Is Fuzziness useful in fMRI Clustering?

H. Fischer, J. Hennig
Dept. of Radiology, Dept. of Computer Science*
University of Freiburg

Introduction

Different clustering techniques have been proposed for analyzing fMRI time series [1-3]. Among these are the K-Means and the Fuzzy C-Means (FCM) algorithm. FCM is based on the fuzzy set theory and adds another parameter m which is said to handle 'fuzzy' data more appropriately. In the fMRI clustering literature this value is typically set to m=1.3 [2] and its actual value is believed not to be critical. This is investigated in this study.

Methods

The basic idea of partitional clustering is to group n objects into k sets in a way that similar objects (pixel time series, ptc) belong to the same set (cluster). This assignment is called partition. Assessing all possible partitions in respect to some quality measures leads to a statistical explosion. Therefore, optimization procedures have been developed, representing clusters by typical objects (cluster centers) that are updated iteratively, starting from some initial setting. Using K-Means, the cluster centers are updated as the average of all belonging ptcs (1), while FCM updates cluster centers as an average of all ptcs (2) weighted by the ptcs' membership values u_{ij}.

$$\vec{c}_i = \frac{1}{N_i} \sum_{\vec{x} \in C_i} \vec{x} \quad for \quad 1 \le i \le k \qquad (1)$$

$$\vec{c}_i = \frac{\sum_{j=1}^{n} (u_{ij})^m \vec{x}_j}{\sum_{j=1}^{n} (u_{ij})^m} \quad for \quad 1 \le i \le k \,. \qquad (2)$$

In order to assess the influence of m, FCM clustering results for different m values are calculated. Then it is asked if these partitions assign the ptcs into similar clusters. This comparison is carried out using the Jaccard statistics [4]

$$S_{Jaccard}(A,B) = \frac{n_{11}}{n_{11} + n_{12} + n_{21}}$$

with n_{11} as the number of object pairs (x_d, x_e) which belong to the same cluster in both partitions A and B and n_{12} as the number of object pairs which belong to the same cluster in A but not in B (n_{21} equivalent). Partitions are similar for $S_{Jaccard} \to 1$ and different for $S_{Jaccard} \to 0$.

Results

K-Means and FCM clustering was applied to a large number of fMRI data sets. A typical example is given in the following. The data was analyzed after normalizing the ptcs by their mean and standard deviation. The Euclidean distance was used to calculate the 8 cluster centers. The first partition - using random initialization - was obtained with m=1.3 as proposed in [2]. Using these cluster centers as initialization, partitions for m=1.05, 1.1, 1.2, 1.4, 1.5, 1.6, 1.7, 1.8 and 2.0 had been processed and the statistics for all possible pairs of partitions had been calculated. The histogram of the Jaccard statistic is shown in Fig 1a and reveals that different m values result in a very different assignment of ptcs to clusters.

The poor reproducibility of the FCM in respect to m is further demonstrated by comparing 10 randomly initialized partitions which serve as a reference. The partitions had been calculated with m=1.3 and even though prototype based clustering is known to be highly sensitive to the initialization, FCM is even much more sensitive to the variation of m, as can been seen in Fig. 1. This sensitivity is further demonstrated in Fig. 2: the cluster with functional activation is plotted for m=1.05, 1.1, 1.2, 1.3 and 2.0 along with the results for K-Means clustering and the average time profile of all t-test significant pixels as a reference.

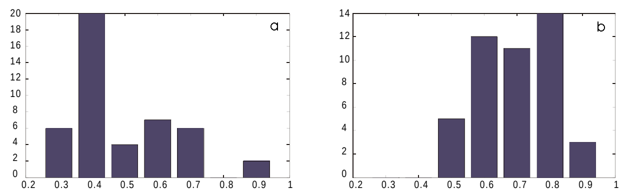

Fig. 1: (left) Histogram of Jaccard statistics for varying m values and (right) for random initialized partitions.

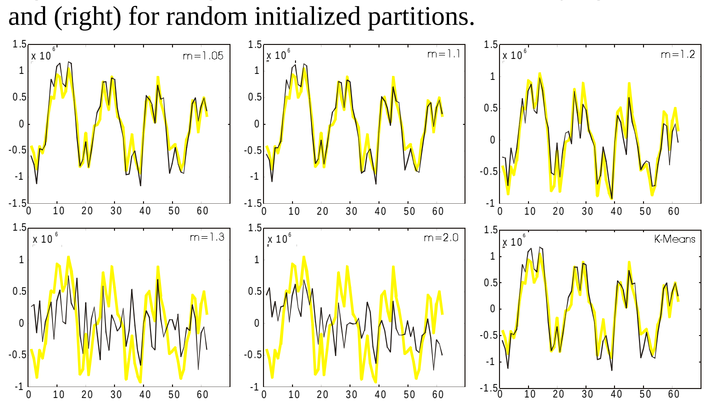

Fig. 2: Cluster center with functional activation for different m values overlayed on top of the average time profile of t-test significant pixels (light grey). For large m values, the information about functional activation is lost. Results for small m values are almost identical to the results obtained by K-Means.

Discussion

The results consistently show that

- different values of FCMs' fuzziness controlled by the parameter m result in significantly different assignments of ptcs to clusters,
- the influence of m is considerably stronger than the influence of random initialization,
- best partitions are obtained for $m \to 1$, while large m values do not produce any useful results,
- for $m \to 1$ FCM approaches the K-Means algorithm.

Conclusion

The FCM obtains best results for very small values of the fuzziness controlled by m. In this case, the algorithm approximates the classical K-Means algorithm which can be seen from (2) and Fig. 2. K-Means is computationally much faster and easier to implement. Therefore, our results suggest that K-Means or some other more advanced data driven approach should be used for paradigm free analysis.

References

1. Ding X. et al., 2nd ISMRM Meeting, p. 630, 1994.
2. Scarth G. et al., 3rd ISMRM Meeting, p. 238, 1995.
3. Wismüller A. et al., 6th ISMRM Meeting, p. 249, 1998.
4. Dubes R., Jain A.K., (1988). Algorithms for Clustering Data, Prentice-Hall, Englewood Cliffs, New Jersey.