DATA FUSION OF HR-MAS AND IN-VIVO INFORMATION WITH APPLICATION IN BRAIN TUMOR RECOGNITION

A. Croitor Sava[1], T. Laudadio[1,2], J.B. Poullet[1], D. Monleon[3], M.C. Martinez-Bisbal[4,5], B. Celda[4,5] and S. Van Huffel[1]

[1] Department of Electrical Engineering, Division ESAT-SCD, Katholieke Universiteit Leuven, Leuven, Belgium ; [2] Istituto per le Applicazioni del Calcolo "M. Picone", National Research Council, (IAC-CNR), Bari, Italy; [3] Fundación de Investigación Hospital Clínico Universitario de Valencia, Spain; [4] Departamento de Química-Física, Facultad de Química, Universidad de València, València, Spain; [5] CIBER of Biongineering, Biomaterials and Nanomedicine, ISC-III, Spain

Anca.croitor@esat.kuleuven.be

Abstract: The purpose of this study is to combine multimodal information coming from magnetic resonance imaging (MRI), magnetic resonance spectroscopic imaging (MRSI) and high resolution magic angle spinning (HR-MAS) in order to develop advanced pattern recognition methods able to help clinicians in diagnosing brain tumors.
Recently, the use of multimodal information in diagnosis and tissue characterization of brain tumors has been shown to improve the performance of the classifiers. In particular, HR-MAS is seen as a promising complementary method, as it is very helpful for the assignment of well-resolved spectra of cellular metabolites and, therefore, can strongly support in vivo MRSI by increasing its clinical value. The pattern recognition technique used in this study is based on a statistical method known as canonical correlation analysis (CCA). Here, the performance of CCA is further investigated by introducing HR-MAS information as prior knowledge. Specifically, we study the impact of using HR-MAS information in combination with in vivo MRI and MRSI information to detect and classify brain tumors.

I. Introduction

MRI is an imaging technique mainly used to produce high quality images of the inside of the human body and is widely applied by clinicians for brain tumor diagnosis. It brings important information regarding the location and the anatomy of the tumor. However, questions regarding the type and the grade of the tumor are still difficult to address by using MRI.

On the other hand, MRSI provides significant biochemical information on the molecules of the organism under investigation. It can therefore complement conventional anatomical imaging. For the moment retrieving accurate estimates of the most relevant metabolite concentrations remains a difficult task because of magnetic field inhomogeneities, relatively low signal-to-noise ratio (SNR), physiological motion, peak overlap and low resolution of the spectra.

The use of HR-MAS data can support *in vivo* MRSI, as previous studies have already shown that the resolution of HR-MAS spectra is higher and, therefore, metabolites characterizing the different tissue types can be more easily distinguished. HR-MAS spectra are characterized by narrow line widths and large SNRs, thereby allowing the identification of an important number of metabolites [1].

In our study, we propose a classification method that uses multimodal information coming from 3 different measurement techniques: HR-MAS, MRI and MRSI. This method is based on CCA, an accurate and efficient tissue segmentation and classification technique that has recently been successfully applied to prostate and brain tumor recognition [2,3]. The performance of CCA is here further investigated by introducing HR-MAS information as prior knowledge.

II. Materials and Methods

Materials

The tissue subspace models used in this study were obtained by exploiting information coming from HR-MAS measurements on 100 biopsies stored after surgery at -80C until the HR-MAS study. The biopsies were gathered in 4 brain tumor classes: 27 glioblastomas (GBM), 18 glioma grade II, 6 glioma grade III, 49 meningiomas (MNG). 1D PRESAT (pulse-and-acquire) data were acquired at 0-4 ^{0}C in 11.7 T (500 MHz for 1H) and 14.1 T (600 MHz for 1H) BRUKER Analytik GmbH spectrometers at spinning rates of 4000 and 5000 Hz, respectively. The spectra were preprocessed: the water components were removed by HLSVD-PRO [4], signals were normalized (divided by the L2 norm of the frequency domain signal between 0.25 and 4.2 ppm), aligned with respect to the Alanine doublet at 1.47 ppm, and corrected for the baseline (by subtracting the product of the signal and an apodization function) [5].

The method was evaluated on a number of 14 presurgical MRI and MRSI studies coming from patients with brain tumor posteriorly diagnosed by

consensus on a histopathological study (2 glioma grade II, 5 glioma grade III, 4 GBM and 3 MNG).

The *in vivo* MR data were acquired with a 1.5 T Siemens Vision whole body scanner. For every patient first 4 MR images were acquired: T1 weighted (TE/TR=15/644ms), T2 weighted (TE/TR=16/3100ms), proton density weighted (TE/TR=98/3100ms) and a Gadolinium enhanced T1 image (15 ml 0.5 M Gd-DTPA). MRSI data were acquired, using a 2D STEAM pulse sequence, with the following parameters: 16x16x1024 samples, TR/TE/TM=2000 or 2500/20/30 ms, slice thickness = 12.5 or 15 mm, FOV = 200 mm, spectral width = 1000 Hz and NS=2. Eddy current correction was performed followed by water removal, baseline correction and subtraction of the residual from the original time domain signal as in the paper by Simonetti et al. [6]. All spectra were normalized with respect to the water signal.

Canonical Correlation Analysis

CCA represents the multi-channel generalization of ordinary correlation analysis, which quantifies the relation between two random variables x and y by means of the so-called correlation coefficient:

$$\rho = \frac{Cov[x, y]}{\sqrt{V[x]V[y]}}$$

CCA can be applied to multichannel signal processing as follows: consider two zero-mean multivariate random vectors

$x = [x_1(t), \ldots, x_m(t)]^T$ and

$y = [y_1(t), \ldots, y_n(t)]^T$ with $t = 1, \ldots, N$ where the superscript T denotes the transpose. The following linear combinations of the components in x and y are defined, which respectively represent two new scalar random variables X and Y:

$$X = \omega_{x_1} x_1 + \cdots + \omega_{x_m} x_m = \omega_x^T x$$

$$Y = \omega_{y_1} y_1 + \cdots + \omega_{y_n} y_n = \omega_y^T y$$

CCA computes the coefficients ω_x and ω_y that maximize the correlation between X and Y [2].

In the tissue segmentation approach proposed in [2,3], the aim is to detect those voxels whose spectra X best correlate with model tissue spectra Y, defined as prior knowledge. CCA is applied to each voxel and the tissue type giving rise to the largest canonical correlation coefficient is assigned to the voxel under investigation. The results are then exploited in order to construct nosologic images in which all the detected tissues are visualized.

Defining the subspace model

The purpose of our study is to analyze the behavior of a method that combines multimodal information. To this aim, a harmonization of the input space is required and this step is performed by a dimension reduction of the available data.

Metabolite selection: The input pattern for developing the tissue subspace model was considered either as a set of quantified values from HR-MAS spectra or as the combination of the quantified values from HR-MAS and imaging intensities from MRI measurements. Tests were performed for both cases. For extracting the relevant features and reducing the input space peak integration, a frequency domain quantification method based on the integration of the area under the peaks of interest, was used. Special attention was paid to the selection of the area to integrate. The following metabolites were considered [7] *L2* (lipids at ≈ 0.9ppm), *L1* (lipids at 1.3 ppm), *Lac* (3CH3-group), *Ala* (1CH3-group), *NAA* (2CH3-group), *Glx* (3CH2-group), *Cr* (N(CH3)-group), *Cho* (N(CH3)3-group), *Tau* (1CH2-group), *mI* (1CH-, 3CH-, 4CH- and 6CH-group) + *Gly* (2CH2-group).

Metabolite correlation: Selecting the same integration intervals for both HR-MAS and MRSI, when extracting the characteristic features, led to misclassification due to the difference in resolution and SNRs between the two types of spectra. Therefore, a direct comparison of *in vivo* short-echo time spectra with the *ex vivo* HR-MAS was performed for each metabolite (see Figure 1 and 2).

Short-echo time spectra show overlapping peaks and a relatively low resolution. A particular region might cover resonances of more than one metabolite. For example Lactate (Lac), Alanine (Ala) and lipids (Lip) overlap around 1.33 ppm (see Fig. 2).

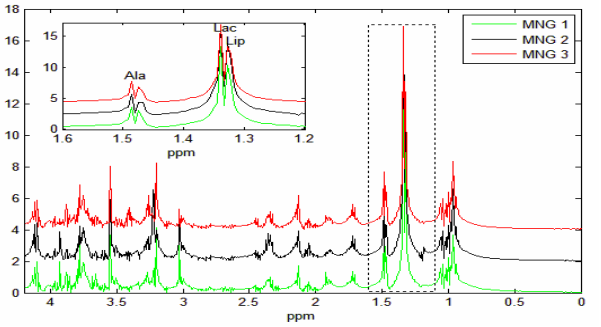

Figure 1 HR-MAS spectrum: Lip1 (1.31ppm) , Lac (1.33ppm), Ala doublet peak (1.47ppm)

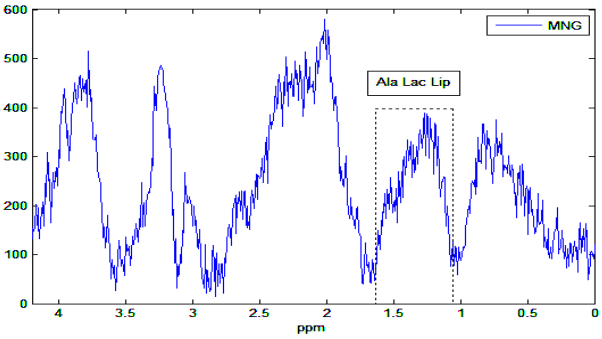

Figure 2 In vivo MRS spectrum: Lip1, Ala, Lac resonances overlap

The amplitude estimates of the short-echo time signals were obtained by applying peak integration within a spectral range of 0.13 ppm symmetrically chosen around the metabolite peak. For the overlapping peaks a combined integration was performed, around

resonances from metabolites and lipids selected as being characteristic of different tumor types [7].

For the HR-MAS spectra, as a clear distinction between the 10 metabolites can be made [8], an automatic method was used to select the area under the peak by computing the regions under the peak tales. The regions of interest were selected to match those typically found *in vivo*. For example, the typical Cho peak detected *in vivo* at 3.20 ppm includes many choline-containing compounds, which can be detected in HR-MAS spectra. In this work, all these compounds have been combined in a unique feature or spectral region. The regions of interest were: *L1* [0.89-0.92ppm], *L2* [1.30-1.32ppm] + *Lac* [1.32-1.34ppm] + *Ala* [1.45-1.49ppm], *NAA*[1.955-2.085ppm], *Glx* (Glu[2.09-2.17ppm] + Gln[2.39-2.50ppm]), *Cr*[3.01-3.03ppm], *Cho+Choline-containing compounds*[3.135-3.245ppm], *Tau*[3.39-3.42ppm], *mI* + *Gly*[3.72-3.8ppm], *Glx* + *Ala* [3.72-3.8ppm] and *Cr* [3.91-3.94ppm].

For the metabolites that were integrated as one feature for the *in vivo* MRS data, due to the overlapping of the resonances, but are clearly distinguishable in the HR-MAS data, each metabolite was integrated separately and the results were summed up in one feature.

Model variable components: The *y* variable in CCA, called subspace model, consists of a multivariate vector.

The first component (*y1*) was defined as the mean of the feature vectors extracted from the HR-MAS signals. The second component (*y2*) was defined as the first principal component of the matrix containing all the mean-centered HR-MAS feature vectors.

$$\begin{cases} y1(n) = \dfrac{1}{M}\sum_{i=1}^{M} S_i(\omega_n) \\ y2(n) = 1^{st}\,PC \end{cases}$$

We propose two approaches for defining the subspace models: using only HR-MAS feature vectors; using HR-MAS feature vectors as well as of four image variables obtained from MRI. The performance of the method was analyzed for both cases.

Applying CCA

During an MRSI acquisition, MR spectra are measured in a grid of voxels. In the proposed tissue segmentation approach, each voxel is assigned to a certain tissue type by computing the correlation between the tissue type models (*y* variable) defined a priori, and the voxel under investigation (*x* variable).

In order to exploit the spatial information the x variable is defined as a multivariate vector that contains information not only from the voxel under investigation but also from the surrounding voxels. Different spatial models can be used to define the x variable. Canonical correlation coefficients are computed with the chosen spatial model and the subspace models of different tissue types. The considered voxel is assigned to the tissue type described by the subspace model with the highest correlation coefficient. The final result is a nosologic image where the detected tissue types are visualized by different colors.

III. Results and discussion

The performance of CCA was analyzed by considering two different approaches: the feature vectors contain 10 HR-MAS metabolite estimates and the feature vectors contain 14 entries (10 HR-MAS metabolite estimates and 4 MRI variables)

A reliable and representative model has to be obtained for each tissue type in order to have reliable results. Therefore, we focused not only on the selection of the features but also on the data quality used in creating our models. Indeed, for most NMR experiments, pH and temperature might influence the homogeneity across peak positions and amplitude, as there is a strong dependency between a spins's resonance frequency and its local environment [9]. Hence, different tests were carried out, where the model vector was constructed, starting from data, by neglecting very different signals, eliminating quite different signals, keeping very similar signals. An increase in performance of the method was observed when the spectra used in creating the model were similar to each other and with low variation with respect to the standard model reported in [10].

After agreeing on the spectra to use to construct the model vectors, the attention was directed to the performance of CCA when using different spatial models: the symmetric 3x3 model, the symmetric 3x3 model without corner voxels and the 3x3 model.

A very clear conclusion could not be drawn regarding which spatial model performs best. With this respect, simulation studies need to be performed for detecting the best performing spatial model.

Figures 3 and 4 show the results obtained for a patient affected by GBM. We compare the two approaches: CCA method based on a 10 features model and on the 14 features model. In both cases the tumor area is correctly detected in the upper-right corner of our grid image.

The 10 features model approach detects in the tumor area GBM surrounded by voxels of grade II. The added value of MRI to our model vector does not bring significant new information in detecting the area of the tumor tissue, but influences the classification of the voxels in the tumor area, detecting a bigger area of GBM tumor surrounded by grade II and grade III voxels (see Figure 3 and 4). The spatial model used is the symmetric 3 x 3 model.

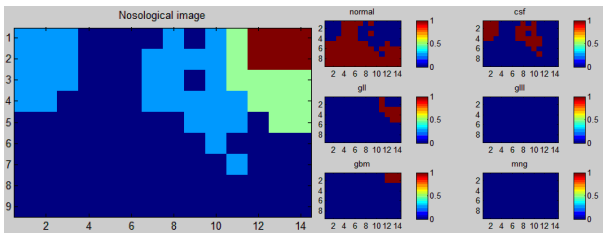

Figure 3 Nosologic image and corresponding tissue correlation maps obtained with CCA using the 10 feature vector approach

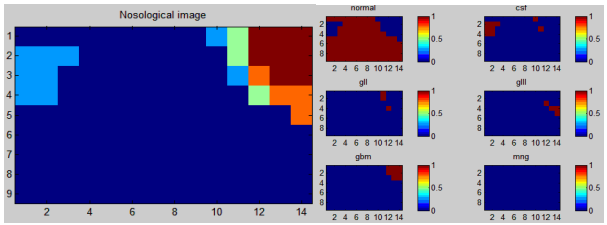

Figure 4 Nosologic image and corresponding tissue correlation maps obtained with CCA using the 14 feature vector approach

IV. Conclusions

In this study CCA was applied to segment and classify short echo-time brain MRSI data by using as prior knowledge information coming from HR-MAS or HR-MAS and MRI data. We studied the possibility of combining multimodal information; we investigated which parameters influence our classification results and the impact of HR-MAS information in combination with *in vivo* information for brain tumor recognition. The model that performs best is the model that uses multimodal information coming from HR-MAS and MRI data.

One of the most challenging problems for the method was the harmonization of all the input spaces due to the fact that we have to manage the use of very different information/data, obtained with different measurement techniques, as well as the use of data coming from different clinical centers. The problem was overcome by extracting common characteristic features from all the different data types.

V. Acknowledgment

Research supported by Research Council KUL: GOA-AMBioRICS, CoE EF/05/006 Optimization in Engineering (OPTEC), IDO 05/010 EEG-fMRI, IOF-KP06/11 FunCopt, several PhD/postdoc & fellow grants;Flemish Government:FWO: PhD/postdoc grants, projects, G.0407.02 (support vector machines), G.0360.05 (EEG, Epileptic), G.0519.06 (Noninvasive brain oxygenation), FWO-G.0321.06 (Tensors/Spectral Analysis), G.0302.07 (SVM), G.0341.07 (Data fusion), research communities (ICCoS, ANMMM);IWT: PhD Grants;Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, `Dynamical systems, control and optimization', 2007-2011);EU: BIOPATTERN (FP6-2002-IST 508803), ETUMOUR (FP6-2002-LIFESCIHEALTH 503094), Healthagents (IST–2004–27214), FAST (FP6-MC-RTN-035801), Neuromath (COST-BM0601), National Grants from the Science and Education Ministery (SAF2004-06297, SAF2004-20971-E and SAF2007-65473) and Health Ministery (ISCIII2003-G03/185) of the Spanish Government and by ADIRM. Generalitat Valenciana (GVA2007-AP64). DM gratefully acknowledges the Ministry of Science and Education for a Ramon y Cajal 2006 contract.

VI.References

1. M. Carmen Martınez-Bisbal, Luis Martı´-Bonmatı´, Jose´ Piquer, Antonio Revert, Pilar Ferrer, Jose´ L. Llacer, Martial Piotto, Olivier Assemat and Bernardo Celda , (2004) 1H and 13C HR-MAS spectroscopy of intact biopsy samples ex vivo and in vivo 1H MRS study of human high grade gliomas, NMR in Biomedicine, Pages 191–205

2. T. Laudadio, P. Pels, L. De Lathauwer, P. Van Hecke, S. Van Huffel, (2005) Tissue segmentation and classification of MRSI data using Canonical Correlation Analysis, Magn. Res. Med., 54, 1280-1284

3. De Vos M, Laudadio T, Simonetti AW et al (2007) Fast nosologic imaging of the brain. J Magn Reson 184:292-301

4. Laudadio T, Mastronardi N, Vanhamme L, Van Hecke P, Van Huffel S. , (2002) Improved Lanczos algorithms for blackbox MRS data quantitation. Journal of Magnetic Resonance.; 157:292-297.

5. J.B. Poullet, M.C. Martinez-Bisbal, D. Valverde, D. Monleon, B. Celda, Carles Arus and S. Van Huffel (2007), Quantification and classification of high-resolution magic angle spinning data for brain tumor diagnosis, Conf Proc IEEE Eng Med Biol Soc 1: 5407-5410.

6. A.W. Simonetti, W.J. Melssen, M. van der Graaf, A.Heerschap, L.M. Buydens, Automated correction of unwanted phase jumps in reference signals which corrupt MRSI spectra after Eddy current correction, J. Magn. Res., 159, 151-157 (2002)

7. Andy Devos, (May 2005), Quantification and classification of magnetic resonance spectroscopy data and applications to brain tumour recognition, Phd thesis, Faculty of Engineering, K.U.Leuven

8. V. Govindaraju, K.Young and A. A. Maudsley, Proton NMR chemical shifts and coupling constants for brain metabolites, NMR Biomed. 2000 May;13(3):129-53

9. G. Reynolds1y, M. Wilson, A. Peet and T. N. Arvanitis. (2006) An Algorithm for the Automated Quantitation of Metabolites in in Vitro NMR Signals, Magnetic Resonance in Medicine 56:1211–1219

10. Sarah J. Barton, Franklyn A. Howe, Andrew M. Tomlins, Simon A. Cudlip, Jeremy K. Nicholson, B. Anthony Bell, John R. Griffiths (1999), Comparison of in vivo 1H MRS of human brain tumours with 1H HR-MAS spectroscopy of intact biopsy samples in vitro, MAGMA; 8: 121–128