

Segmentation of Chemical Shift Images with Mixture Modeling

A.W. Simonetti, R. Wehrens, L.M.C. Buydens

Laboratory of Analytical Chemistry, Nijmegen University, Toernooiveld 1, 6525ED
Nijmegen, The Netherlands
A.Simonetti@sci.kun.nl

Introduction

Chemical shift imaging (CSI), a method which samples ^1H NMR-spectra from a grid of volume elements, produces an overwhelming amount of data. Each spectrum contains information about several metabolites in the sampled area. One approach for interpretation of this large amount of data is segmentation of the CSI grid in clusters which share the same features, followed by classification of each segment to a specific tissue type. We used mixture modeling to perform segmentation of CSI images for automatic identification of malignant areas in the human brain.

Mixture modeling is an approach to clustering in which the data are described as mixtures of distributions, usually Gaussians [Mc Lachlan, G.J., *Finite Mixture Models*, Wiley NY (2000)]. The parameters of the Gaussians are obtained through application of the EM algorithm [Dempster, A.P., J. Royal Statist. Soc. B 39(1) (1977) 1-38].

There are a number of advantages of mixture modeling over other more common forms of unsupervised clustering; it is possible to derive uncertainty estimates for individual classifications; with the Bayesian Information Criterion (BIC) [Schwarz, G., Ann. Statist. 6 (1978) 461-464] the optimal number of clusters as well as the clustering method can be chosen automatically -this is often a problem in other unsupervised approaches- and the visualization of the clusters is possible in the space of the original variables.

Results & Discussion

High resolution MRI images and CSI images were acquired from four patients with a histologically proven oligodendroglioma tumor on a 1.5 T Siemens whole body system. From each spectrum within the selection box 5 resonances were analyzed by numerical integration of 16 points under each resonance: myo-inositol (3.6 ppm); choline (3.2 ppm); creatine (3.0 ppm); N-aspartyl-aspartate (2.0 ppm) and lipids/lactate (1.3 ppm). These 5 quantitated resonances were used as input for the mixture modeling.

Segmentation was performed with the MCLUST software, which gives easy access to mixture modeling for clustering. The number of clusters were fixed between 3 and 5, which is a reasonable number from a medical point of view. From the BIC value

calculated by MCLUST the optimal number of clusters as well as the best model was automatically selected for each patient.

In Figure 1 the BIC values for patient G are presented. An ellipsoidal model with equal cluster volume and shape (EEE) gives the highest BIC value for 3 clusters, therefore this 3 cluster model is used.

In Figure 2 the segmentation (left) and uncertainty (right) values for each patient are depicted below a MRI image. Clearly the segmentation corresponds to areas suspected for malignancy for all patients. In all cases the segments can also be identified by visual inspection of the spectra. Also the agreement between clustering pixels in MRI images (not shown) and the segmentation presented is very good. In the case of patient M the 3 voxel segment at the top left of the tumor is suspected to be a bleeding. Segmentation for patient W is not convincing. This can be due to a low quality of the spectra.

The uncertainty values for all voxels are quit low. Uncertainties below 10% are common (white). The values can be used to expel voxels from a cluster. For example the isolated and right lower voxel belonging to the dark cluster in patient D may be removed from the cluster due to their high uncertainty value.

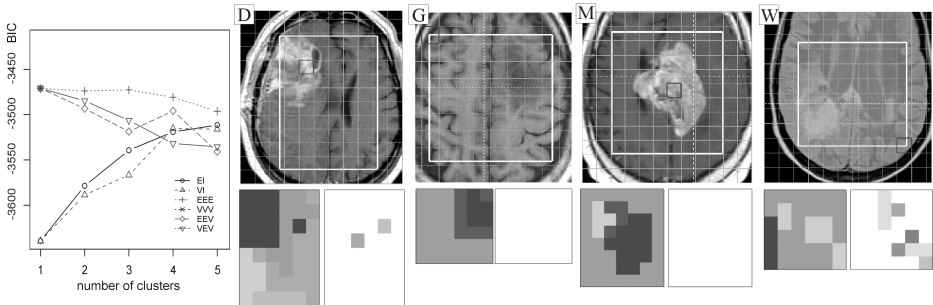


Fig. 1. BIC values for patient G. Clustering with 6 different models is tested.

Fig. 2. Segmentation (left) of CSI images from four patients. The number of clusters is automatically defined by MCLUST. For G and W 3 clusters are optimal, for D 5 and for M 4. The uncertainty values (right) for voxels are mostly below 10 % (white).

Conclusion

Mixture modeling can be used to identify malignant regions in the brain by segmenting CSI data. The characteristics of all spectra within one *segment* may be used instead of the spectrum in one *voxel*, which may lead to a more reliable identification of the tumor type by a (supervised) classification algorithm. If several clusters can be identified within the tumor, it may even be possible to take the heterogeneity of the tumor into account.

Voxels with a high uncertainty value can be identified and removed from a cluster, thus improving classification. CSI image segmentation may be performed automatically using the BIC criterion.