Non-negative blind source separation techniques for brain tumor tissue typing using *in vivo* MRSI data

Anca R. Croitor Sava[1,2], Diana M. Sima[1,2], Sofie Van Cauter[3,4], Uwe Himmelreich[4], Sabine Van Huffel[1,2]

[1] Department of Electrical Engineering, ESAT-SCD, KU Leuven, Kasteelpark Arenberg 10, box 2446, 3001, Heverlee, Belgium; [2] IBBT-K.U.Leuven Future Health Department, Kasteelpark Arenberg 10, box 2446, 3001, Heverlee, Belgium; [3] Department of Radiology, University Hospitals of Leuven, Herestraat 49, 3000, Leuven, Belgium; [4] Biomedical NMR Unit/Molecular Small Animal Imaging Center, Department of Medical Diagnostic Sciences, Katholieke Universiteit Leuven, Herestraat 49, 3000, Leuven, Belgium

1. Introduction

Although the magnetic resonance spectroscopic imaging (MRSI) technique is very interesting due to its non-invasive nature, as well as due to the possibility of exploring spatial information, the accuracy of MRSI in differentiating and grading glial brain tumors is limited by the significant variability of *in vivo* spectra due to the intra-tumoral heterogeneity. In gliomas one can observe distinct histopathological tissue properties, such as viable tumor cells, necrotic tissue or regions where the tumor infiltrates normal brain. Moreover, a significant variability within the magnetic resonance spectroscopic (MRS) spectral profiles belonging to the same brain tumor tissue type can be observed due to the heterogeneity that characterizes brain tumors. Thus, the tissue under investigation might present contributions from various tumor tissue types. The observed spectra are, therefore, a combination of different constituent sub-spectra, since the measured signal is the response to the stimulation of the entire tissue sample. The overall gain with which a tissue type contributes to a spectrum is proportional to its abundance, i.e., its proportion in the entire mixed tissue sample. This concept can be summarized by describing the spectra available from *m* samples, which are stacked as *n*-dimensional row vectors in an *mxn* matrix *X*, where *n* is the number of observations (or data points) in each spectrum:

$$X = AS \tag{1}$$

where *S* is a *kxn* matrix whose *k* rows are the unknown pure tissue spectra (also referred to as sources) and *n* the number of observations of each source. *A*, an *mxk* matrix, contains the concentrations, or abundances, of the constituent pure tissue sources in each sample. Brain tumor tissue classification problems arising from *in vivo* MRS can benefit from such a solution.

This study proposes a screening between the intratumoral histopathological tissue properties within MRSI data by quantifying the abundance within each MRSI voxel for each intratumoral histopathological tissue property. We assume that such an approach would greatly assist in an improved diagnosis, prognosis and treatment planning in gliomas. Non-negative matrix factorization with sparcity constraints (NNMFSC) (Kim and Park, 2007) will further be considered for this study, as it has been shown to be most accurate in solving the addressed problem. Additionally, nosologic images are drawn based on the extracted abundance maps, reflecting the presence of necrosis, viable tumor cells or infiltrations in the MRSI grid.

2. Data

MRSI data of 7 patients with gliomas, histopathologically confirmed according to WHO classification, were acquired at the University Hospital of Leuven (UZLeuven) on a 3T Philips scanner (Achieva and Intera, Philips, Best, The Netherlands), using a PRESS pulse sequence as the volume selection technique. The acquisition parameters are: FOV: 16x16cm, VOI: 8x8cm, samples size: 2048, number of signal averages: 1, shimming: pencil beam shimming, first and

second order. For each patient both water suppressed and unsuppressed proton MRSI data were acquired, as well as MR images (T_1-weighted MRI without contrast, T_2-weighted MRI, T_1-weighted MRI with contrast, FLAIR, DWI with sequence b-value of 0 mm/s² , DWI with sequence b-value of 1000 mm/s²).

The spectra were preprocessed using the Matlab platform and by water removal with HLSVD-PRO (Laudadio et al., 2002), baseline correction and normalization. The baseline was corrected using an apodization function and the normalization was performed with respect to the water signal. Magnitude spectra are computed by taking the absolute value of the Fourier transformed time-domain signals. Contributions outside the frequency interval [0.25, 4.2] ppm were filtered out in order to keep only the contribution of the metabolites of interest and to avoid expensive computations.

3. Methods

For each MRSI grid, NNMFSC is applied separately on the magnitude MRSI spectra in the region of interest between 0.25 ppm and 4.2 ppm and on sets of features obtained from the spectra.

A two step approach is implemented. This choice is motivated by the high heterogeneity of the data reflected both at MRSI grid level, as well as at the intra-tumoral level. Also, since there are different tumor types and grades, in the first step we separate normal brain tissue voxels from voxels with abnormal brain tissue, which should therefore contain predominant tumor tissue. For this step, the value of k is set to 2 and the two obtained NNMFSC components should typically represent the normal brain tissue and the abnormal tissue pattern. Then, based on the highest abundance, each voxel is assigned to normal or abnormal tissue. For selecting the abnormal tissue voxels one may use MR images. Still, as shown in previous studies, glial tumors are highly infiltrative and the area surrounding the enhancing region in glioblastoma and the region where the gliomas infiltrate, appeared normal on the MR images. Additionally, for the infiltrated non-enhancing/non-*T*2 hyperintense areas, abnormal Cho/NAA ratio levels were reported for gliomas (Di Costanzo et al., 2008). Elevated levels of Cho were also observed in the surrounding region for gliomas in (Fan et al., 2004).

The NNMFSC results are validated by comparing them against MRI intensity enhancement and by evaluating the Cho/NAA ratios within the grid. Spectra from areas with obvious intensity enhancements were selected as tumor spectra.

In the second step, NNMFSC is applied within the region with abnormal tissue with the purpose of identifying for each voxel within this region its predominant intratumoral histopathological property corresponding to necrotic, high cellular or infiltrations. The results are then used to construct nosologic images, by assigning each voxel to the tumor tissues type with the highest abundance coefficient. In this step the number of components to be extracted, k, is set to 3.

As mentioned above, two types of experiments were designed in this study. One where we work with the full magnitude spectra and one where we work with metabolic features extracted from each spectra. For the second case, AQSES-MRSI (Croitor Sava et al., 2011a), an advanced metabolite quantification method for MRSI data, with which the available spatial information is exploited, was considered.

For the metabolite features case, we considered a different number of metabolites for each step. Thus, 11 observations, representing the concentration of the most representative metabolites in separating tumor from normal tissue (NAA, Glu, Cr, PCho, Glc, Lac, Ala, Myo, Tau and Lips at 0.9 and 1.3ppm) were used in the first step. Only 6 observations representing the concentration of the most representative metabolites in identifying intratumoral histopathological variability (Lips at 0.9 and 1.3ppm, Lac, Cr, NAA and PCh) were further considered in the second step. This choice was based on the conclusions drawn in previous studies which showed that there is a high correlation between the concentrations of Lips, Cr, NAA and PCh0 and the histopathological tissue properties (Cheng et al. 2000; Andronesi et al., 2008; Opstad et al., 2008a; Croitor Sava et al., 2009).

Although a direct validation of the results using histopathological analysis is not feasible when analyzing *in vivo* MRSI data, for all voxels, the results were confronted with the tissue labeling proposed by the experts in spectroscopy.

4. Results

For each individual MRSI examination, a set of two components have been obtained in the first step, as well as an abundance matrix. The results on an MRSI image of a patient with glioblastoma are presented in Figure 1 which displays the extracted components and the two corresponding maps of the abundance of each component for the different voxels. The abundance matrix is drawn using a grey scale map showing the values of each component for the different voxels. It takes values between 0 and 1, represented by black and white, respectively, as shown in Figure 1.b.

The contour of the abnormal tissue area is very similar to the contour detected in the MR image. The first component is present in high abundancy in the center of the tumor, in the necrotic area. Its metabolic profile is very different from the normal tissue component. It presents elevated levels of Lips and Lac, while the concentration of the other metabolites is very low. The abnormal tissue area contains many voxels where a mixture of these two components is obtained. The second component, the normal tissue component, is present only in the healthy part of the brain and shows a high degree of similarity with the measured spectra obtained from healthy tissue, with high NAA/Cho levels and with metabolic profiles very close to the conclusions reported in the literature (Devos, 2005b).

We also compared the results obtained on the full magnitude spectra with the results on feature vectors. See Figure 2, where each voxel is labeled based on the highest abundance component. Results obtained on magnitude spectra and metabolite feature space show that with both approaches we obtain very similar tissue assignment.

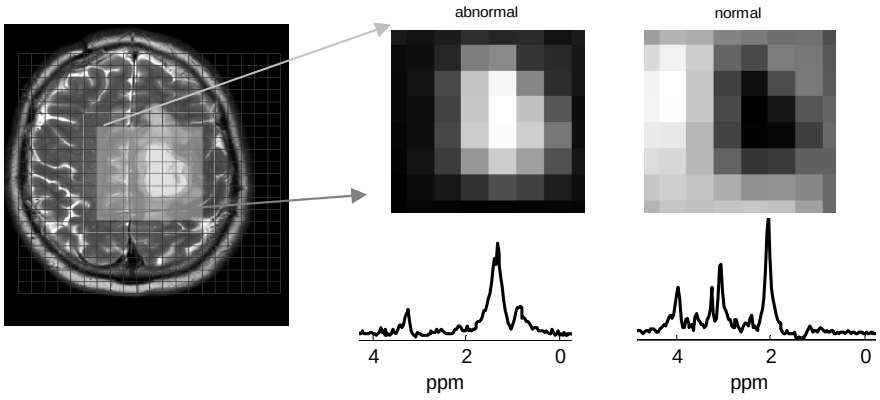


Figure 1 *a. MR T2-weighted image of a glioblastoma patient. The contour of the MRSI grid is marked with light grey on the image. In b. the results of the first step using NNMFSC are visualized: abundance maps showing the voxels identified as abnormal tissue and predominantly normal brain tissue (the upper part of the image); lower part shows the component profiles as identified with NNMFSC.*

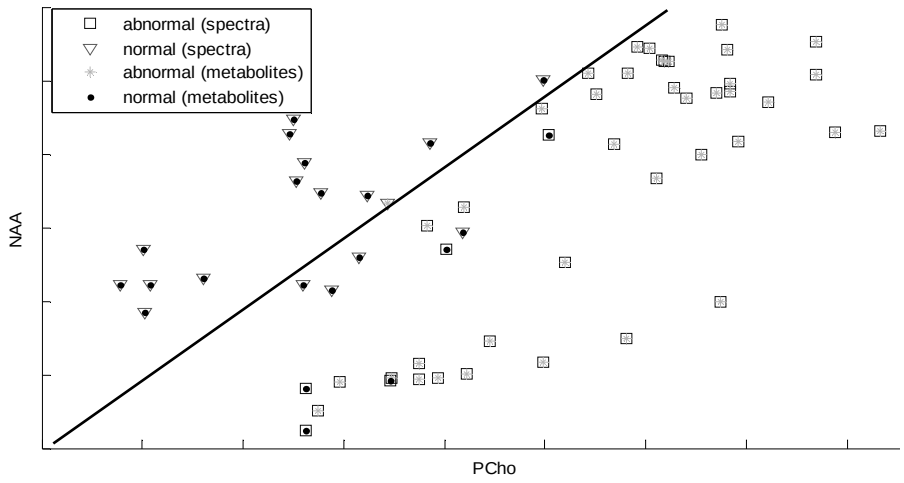
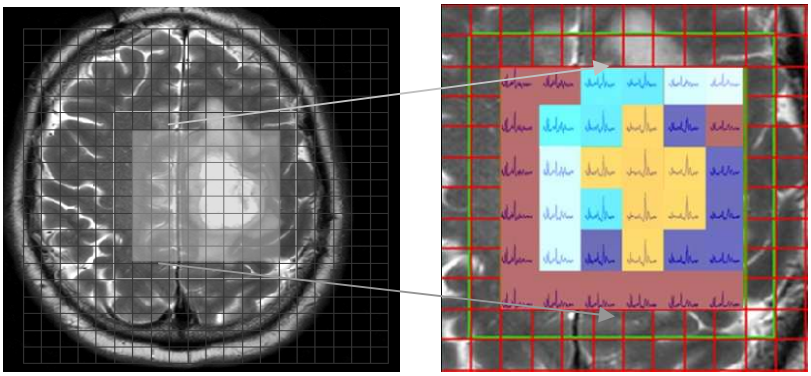


Figure 2 *Plot showing the label of each voxel, within the MRSI grid, based on its highest abundance component. Results on magnitude spectra and metabolite feature space are compared.*

In the second step, NNMFSC results applied in the abnormal tissue region, for *k=3*, are converted to nosologic images. The voxels identified as predominantly necrotic tissue, voxels with viable tumor cells and voxels with predominant normal brain tissue are visualized. Each histopathological tissue class is represented by a different color and therefore the images are easily interpretable. See Figure 3 and Figure 4, where the results on the GBM patient after the second step are presented.



b.

Figure 3 a. *MR T2-weighted image of a glioblastoma patient. The contour of the MRSI grid is marked with light grey on the image. b. Nosologic image showing the voxels identified as predominantly necrotic tissue (orange color), voxels with viable tumor cells (blue color; dark blue stands for high cellularity; lighter blue reflects higher levels of infiltrations from normal tissue, the lighter the blue the higher the percentage of infiltrations) and voxels with predominantly normal brain tissue (dark red). This image was created using NNMFSC, with metabolite concentrations as features.*

Figure 4 illustrates the components obtained in the second step. The border and the highly cellular tissue component are surrounding the necrotic tissue area. Where the infiltrations with normal tissue are high, also the levels of NAA and Cr are elevated, while in the highly cellular regions high levels of PCho are observed.

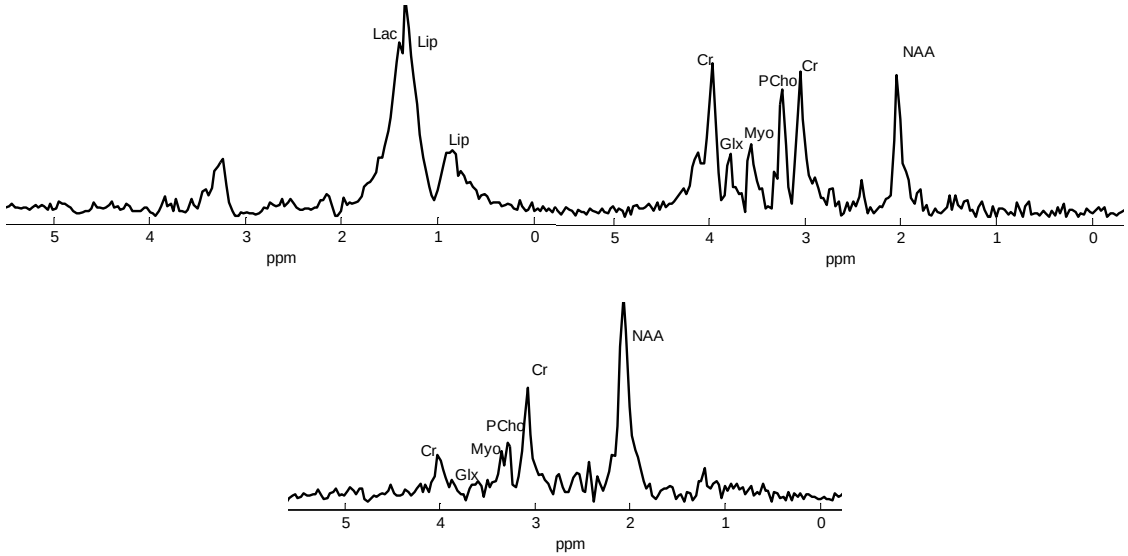


Figure 4 *The NNMFSC 2nd step sources obtained for k=3.*

5. Conclusions

The nBSS method proposed for obtaining characteristic profiles for each tumor tissue histopathological property and their abundance within a spectrum can reliably answer the problem of source separation when analyzing MRSI data. Moreover, we can decompose the observed MRSI grid into constituent tumor tissue sources with different predominant intratumoral histopathological properties and further quantify the abundance of each considered tissue source which can then be explored as nosologic images. This approach can provide relevant additional information for a better interpretation and classification of *in vivo* MRSI data and therefore enhance its contribution to brain tumor classification. In particular, this method can be of added value in addressing difficult questions such as the grading of glial tumors or differentiating metastasis from glioblastoma. Also it requires no previous training set, which often can be a problem when dealing with new measurements or with rare tumors. The finding in this study can provide relevant additional information for a better interpretation and classification of brain tumor tissue in *ex vivo* high resolution MRS and *in vivo* MRS(I), respectively.