Clustering of Brain Tumors by Segmentation of Magnetic Resonance Data

P. W. Krooshof[1], G. J. Postma[1], W. J. Melssen[1], L. M. Buydens[1]

[1]Analytical Chemistry, Radboud University Nijmegen, Nijmegen, Gld, Netherlands

Introduction

Diagnosis of brain tumors is made by clinical symptoms, radiological appearance (MRI, CSI) and histopathological diagnosis [1]. Previous research [2] has shown that clustering of Magnetic Resonance data facilitates the diagnosis of brain tumors. However, the study did not include spatial information in the clustering of Magnetic Resonance Imaging (MRI) data. Additionally, the clustering results were not stable.

The aim of this study is to investigate the added value of the inclusion of spatial information in the clustering of multivariate MRI images to facilitate the interpretation of MR data.

Materials and Method

Multivariate images of a slice of the brain are studied, which consist of a stack of MRI images with a resolution of 256x256 pixels. MR imaging was performed on a 1.5 T Siemens Vision MR whole body system. In addition to T_1-, proton density and T_2-weighted images, a T_1-weighted Gadolinium contrast enhanced MRI was obtained (SE, TR=600, TE=14 ms). Patients with different tumor types and grades are included in a database setup by the ETUMOUR project [3].

A (standard) clustering procedure for large data sets (with high computational demands) has been used for the segmentation of the multivariate images [2]. This procedure is based on Mixture Modeling, which describes data as mixtures of multivariate normal distributions.

In the first step of the clustering procedure computational demands can be reduced by performing fuzzy c-means clustering [4] to obtain a reduced number of clusters (e.g. ~240), instead of using the ~65,000 individual pixels. In the second step, the initial parameters for Mixture Modeling are estimated by model-based hierarchical clustering, which hierarchically groups the 240 clusters up to one cluster. In the final step, Mixture Modeling is performed by using the initial parameters, to obtain e.g. 2 to 10 cluster models. The model with the optimal number of clusters fits the data best and this model is selected by using the Bayesian Information Criterion (BIC). This clustering procedure, however, is not stable (see below). Furthermore, no spatial information is used in these methodologies.

A recently developed methodology [4] uses spatial information to obtain the initial partitioning for the hierarchical clustering.

1. Obtaining the Initial Partitioning

The image is partitioned into a number of homogenous regions, e.g. about 240, by grouping adjacent pixels if their distance (in feature space) is smaller than a predefined threshold. Several small regions may not be selected by this method, as they are not homogenous and may contain noise, artefacts or spatially isolated pixels. Subsequently, the second step of the standard procedure (model-based hierarchical clustering) is applied, to obtain the initial parameters for Mixture Modeling. Finally, Mixture Modeling is performed on the resulting 2 to 10 cluster models and the best model is selected by using the BIC.

2. Classification of Small Regions

Markov Random Field (MRF) filtering is used to take into account the spatial relation between pixels, to result in a stable clustering. Additionally, if small regions are not selected by the clustering procedure, MRF is used to classify these regions.

Results and Discussion

The standard clustering procedure (i.e. without inclusion of spatial information) gives the results (for two runs) shown in Fig. 1. As shown, small differences exist in the two clustering results and are explained by the random selection of clusters by fuzzy c-means clustering.

With the MR Image partitioned into a number of clusters by searching for homogeneous regions (including spatial information), about 240 regions are found. These homogeneous regions are presented in Fig. 2. When the additional steps of the clustering procedure are performed on these regions, the clustering results of Fig. 3 are obtained. The results of several runs are identical, which indicates the stability of the method, and small noisy regions disappear in the clustering result.

Conclusions

Spatial information can be used effectively to obtain estimates for initial Mixture Modeling parameters. Because the use of spatial information results in a stable partitioning into homogeneous regions, a unique clustering is obtained when these regions are further clustered.

Future research should investigate the radiological value of the clustering strategy.

Acknowledgements

This work is funded by the EU: ETUMOUR project, FP6-2002-LIFESCIHEALTH 503094

References

1. M.C. Preul, *et al*, "Accurate, Noninvasive Diagnosis of Human Brain Tumors by using Magnetic Resonance Spectroscopy, *Nature Medicine*, Vol. 2 (1996), 323-325
2. R. Wehrens, *et al*, "Mixture Modelling of Medical Magnetic Resonance Data", *J. Chemometrics*, Vol. 16-6 (2002), 274-282
3. http://www.etumour.net/
4. T.N. Tran, *et al*, "Strategies for Mixture Model Clustering of Multivariate Images", *submitted*

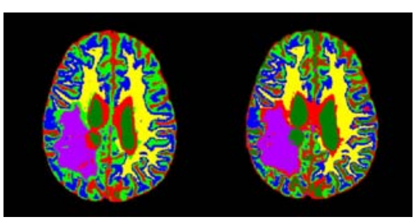

Fig. 1: Two clustering results with the initial partitioning obtained by fuzzy c-means.

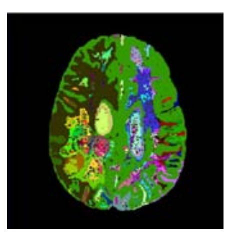

Fig. 2: Representation of the homogenous regions (different colors) in the multivariate MR image

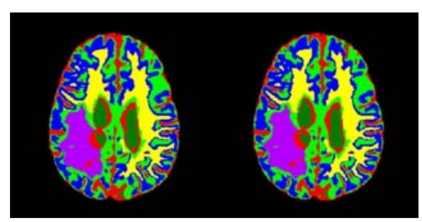

Fig. 3: The stable six-cluster model after MRF classification (two runs are shown).