

Detecting Clusters in fMRI Data Using the Dendrogram Sharpening Algorithm

L. Stanberry¹, R. R. Nandy¹, D. Cordes¹

¹University of Washington, Seattle, WA, United States

Synopsis

Dendrogram sharpening (DSh) makes it easier to identify the modality regions, which are, in essence, areas of activation in the human brain while analyzing fMRI data. The idea of the algorithm is to simply remove data from low density regions in order to obtain a clear representation of the data structure. For this paper, a hierarchical clustering method based on a single linkage algorithm was used to group the data, then the DSh algorithm was applied twice. After cluster cores were identified, the classification algorithm was run on voxels, set aside during DSh, attempting to assign them to the found clusters.

Introduction

In this work, the problem of identifying functionally connected regions of the brain, at rest or during sustained motor activity, is considered. The data collected is often contaminated by physiological noise, as well as by motion artifacts from scanner instabilities and subject motion. Upon grouping data into a hierarchical cluster tree, clusters could be identified either by specifying their number or by choosing the appropriate inconsistency coefficient. The number of clusters present in the data is not known before hand, and even slight variation of the inconsistency coefficient significantly effects the results. Sharpening methods remove observations in low density regions deepening the valley between clusters, thus making it easier to identify modal peaks. DSh in particular discards all small-sized children-nodes with a large-sized parent node. Then each large enough node is tested for inconsistency with respect to its left and right child using different thresholds, depending on the agglomeration value of the nodes constituting the child. This approach does not require prior knowledge of the number of clusters or their location. The final classification, performed on the voxels left aside during sharpening, attempts to assign the rest of the data to the identified clusters if possible.

Theory

First, the hierarchical clustering algorithm, based on the single linkage method, is used to group the data. As a measure of similarity between the voxels, a distance, based on the correlation coefficients of the EPI time series, was specified, as described in [1]. The result of linking is a binary dendrogram tree. The sharpening process described in [2] is controlled by two parameters, *fluff*-value and *core*-value, where *fluff*-value is the maximum size of a child cluster that will be discarded if it has a parent node of a size not smaller than a *core*-value. This is a recursive algorithm that begins with only the root node of the whole dendrogram and continues to invoke on the children of each node since each child node is the root node of a dendrogram tree itself. Initially the *fluff*-value equals 2 and during the second sharpening, applied to the dendrogram resulting from the first computation, the *fluff*-value is set to 10. Upon completion of sharpening, we begin cluster identification using the method of inconsistent edges described in [2]. We exploit the correspondence between the single linkage dendrogram (SLD) and minimal spanning tree (MST). The length of the edge in the MST is an agglomeration value of the corresponding node in the SLD. The length of the edge (or the agglomeration value of the corresponding node) is compared to the length of the edges represented by the left child and then by the right child. The value of median edge length of the left (right) subtree plus twice the interhinge spread is the proposed threshold, beyond which edge is considered inconsistent with respect to its left (right) child. The recursive procedure begins with analyzing the root node of the SLD and then invokes on its children. After the clusters were identified we could classify the voxels, set aside during DSh. We are using the original SLD obtained prior the sharpening procedures moving from the bottom of the tree to its top.

Methods

Data was collected on a 1.5 T MR scanner (GE, Waukesha) equipped with echo-speed gradients and a standard birdcage head coil. Scan parameters were set as 64x64 imaging matrix, TR/TE 400ms/40ms, FA 50 deg, FOV 24cm x 24cm, BW ± 62.5 kHz, slice thickness 7mm, gap 2mm, 2275 time points, 4 axial slices covering motor/somatosensory cortex. Two different paradigms were used. First one consisted of 5 minutes continuous finger-tapping exercise with 5 minutes of rest before and after the motor task activity. Second paradigm consisted of 5 minutes equally paced finger-tapping exercise with 5 minutes of rest before and after the motor activity. Acquired data was grouped into a hierarchical tree using the single linkage method. The sharpening algorithm was performed on the original data with parameters: *fluff*-value equals 2, *core*-value equals 40. Second sharpening was applied to the dendrogram, resulted from the first computation, with the *fluff*-value set to 10.

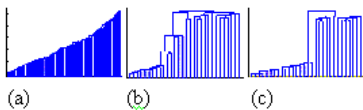


Fig.1 SLD (a) original, (b) sharpened once, (c) sharpened twice

Results

It is clear from Fig.1 that the sharpened dendrogram better reveals the structure of the data. Clusters identified after the DSh contained from 8-40 voxels. All clusters were obtained using frequency contributions between 0.02Hz and 0.1Hz. Most of the major clusters gave clearly identifiable patterns. Fig.2 displays the results of the sharpening. Activation was found in the motor cortex area during motor task activities as well as during the resting state.

Conclusion

The dendrogram sharpening technique is a very helpful tool for analysis of activation patterns in fMRI data. The single linkage method, used to group the data, is preferable over the complete or average linkage algorithms, because of its ability to reveal the modal regions of the data. The large amount of fMRI data makes it difficult to detect the functional connectivity between the voxels. Sharpening significantly simplifies the identification by reducing the data set and preserving its structure. Though DSh is sensitive to the parameters it is possible to empirically identify appropriate values for the particular data set, that would yield a meaningful result.

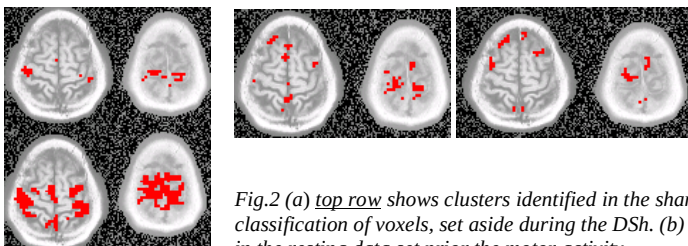


Fig.2 (a) top row shows clusters identified in the sharpened data during paced motor activity, bottom row displays same clusters after final classification of voxels, set aside during the DSh. (b) clusters identified in the resting data set after the motor activity, (c) clusters identified in the resting data set prior the motor activity.

References

- [1] Cordes D. et al. Magn Reson Imag 2002;20:305-317 [2] McKinney S. Autopaint: A toolkit for visual. data in four or more dims, PhD thesis, UW,1996.