the curve corresponding to the arclength. Circles (red): projections on the slicing plane. The bottom-right legend shows an embedding of the curves in the axis-aligned plane in according to the current slice.

3 Slicing and cluster stability

The basic idea of our method is to assess the behavior of curves, that comprise a candidate cluster, relative to each other in a number of cross-sections (slices). A clustering pattern, consistent over neighboring slices, is found, the cluster is considered incoherent, and split, with the splitting details deduced from the slices. Otherwise, the cluster is considered coherent, and preserved intact.

3.1 Slicing a candidate cluster

Given a cluster of undirected curves X_1, \dots, X_N in \mathbb{R}^d , where $d = 3$, we first orient the curves (in the sense of assigning start- and end-point designations), in a way that makes orientations consistent within the cluster. This can be done, for instance, by computing the start-to-end vector for each curve and iteratively re-orienting the curves until all the vectors are in the same half-space.

Next we sample each curve uniformly along its patch, in the following manner. Let S_i be the arc-length of the curve X_i ; i.e., the curve can be parameterized as $\mathbf{X}(s)$ where s goes from 0 to S_i . We take M samples $x_i^{(1)}, \dots, x_i^{(M)}$ from each curve such that $x_i^{(m)} = X_i(\alpha S_i)$ where $\alpha = \frac{m}{M}$. Now for each m , we have a point set $X^{(m)} = \{x_1^{(m)}, \dots, x_N^{(m)}\}$.

Each point set is then projected onto a $(d - 1)$ -dimensional *slicing plane*. Intuitively, we would like this plane to be normal to the “cluster tangent” at a given arclength. We estimate this by computing the tangent $\tau_i^{(m)}$ to the i -th curve at $x_i^{(m)}$ and averaging this direction over the curves, to yield $\tau^{(m)}$. The slicing plane is spanned by the columns of the $d \times (d - 1)^{(m)}$ matrix \mathbf{P}_m that are set to be orthonormal and orthogonal to $\tau^{(m)}$. Furthermore, we require the plane to pass through the mean location $\mu^{(m)} = \frac{1}{N} \sum_{i=1}^N x_i^{(m)}$. The geometry of this construction is illustrated in Figure 1.

This slicing mechanism is motivated by the following intuition. Suppose that the cluster is coherent, i.e. the curves comprising it follow similar paths. In that case, the slicing plane will be close to normal to the tangent of each individual curve, and moreover the points corresponding to the m -th sample will be close in space. On the other hand, if the cluster contains a number of distinct subgroups, there will be two sets of samples, which will be clustered around distinct submeans. Moreover, in such a case we can expect that projecting those samples onto the slicing plane, orthogonal to the cluster, will emphasize the separation between the subgroups.

Given the m -th slicing plane, we represent the cluster by a set of $(d - 1)$ -dimensional projections, which we call the “slice”, $Y^{(m)} = \{\mathbf{y}_1^{(m)}, \dots, \mathbf{y}_N^{(m)}\}$:

$$\mathbf{y}_i^{(m)} = \mathbf{P}_m^T (\mathbf{x}_i^{(m)} - \mu^{(m)}) \quad (1)$$

Before we proceed, we note that an alternative slicing technique could be based on fitting a plane to $X^{(m)}$ using Principal Component Analysis [13]. A straightforward application of this in our experiments has proved inferior to the tangent-based technique described above, due to the effect of outliers. This could be alleviated by using *robust* PCA; we do not, however, pursue it further in this paper.

3.2 Cluster co-membership within slices

We now treat each slice separately. Effectively, each slice is an *embedding* of the set of curves into $(d - 1)$ -dimensional space. The set of points in this space can in principle be clustered using any off-the-shelf clustering method. However, we are specifically interested in determining whether there is an “interesting”

partition in the slice. Therefore, of most interest to us are methods that allow automatic determination of the number of clusters. In the experiments present below, we used for this purpose the Gaussian mixture clustering, accompanied by the Bayesian Information Criterion (BIC) for setting the number of components [14]. The details on these techniques can be found, e.g., in [15].

3.3 Spatial and temporal coherence

Once a Gaussian mixture model has been fit to the slice $Y^{(m)}$ for a range of values of k , we select the optimal model based on BIC. With the mixture model, we cluster the data by assigning each point to the component with the highest responsibility (i.e., the highest posterior probability of the point drawn by the associated Gaussian distribution). We will denote the label of $\mathbf{y}_i^{(m)}$ by $c_i^{(m)}$.

There is of course no direct relationship between the cluster labels across slices, since those are assigned arbitrarily. Even if the same partition of the curves to two clusters is reached in two slices, a given set of curves could be labeled 1 in one slice and 2 in the other slice. The information of interest to us is conveyed by the *co-membership* of a given *pair* of curves in a particular slice. Specifically, we define $J_{ij}^{(m)}$ such that

$$J_{ij}^{(m)} = \begin{cases} 1 & \text{if } c_i^{(m)} = c_j^{(m)}, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Intuitively, the m -th slice “votes” for X_i and X_j being in the same cluster if $J_{ij}^{(m)} = 1$. One could now simply combine the values of this vote across all slices. However, it is possible for a particular pair of curves to be separated by the clustering in a given slice m simply due to the randomized nature of the procedure (placement of slicing plane, random initialization of the EM, etc.) despite genuinely belonging to the same subgroup. If indeed such an accidental result occurs, we expect that it will not persist in the neighboring slices, $m - 1$ and $m + 1$. This suggests the notion of *temporal coherence*. We formalize it with the value $T_{ij}^{(m)}$ defined as

$$T_{ij}^{(m)} = \begin{cases} 1 & \text{if } J_{ij}^{(m)} = J_{ij}^{(m-1)} \text{ and } J_{ij}^{(m)} = J_{ij}^{(m+1)}, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

Finally, we can now represent the “vote” of the m -th slice regarding the similarity of curves i and j , weighted by the temporal coherence of that slice (with respect to that pair!):

$$W_{ij}^{(m)} = (1 - J_{ij}^{(m)}) \cdot T_{ij}^{(m)}. \quad (4)$$

$W_{ij}^{(m)}$ takes the value of 0 or 1. When it is 1, it indicates that the m -th slice supports a split where X_i and X_j are separated. $W_{ij}^{(m)}$ with 0 value, on the other hand, does not mean that the slice supports keeping X_i and X_j in the same cluster. It simply means that there is no evidence to the contrary. This may be

due to the two curves being separated in the slice (zero co-membership), or the lack of temporal coherence, i.e., the co-membership of i and j being unstable in this slice—or both. This sort of asymmetric reasoning is similar in spirit to the statistical hypothesis testing formalism, in which the null hypothesis is either rejected or not, but never “accepted”. In our case, this reflects the notion that not splitting the cluster is the default action.

We are now ready to describe the algorithm for evaluating the coherence of the cluster. Given the set of curves, we calculate a set of M slices, and cluster each of them using the Gaussian mixture clustering, with BIC model selection. We then compute the cluster co-membership values $J_{ij}^{(m)}$ and the temporal coherence $T_{ij}^{(m)}$ for each $m = 1, \dots, M$ and $i, j = 1, \dots, N$. This yields for each pair i, j a set of M votes $W_{ij}^{(m)}$. We combine the evidence regarding X_i and X_j across all slices in a single value:

$$W_{ij} = M - \sum_{m=1}^M W_{ij}^{(m)}. \quad (5)$$

This provides us with a measure of similarity for each pair of curves, organized in the form of an $N \times N$ matrix \mathbf{W} . By default, its diagonal is set to zero. For example, a value of 0 indicates that the two curves are consistently separated in *all* slices; the value of M indicates that no evidence for separating the two curves is provided by any of the slices. The form of the similarity matrix is the same as that calculated with some alternative distance measure (see the next section for an example). There is however an important distinction: our similarity values are computed *in the context of the given cluster*, specifically with the objective to evaluate potential splits. This is in contrast to “global” distance measures, that operate on the same scale throughout the data set.

4 Refining hierarchical clustering results

Given initial “hypothesized” (candidate) clusters as a result of hierarchical clustering, we evaluate the decision of splitting the cluster by applying the spectral clustering algorithm described in [16] to \mathbf{W} . Briefly, the algorithm is based on eigendecomposition of the symmetric matrix $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$, where \mathbf{D} contains the sums of corresponding rows of \mathbf{W} on its main diagonal and zeros everywhere else. When two clusters are requested, the algorithm divides the data according to the sign of the corresponding entries of the $N \times 1$ second largest eigenvector of the matrix above. Note that this can lead to either two clusters, or a single one (if all the entries in the eigenvector are of the same sign). The results of this final clustering are returned to the user.

4.1 Expert Evaluation

In order to assess the utility of our method for a practitioner, we conducted a comparative evaluation study with a domain expert. The fiber tracking data used in our experiments were obtained from DTI brain data sets scanned from four volunteers. We first obtained initial candidate clusters by applying the single linkage hierarchical clustering algorithm, with a distance measure adapted from [3]. We used the cut-off threshold of 3 (set heuristically to produce reasonable cluster sizes). The expert has significant experience with DTI and uses

fiber-track models (integral curves) generated from DTI data sets in clinical research regularly. We compared three methods. One is our slicing-based method described above; we used $M = 50$ slices for each candidate cluster. The second method consists of applying the single-linkage hierarchical clustering algorithm on an affinity sub-matrix for the candidate cluster, with the objective of obtaining two sub-clusters. In other words, this method finds the optimal split of the appropriate subtree in the original dendrogram. The third method applies the same spectral clustering algorithm as does our method, but the affinity matrix is computed from the curve distances D_{ij} rather than using slices.

We displayed 93 cases to the expert. In each case where at least one of the methods produced a different split. The expert was shown the results from the three methods side-by-side, and asked to rank them. The evaluation was blind (i.e., the expert was not told which of the methods produced each result). We used a streamtube representation for the curves and juxtaposed clusters with the surface of lateral ventricles (areas of cerebrospinal fluid in the brain) extracted from the original MRI data set to provide an anatomical landmark. The expert were able to interactively manipulate the viewpoint, zoom-in and out, and rotate the models. The ranking decisions were based on the following criteria: anatomical correctness (whenever the expert recognized a candidate cluster), anatomical and physiological plausibility, and amount of information conveyed by the resulting clustering. The goal in this scenario is to evaluate whether the resulting clustering decision helps to identify biologically distinct structures in the DTI data.

4.2 Results

Out of 93 cases considered, 20 were ranked as three-way ties (i.e., undecided, equally good, or equally arbitrary). In the remaining 73, the expert ranked our slicing-based method the best method in 49 cases while only in 9 of these cases there was another method ranked the same. Furthermore, our method ranked worse than the other two methods only in one case. We summarize these results in Figure 2. In Figure 3, we show examples where our method produced better

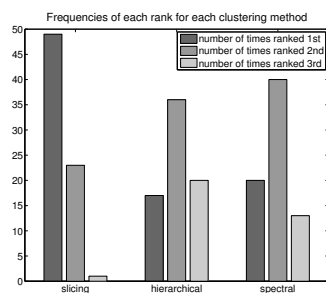


Fig. 2. Results of expert evaluation of cluster refinement in terms of method ranking.

results (according to the expert feedback) than the other two methods.

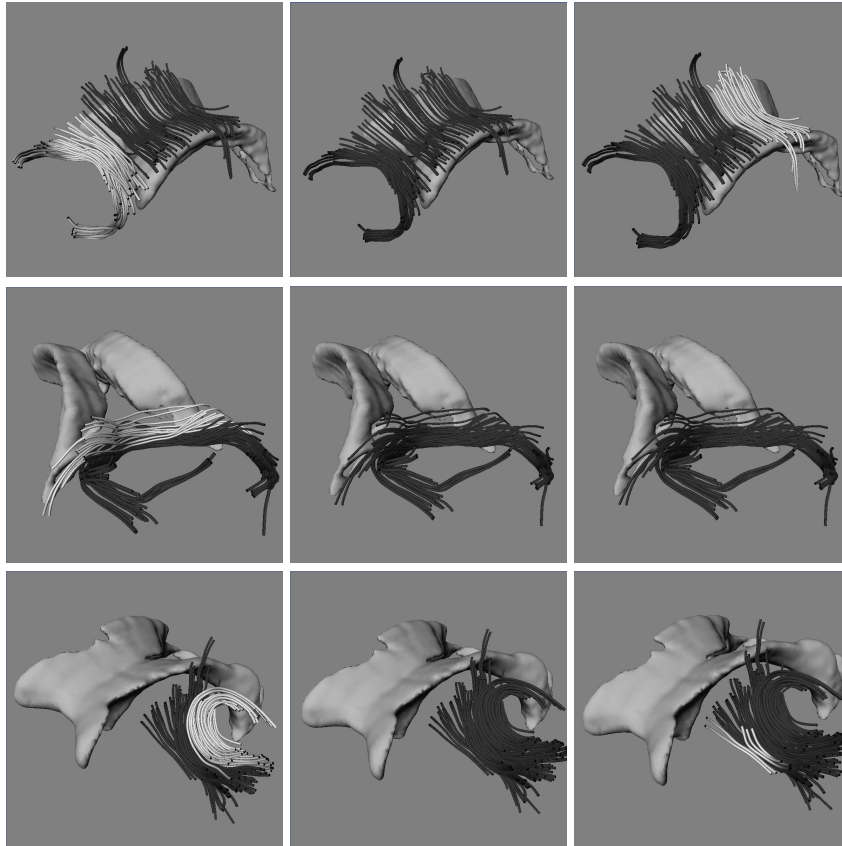


Fig. 3. Examples of clusters where our method was ranked to be the best by the expert. Split, if exists, is visualized by two colors (dark gray and white). Left: the slicing-based split. Middle: 2-way split on the dendrogram used in hierarchical clustering. Right: spectral clustering.

5 Discussion and conclusions

The main contribution of this paper is the novel coherence measure that is obtained by combining an intuitive geometric idea, slicing, with known statistical machine learning techniques. An important property of our method is its reliance on the context of the cluster in evaluating similarity between curves. The quantities computed in each slice, and across slices, are tied directly to the clustering objective. This is in contrast to the more standard setup in which distances and thresholds are defined in the global context. While our motivation and the experimental evaluation have been on 3D DTI integral curves, our method generalizes to higher and lower dimensions easily and may apply to curve data in other domains. There is a number of technical aspects that we believe could be improved. Specifically, a more robust slicing and projection method that explicitly down-weights outliers could help reduce uncertainty in the per-slice quantities.

A more challenging extension of the idea of slicing that we would like to pursue is to build a semi-parametric generative model for a cluster. Also, although our primary interest is in the analysis of DTI data, we believe our method is general and can be applied to any domain where data instances are represented by curves or trajectories.

References

1. Basser, P.J., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the nmr spin echo. *Journal of Magnetic Resonance B* (1994) 247–254
2. Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine* **44** (2000) 625–632
3. Zhang, S., Demiralp, C., Laidlaw, D.H.: Visualizing diffusion tensor MR images using streamtubes and streamsurfaces. *IEEE TVCG* **9**(4) (2003) 454–462
4. Moberts, B., Vilanova, A., van Wijk, J.J.: Evaluation of fiber clustering methods for diffusion tensor imaging. In: *Procs. of Vis'05*. (2005) 65–72
5. Shimony, J., Snyder, A., Lori, N., Conturo, T.: Automated fuzzy clustering of neuronal pathways in diffusion tensor tracking. In: *Proceedings of ISMRM*. (2002)
6. Ding, Z., Gore, J.C., Anderson, A.W.: Case study: reconstruction, visualization and quantification of neuronal fiber pathways. In: *VIS '01: Proceedings of the conference on Visualization '01*, IEEE Computer Society (2001) 453–456
7. Corouge, I., Gouttard, S., Gerig, G.: Towards a shape model of white matter fiber bundles using diffusion tensor mri. In: *International Symposium on Biomedical Imaging*. (2004) 344–347
8. Brun, A., Knutsson, H., Park, H.J., Shenton, M., Westin, C.F.: Clustering fiber traces using normalized cuts. In: *MICCAI*. (2004) 368–375
9. O'Donnell, L., M, K., Shenton, M.E., Dreusicke, M., Grimson, W.E.L., Westin, C.F.: A method for clustering white matter fiber tracts. *AJNR* **27**(5) (2006)
10. Breckenridge, J.N.: Replicating cluster analysis: Method, consistency, and validity. *Multivariate Behavioral Research* **24**(2) (1989) 147–161
11. Fridlyand, J., Dudoit, S.: Applications of resampling methods to estimate the number of clusters and to improve the accuracy of a clustering method. Technical report, Division of Biostatistics, School of Public Health, UC Berkeley (2001)
12. Lange, T., Roth, V., Braun, M.L., Buhmann, J.M.: Stability-based validation of clustering solutions. *Neural Computation* (2004) 1299–1323
13. Jolliffe, I.T.: *Principal Component Analysis*. 2nd edn. Springer (2002)
14. Schwarz, G.: Estimating the dimension of a model. *The Annals of Statistics* **6**(2) 461–464
15. Bishop, C.M.: *Pattern Recognition and Machine Learning*. Springer (2006)
16. Ng, A.Y., Jordan, M.I., Weiss, Y.: On spectral clustering: Analysis and an algorithm. In: *Neural Information Processing Systems*, MIT Press (2001) 849–856