Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 184 (2007) 292–301
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Fast nosologic imaging of the brain

M. De Vos [a,*], T. Laudadio [a], A.W. Simonetti [a], A. Heerschap [b], S. Van Huffel [a]

[a] *Katholieke Universiteit Leuven, Department of Electrical Engineering, Division ESAT–SCD (SISTA), Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium*
[b] *University of Nijmegen, University Medical Center St. Radboud, Department of Radiology, Geert Grooteplein Z18, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands*

Received 10 April 2006; revised 2 October 2006
Available online 22 November 2006

Abstract

Magnetic resonance spectroscopic imaging (MRSI) provides information about the spatial metabolic heterogeneity of an organ in the human body. In this way, MRSI can be used to detect tissue regions with abnormal metabolism, e.g. tumor tissue. The main drawback of MRSI in clinical practice is that the analysis of the data requires a lot of expertise from the radiologists. In this article, we present an automatic method that assigns each voxel of a spectroscopic image of the brain to a histopathological class. The method is based on Canonical Correlation Analysis (CCA), which has recently been shown to be a robust technique for tissue typing. In CCA, the spectral as well as the spatial information about the voxel is used to assign it to a class. This has advantages over other methods that only use spectral information since histopathological classes are normally covering neighbouring voxels. In this paper, a new CCA-based method is introduced in which MRSI and MR imaging information is integrated. The performance of tissue typing is compared for CCA applied to the whole MR spectra and to sets of features obtained from the spectra. Tests on simulated and *in vivo* MRSI data show that the new method is very accurate in terms of classification and segmentation. The results also show the advantage of combining spectroscopic and imaging data.

Keywords: Magnetic resonance spectroscopic imaging; Magnetic resonance imaging; Tissue segmentation; Canonical correlation analysis; Classification; Brain tumors

1. Introduction

Magnetic resonance spectroscopy (MRS) is increasingly being used in the diagnosis of brain tumors, to choose the most appropriate therapy and to perform therapy follow-up. With single voxel proton MRS, tissue concentrations of metabolites can be determined in a well-defined voxel in a non-invasive way. This can reduce the need for a biopsy, which is not without risks for the patient [1]. However, analysis of MR spectra requires a lot of expertise from the radiologist. To reduce the need of human interaction, several classifiers have been proposed that try to distinguish tumors in an automatic way, e.g. [2–7]. Although single voxel proton MRS is a relatively fast method to characterise the metabolic signature of the tumor, radiologists are also interested in its shape, size and (metabolic) heterogeneity. This information can be obtained with multivoxel proton MRSI, in combination with high-quality anatomical MR images [8]. During an MRSI examination, metabolic information is measured in a grid of voxels rather than in one voxel. This allows radiologists to extract histopathological information from each voxel as well as information about the heterogeneity of the tumor. However, due to the large amount of data that is acquired and needs to be processed during the investigation of a patient, there is an even stronger need for automatic 'classification' strategies than for single voxel MRS. More correctly, one should speak of 'segmentation' strategies as the goal is to segment an MRS image into regions of different tissues. We will explain how the proposed segmentation method is also able to label the

* Corresponding author. Fax: +32 16 32 19 26.
E-mail address: maarten.devos@esat.kuleuven.be (M. De Vos).

1090-7807/$ - see front matter
doi:10.1016/j.jmr.2006.10.017

detected segments simultaneously. A variety of methods is available in the literature, and in particular pattern recognition methods are commonly used to segment MRSI data [3,9]. The result of a segmentation is a nosologic image [10]. Such an image summarizes the available information by coloring each pixel according to the determined histopathological class. In this way, nosologic images can be routinely and easily used by radiologists for tumor diagnosis.

In this paper, we extend this study in various ways. First of all, we focus on short echo-time spectra of the brain. We use short echo-time spectra, because they are more relevant than long echo-time spectra for tumor classification and grading due to the presence of signals of more metabolites, such as myo-inositol, glutamate and glutamine [7,11–14].

Second, for brain tumor diagnosis, it is not sufficient to segment an MRSI image in normal and tumor tissue. Several types and grades of brain tumors may exist in different patients. For choosing the most appropriate therapy, it is very important to distinguish between different types of tumor. In order to deal with different tumor types, we introduce a two-steps CCA approach. We will explain how the first step is a tumor typing step, and how in the second step a segmentation is obtained.

Third, it is shown how to combine multimodal information and the influence of adding MRI information to the MRSI data is investigated. Four MR images with different contrasts (see Section 3) were acquired for every patient. This extra information is shown to improve the classification performance in other pattern recognition studies [9,10]. To integrate MRSI and MRI information in a CCA-based method, a dimension reduction has to be performed on the MR spectra by the extraction of characteristic features [15]. Most other classifiers or pattern recognition methods also use features as input variables, because they cannot deal with high-dimensional input spaces. As the CCA-method can be applied either to a MR spectrum or to a set of features, it is very interesting to compare the results obtained with different CCA-approaches.

In this paper, we start with a short overview of the theory of CCA and how CCA is applied for nosologic imaging of the brain (Section 2). In Section 3 we describe how the simulations are constructed. Further, the accuracy of the method is investigated on baseline-corrected MR spectra. We also compare these results with the results of the CCA-method on features, without and with MRI information. The results on simulated and *in vivo* MRSI are discussed in Section 4, respectively. Finally, the main conclusions are formulated in Section 5.

2. Method

2.1. Canonical correlation analysis (CCA)

We shortly describe the mathematical idea of CCA [16]. For more details, we refer to [17]. A reliable implementation is described in [18]. Ordinary correlation analysis quantifies the relation between two variables x and y by means of a correlation coefficient ρ:

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

CCA is a multivariate extension of ordinary correlation analysis. Consider 2 multivariate zero-mean vectors$\mathbf{x} = [x_1(f),\ldots,x_m(f)]^{\mathrm{T}}$ and $\mathbf{y} = [y_1(f),\ldots,y_n(f)]^{\mathrm{T}}$, $f = 1,\ldots,N$. Two new scalar variables, X and Y, are generated as a linear combination of the components in $\mathbf{x}$ and $\mathbf{y}$:

$$X = w_{x_1}x_1 + \cdots + w_{x_m}x_m = \mathbf{w}_x^{\mathrm{T}}\mathbf{x}$$
$$Y = w_{y_1}y_1 + \cdots + w_{y_m}y_m = \mathbf{w}_y^{\mathrm{T}}\mathbf{y}$$

CCA computes the coefficients $\mathbf{w}_x$ and $\mathbf{w}_y$ that maximize the correlation between X and Y. These coefficients are called regression weights and X and Y are called canonical variates. The resulting correlation coefficient is called the canonical correlation coefficient.

2.2. Applying CCA to spectra

During an MRSI acquisition, MR spectra are measured in a grid of voxels. Based on the information contained in these MR spectra, the goal is to segment the image in different regions. All the voxels in one region belong to the same tissue type. Each voxel can be assigned individually to a certain tissue type by using a sophisticated single voxel classifier like LS-SVM [6] or simply by computing the correlation or distance between the spectrum in the considered voxel and a known tissue type model [10]. However, with this approach, the spatial information provided by multi-voxel MRS is not used. When all the surrounding voxels contain healthy tissue, the considered voxel also has a high probability to contain healthy tissue. CCA exploits this spatial information by considering a multivariate vector $\mathbf{X}$ that contains the spectrum of the voxel under investigation and some spectra of the surrounding voxels. Different spatial models can be used. In this study, the spatial 'star' model is used, defined as:

$$\mathbf{X} = \begin{pmatrix} x_5(f) \\ \frac{x_5(f)+x_2(f)}{2} \\ \frac{x_5(f)+x_4(f)}{2} \\ \frac{x_5(f)+x_6(f)}{2} \\ \frac{x_5(f)+x_8(f)}{2} \end{pmatrix}$$

with $x_i(f)$ defined in Fig. 1. In this case, it can be seen that the information contained in the voxel under investigation (x_5) is always used at least as much as the information in the surrounding voxels.

During the classification of a voxel, CCA also integrates as much spectral information as possible in a spectral model, the multivariate $\mathbf{Y}$. The first component y_1 of this spectral model for each tissue type is constructed as the mean of a database containing *in vivo* spectra, characteristic of that

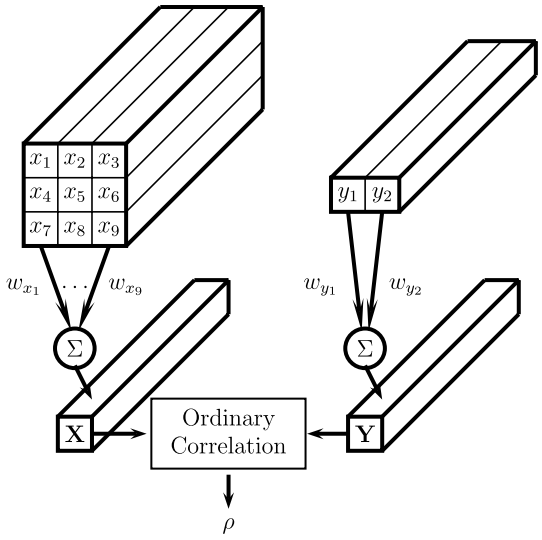

Fig. 1. Diagram of applying CCA to MRSI. We want to assign voxel x_5 to a tissue type. The other x_i are surrounding voxels that will also be used in the analysis of voxel x_5. y_1 is the model spectrum of a certain tissue, and y_2 takes into account realistic variations. CCA computes the regression weights w_{x_i} and w_{y_i} in order to maximize the correlation between X and Y.

tissue type. To model natural inter-spectral and inter-patient variability, a variational component y_2 is computed as follows. First the computed mean is subtracted from all *in vivo* spectra in the database. Secondly, PCA is performed on this mean-subtracted database. The first principal component contains most of the variations, and is used as y_2.

To obtain a segmentation with CCA, canonical correlation coefficients are computed with a spatial model and the spectral models of different tissue types. A visual representation of the analysis of one voxel (x_5) with a certain tissue type is given in Fig. 1.

The computed canonical correlation coefficients indicate how likely it is that the considered voxel belongs to the different tissue types. In the case of three tissue types, these three canonical correlation coefficients can be displayed in the three-dimensional coordinate system (first canonical correlation coefficient, second canonical correlation coefficient and third canonical correlation coefficient). The distribution of the voxels in this coordinate system is displayed for a simulated MRS image with normal, mixed and tumor tissue in Fig. 2. Three clouds of points can be distinguished, corresponding to the three present tissue types. The voxels, indicated by circles, have a high correlation with normal tissue, the voxels, indicated with stars, have a high correlation with tumor tissue and the voxels, indicated with diamonds lie between both clouds. A clustering method [19] can be applied to these canonical correlation coefficients. However, in order to avoid well-known problems of clustering algorithms (e.g. initialisation, number of clusters, ...), another criterion is used to assign a voxel to a certain tissue type: the highest canonical correlation coefficient.

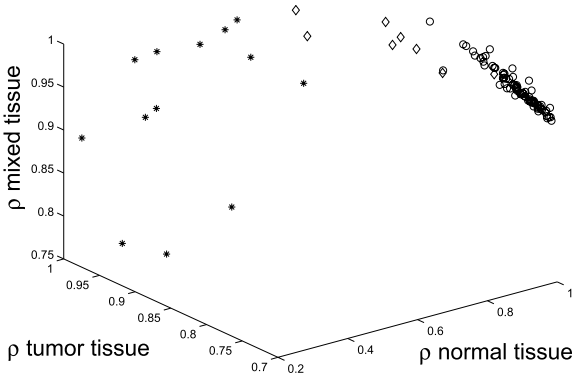

Fig. 2. All voxels in an MRSI grid are displayed in the coordinate system of the computed canonical correlation coefficients with different tissue types. Three clouds corresponding to the three present tissue types are visible.

2.3. A two-steps CCA-based method

As there exist several types of brain tumor malignancy, a model spectrum has to be constructed for each tumor type and for each corresponding mixed tissue since it can be expected that only one tumor type is present in a patient. This mixed tissue is composed of fractions of the model spectrum for healthy and tumor tissue.

CCA is able to discriminate between different tumor types, but we noticed that voxels with mixed tissue are not always correctly detected. To avoid this problem a two-steps method is proposed. The first step is a tumor typing step and the second the segmentation step. In the first step only the model spectra of normal tissue, CSF and the different tumor spectra are used. As pathologists type the tumor with respect to the most malignant tissue, this step is performed to detect the most malignant tissue present. By introducing this second step, we also ensure that no different tumor types will be found (e.g. meningioma and a glial tumor) inside the same tumor mass. In the second step, CCA is applied with the model spectrum from CSF, normal tissue, the detected tumor tissue and corresponding mixed tissue. In this step, the heterogeneity of the tumor is shown.

2.4. Feature-based approach

When we want to integrate the available MRSI and MRI information in the CCA-based method, harmonisation of both input spaces is necessary. This multimodal integration requires 2 steps, which are described.

At most 4 images with different contrasts are acquired during a study. These 4 extra variables per voxel will not influence the result, when they are added to the spectra of 1024 variables. The spectra have to be summarized into representative features first. This dimension reduction can be done by quantification of the spectra. This is meaningful, since the quantified amplitudes are a direct estimate of the concentrations of the most important metabolites.

For the present study AQSES [20] was used, that fits an *in vivo* spectrum in the complex time domain as a linear combination of *in vitro* metabolite spectra. To avoid overfitting, only the 8 most important metabolites (myo-inositol, choline, creatine, glutamine, glutamate, NAA, alanine and lactate) were quantified, together with the lipid peaks at 0.9 and 1.3 ppm. These metabolite levels formed the first 10 values in the set of features.

In addition to feature reduction, also the spatial resolution of both imaging modalities has to be matched. This can be done in several ways. We decided to lower the resolution of the MR image to MRSI voxel size.

The images with 4 contrasts were aligned with respect to the spectroscopic image and the MR pixel intensities were averaged out over a spectroscopic voxel. These 4 image variables formed the last 4 values in the set of features.

Although the application of the CCA method either on spectra or on features does not differ, the implementation does. The aim of the spatial model of CCA is to reduce the noise by 'averaging out' some voxels. In order to obtain a realistic solution, all the regression weights of X should have the same sign. When the CCA-method is applied to spectra, the mathematically optimal solution satisfies most of the time this constraint. However, in the feature case, this constraint is violated with the implementation given in [18], and has to be imposed explicitly. In [21,22] it is proven that an elegant constrained CCA solution exists in the special case the regression weights are imposed to be non-negative. Such a constrained CCA solution equals an *unconstrained* CCA solution when one or several variables have been excluded.

2.5. Algorithmic description of feature-based approach

For reasons of clarity, we resume our method:

- STEP 1: Tumor typing:For every voxel:
  1. Compute constrained CCA with spectral model of normal tissue
  2. Compute constrained CCA with spectral model of CSF
  3. Compute constrained CCA with spectral model of first tumor type
  4. ...
  5. Compute constrained CCA with spectral model of last tumor type

  and type the voxel according to the highest canonical correlation coefficient.The tumor type is detected and used in STEP 2.
- STEP 2: Segmentation:For every voxel:
  1. Compute constrained CCA with spectral model of corresponding mixed tissue
  2. Compare this canonical coefficient with the coefficients, computed in STEP 1 of normal tissue, CSF and detected tumor tissue

  and type the voxel according to the highest canonical correlation coefficient. This gives the segmentation.

3. Materials

3.1. Data acquisition

In this study, data from 24 patients with a brain tumor and from 4 volunteers are used. These data were obtained within the EU project INTERPRET [23], and acquisition has been approved by the ethical committee of the University Medical Center Nijmegen, and from all patients written informed consent was obtained. Each patient's tumor was diagnosed by consensus on a histopathological study. The diagnosis was done according to the rules of the World Health Organization (WHO). 10 patients had a glial brain tumor of grade II (consisting of astrocytic, oligoastrocytic and oligodendroglial tumors), 5 of grade III (consisting of astrocytic, oligoastrocytic and oligodendroglial tumors) and 7 of grade IV (glioblastoma) and 3 patients were diagnosed with a meningioma tumor. The MR data were acquired with a 1.5 T Siemens Vision whole body scanner, using a CP-head coil. For every patient first 4 MR images were acquired: T_1 weighted (TE/TR = 15/644 ms), T_2 weighted (TE/TR = 16/3100 ms), proton density weighted (TE/TR = 98/3100 ms) and a Gadolinium enhanced T_1 image (15 ml of 0.5 M Gd-DTPA). Also both water suppressed and unsuppressed proton MR Spectroscopic Images were acquired. This MRSI data set was acquired using a 2D STEAM pulse sequence with the STEAM box positioned in a transversal plane through the brain showing the largest tumor in the Gd-image. MRSI parameters were: $16 \times 16 \times 1024$ 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 [24], followed by water removal with HLSVD-PRO [25], and finally a simple baseline correction was done by filtering the time signal with a highpass filter with a filterwidth of 5 ms, followed by the subtraction of the residual from the original time domain signal. Zero order phase was already corrected during eddy current correction. Manually, the first order phase of the average spectrum of a patient was used to reset the first order phase of all spectra of that patient. Finally, all spectra were normalised with respect to the water signal [26].

3.2. Construction of simulated MRS images

From the *in vivo* data set, a large database of representative spectra for the different tumor types, normal tissue and CSF was created for simulating purposes. This database was divided in two parts. From each tissue type, 25 spectra were used to construct the spectral model, shown in Fig. 3. As can be seen, the spectra show that the different tissue types are characterised by different contributions of the metabolites of interest. The other spectra were used to simulate 10×10 MRS images for each tissue type as shown in Fig. 4. The mixed spectra for all tumor types in these MRS images were artificially generated by averaging out a normal and a tumor spectrum.

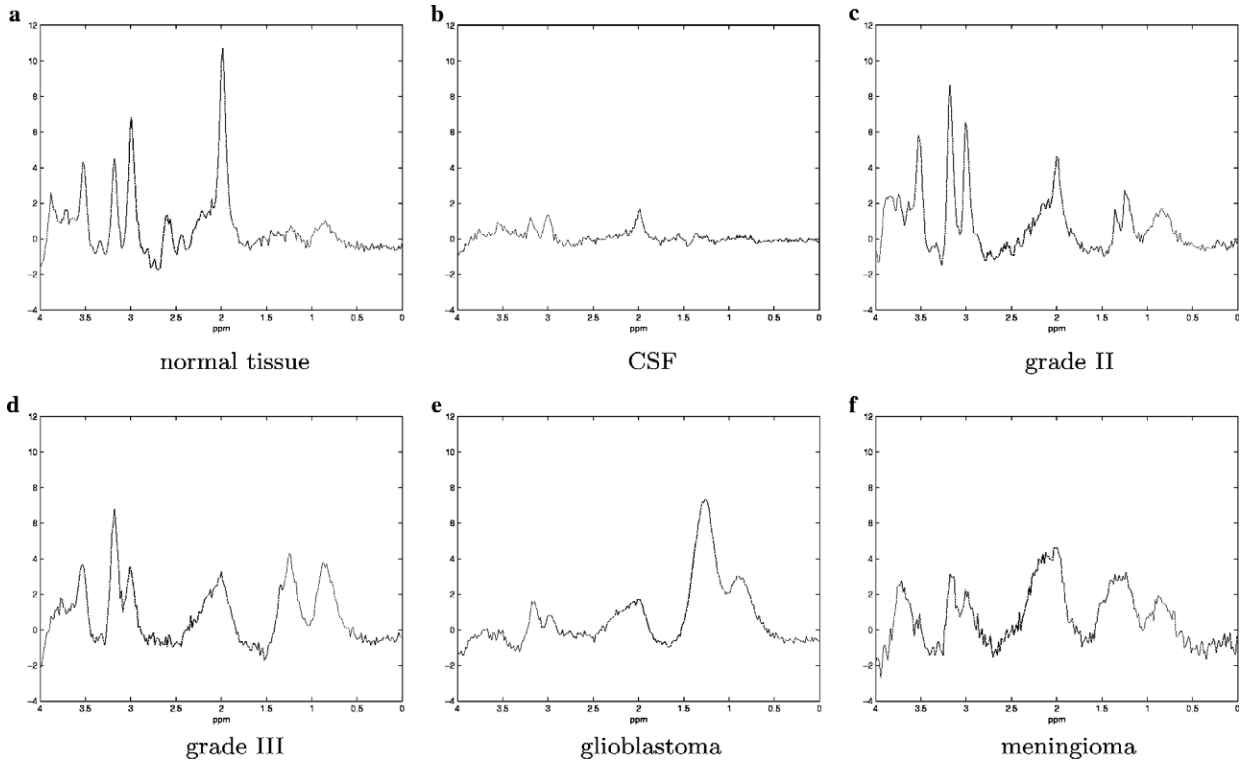

Fig. 3. The first component ($\mathbf{y}_1$) of the spectral model for the different tissue types, present in this study. Note that the residual metabolite signals in the spectrum of CSF are due to a partial volume effect (contamination with other brain tissue). (a) Normal tissue, (b) CSF, (c) grade II, (d) grade III, (e) glioblastoma, (f) meningioma.

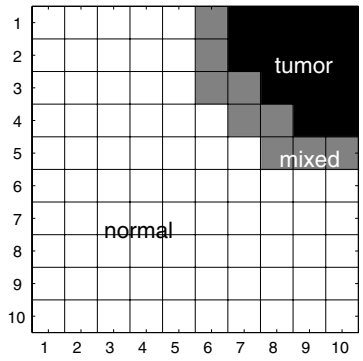

Fig. 4. Artificially created MRS image, partitioned into normal, mixed and tumor tissue. This partition was used for all the simulated MRS images.

Three types of simulated MRS images were created in order to evaluate the performance of CCA. In the first experiment MR spectra were used as input, the second experiment was done with a set of 10 quantified metabolite concentrations and the third experiment used a set of 14 features (10 metabolite concentration variables and 4 image variables) as input. For the experiments, the database was randomly split 5 times in order to obtain different model spectra and simulated MRS images. As there were four different tumor types in the database (grade II, grade III, glioblastoma and meningioma), the CCA-method was evaluated on 20 images with different spatial models. For comparison the result of ordinary correlation was computed.

For experiment 2, the whole database was quantified. Using the same partitioning of the database as in the spectra case, the 20 corresponding grids of sets of metabolite levels were created.

For experiment 3, the image variables were added to the quantified metabolite levels.

4. Results and discussion

4.1. Simulation study

Fig. 5 demonstrates the necessity of the two-steps method. The left figure shows the result on a simulated MRSI example of grade III when all the constructed spectral models are used. A large region with normal tissue (1) and a region with tumor tissue of grade III (4) are detected. In the transition region, voxels are assigned to grade II (3), mixed tissue of grade II (7), mixed tissue of grade III (8) and mixed tissue of meningioma (10). We simulated the image with only grade III spectra and spectra of mixed grade III. A major part of the voxels are assigned incorrectly. In Fig. 5b, the first step of the proposed two-steps method is applied: only spectral models of normal and pure tumor are used. A region with normal tissue (1) and a region with grade III tumor (4) are segmented. In addition, two voxels are assigned to grade II (3). As clinicians type the

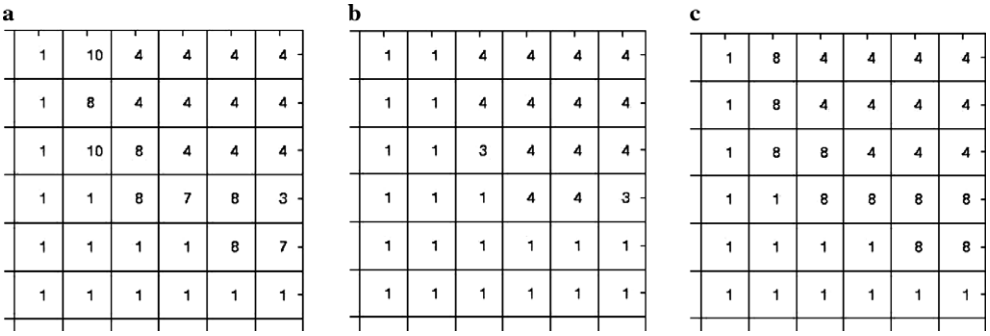

Fig. 5. The result of CCA on spectra with spatial 'star' model on a simulated MRSI example with a grade III tumor. Only the tumor region is shown. (a) When all the constructed spectral models are used, (b) step one for tumor labeling: only spectral models of normal and tumor tissue are used, (c) step two for tissue segmentation: only spectral models of normal tissue, the detected tumor and corresponding mixed tissue are used. Explanation of numbers: 1 normal; 2 CSF; 3 grade II; 4 grade III; 5 glioblastoma; 6 meningioma; 7 mixed grade II; 8 mixed grade III; 9 mixed glioblastoma; 10 mixed meningioma.

tumor with respect to the most malignant tissue present, the detected tumor type is convincingly grade III. In our second step (Fig. 5c), the MRS image is analysed with spectral models of normal tissue (1), tumor (4) and mixed (8) tissue of grade III. Three homogeneous regions are segmented. Compared to Fig. 5a, this result is certainly more correct and more easy to interpret.

On all the simulated MRS images, CCA found in the first step the correct tumor type. The simulation may be biased by the fact that the simulated MRSI are constructed with a database containing only carefully selected spectra. In the rest of this section, we only discuss the segmentation accuracy.

Before discussing more general results, it is useful to have a closer look at the segmentation result on one simulated MRS image. Fig. 6 compares the result of ordinary correlation (a (spectra) and b (14 features)) with the result of CCA (c (spectra) and d (14 features)) using the spatial 'star' model on a grade II tumor. The red, green and blue regions represent respectively the detected tumor, mixed and normal tissue. The example clearly shows the advantage of using spatial information: ordinary correlation does not segment homogeneous regions, but assigns a lot of isolated voxels to the wrong tissue type. Especially the distinction between tumor tissue and mixed tissue is not done correctly. The CCA-based method detects correctly the tumor region, surrounded by a band of mixed tissue. This results in a segmentation accuracy of 98% for CCA. It is observed that most of the misclassified voxels are corner voxels. This is logical, because corner voxels are surrounded by less voxels of the same tissue type. On Fig. 6, we cannot see differences between the CCA-method on spectra and features. Both perform in this case similarly. It is however interesting to note that misclassified voxels in the spectra case not necessarily correspond to misclassified voxels in the feature case.

To compare the different CCA approaches in a general way, the segmentation accuracy of the 20 segmented images is obtained by simply counting the number of correctly classified voxels. The average value is given in Table 1. The columns of the table correspond to three kinds of input vectors: using the MR spectra, using the quantified metabolite levels or using the features.

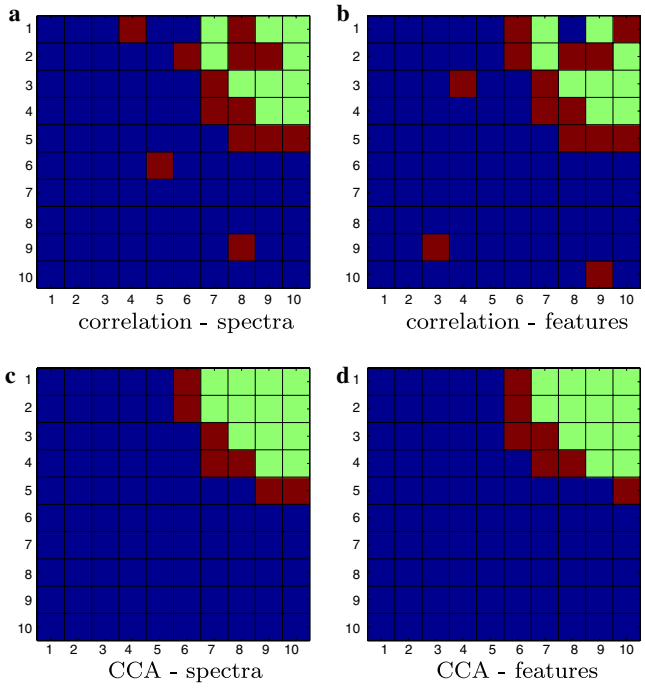

Fig. 6. The result of (a) ordinary correlation on spectra, (b) ordinary correlation on features, (c) CCA on spectra, (d) CCA on features.

Table 1
The segmentation accuracy of the segmented image by CCA and cross-correlation

	Spectra	Metabolite levels	Metabolite levels + image vars
CCA	97.3	95	98.2
Ordinary correlation	91	93.7	95.5

The displayed value is the mean value of the results on 20 artificially generated MRS images with different tumor types. Values in %.

The lower performance of segmentation accuracy of the ordinary correlation analysis compared to CCA is

confirmed. When the different input strategies are compared, it is clearly seen that quantifying the spectra before applying CCA reduces the segmentation accuracy. This approach should never be used, as the quantification step can introduce errors. However, quantifying the spectra makes sense when image information is added. The values displayed in the third column are higher than the values in the second column, which shows that adding image information improves the segmentation result. The accuracy with feature sets of 14 variables is slightly higher than the accuracy of the spectra case. However, the differences are very small. Keeping in mind that we only work with 20 carefully constructed simulated examples, we can state that the spectra and feature approaches perform similarly in terms of accuracy.

4.2. In vivo

Also several *in vivo* MRS images were analysed. The CCA-results on 4 patients are discussed below. We focus in this part on differences between the two valuable approaches, namely CCA on spectra and CCA on feature sets of 14 variables. Care was taken that the model spectrum was constructed with 25 spectra of other patients. So independency is guaranteed. The colors of the segments in Figs. 7–10 have a specific meaning. As the used tissue types differ in the tumor typing and segmentation steps, it is not possible to use in both steps the same color legend. After the tumor typing step, dark blue is used for normal tissue, light blue for CSF, green for grade II, yellow for grade III, orange for glioblastoma and red for meningioma. After the segmentation step, blue represents normal tissue, green mixed tissue, orange the detected tumor tissue and red is used for CSF.

Fig. 7 shows the result of CCA on features for a patient with a glioblastoma tumor. In the first step several glioblastoma voxels are detected, but also several voxels at the boundary of the glioblastoma region are assigned to grade III. This can be expected since tumors are known to be very heterogeneous. The border between normal glial tissue and the center of the tumor is often lower in grade. The diagnosis of the method is correctly glioblastoma. In the second step most 'grade III voxels' in Fig. 7a are assigned to mixed tissue in Fig. 7b as the model spectra of grade III and mixed glioblastoma are very similar. It can also be seen that the segmented regions with CSF correspond nicely to the ventricles in the MR image.

In Fig. 8 the results with spectra and features are compared after the tumor typing step on a tumor of grade III. Although the detected tumor region is large, the CCA method on spectra (Fig. 8a) only detects one voxel of grade III in the first step. This tumor is probably very heterogeneous and only a very small part of the tumor will really be of grade III. For this patient, the diagnosis based on the most malignant tumor present, will still be correct (grade III) but this example illustrates the limitations of methods that improve robustness by including spatial information. When the tumor region is too small, the tumor will be hidden by the surrounding voxels. Compared to the result obtained with features (Fig. 8b) we can clearly state that the result of CCA on features is better, as there is more convincingly a grade III region detected (like in [9]). In the second step, both approaches gave similar results.

Fig. 9 compares the segmentation results of both approaches on a meningioma tumor. The first step is not shown as a region with meningioma is convincingly detected by both CCA approaches. In the second step, the present meningioma tumor is nicely segmented in both figures. But it is also seen that the CCA method based on spectra segments three lines as a (homogeneous) region of mixed tissue. These lines contain normal spectra. However, the spectra in the lower part are of minor quality compared to the spectra in the upper part of the grid.

Finally, Fig. 10 compares on the same patient the results of both approaches after step 1 and step 2. For this patient no histopathological agreement was reached. The pathologists who validated the tumor disagreed between grade II and grade III. Both CCA on spectra and CCA on features classify the large area as grade III. The result of ordinary correlation is not shown, but ordinary correlation assigned two voxels to glioblastoma tissue (grade IV). These are clearly misclassifications, that will lead to the wrong diagnosis. These misclassifications are avoided by the use of spatial information. The most interesting part of this example is however the right ventricle. CCA on spectra assigns these CSF-voxels wrongly to grade II (step 1) or mixed tissue (step 2). When features are used as input, the two ventricles are nicely segmented and correctly classified. This is certainly due to the incorporation of the image variables because in the MR images the ventricles are clearly visible. It can also be noted that in this figure and in Fig. 8, the segmentation boundaries differ between both approaches. However, it cannot be said which segmentation is the best, because no golden standard is available. A comparison between the obtained segmentation and the high-resolution MR images shows that the boundary of mixed voxels obtained with the method makes sense.

In [18], an extensive study was done in order to demonstrate the computational efficiency of the CCA-based segmentation. A segmentation based on the original MR spectra took some seconds, the segmentation based on features in this study was even faster due to the limited number of features. However, quantitation of the spectra takes 10–20 min with AQSES. Taking this into account, the overall segmentation on the original MR spectra is the fastest segmentation.

Finally, we have to remark that the use of this accurate method has one drawback. A reliable and representative model spectrum has to be constructed for the present tissue types, in order to obtain a meaningful result. To construct all the spectral models, a database of carefully validated spectra has to be generated. This takes time, especially for rarely present tumor types. Constructing a spectral model is even 'impossible' when there are 'unexpected'

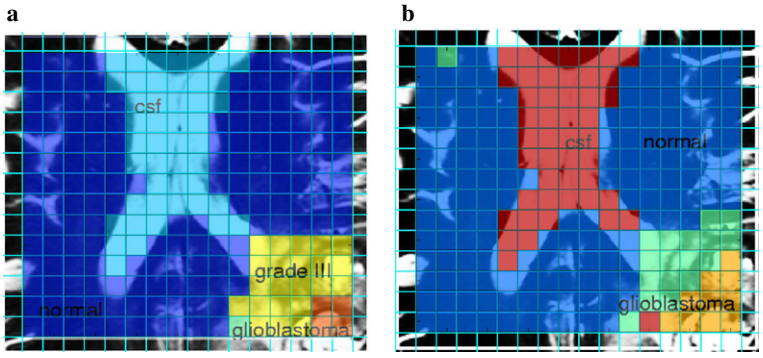

Fig. 7. The result of CCA on features (a) after the tumor typing step and (b) after the segmentation step.

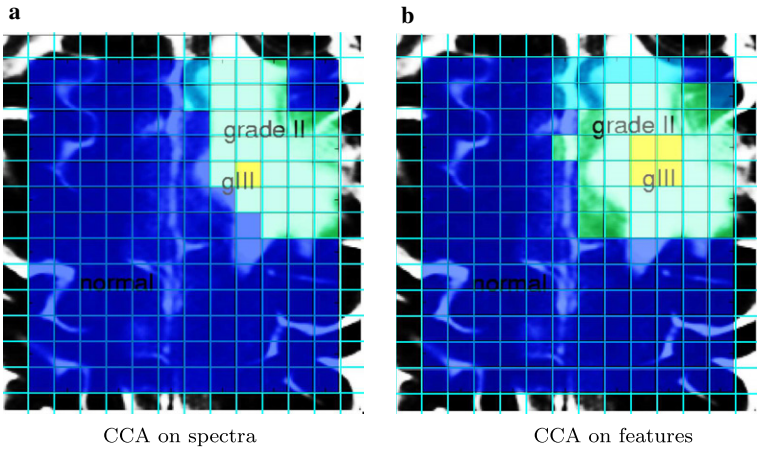

Fig. 8. The result of CCA after the tumor typing step (a) applied on spectra and (b) applied on features.

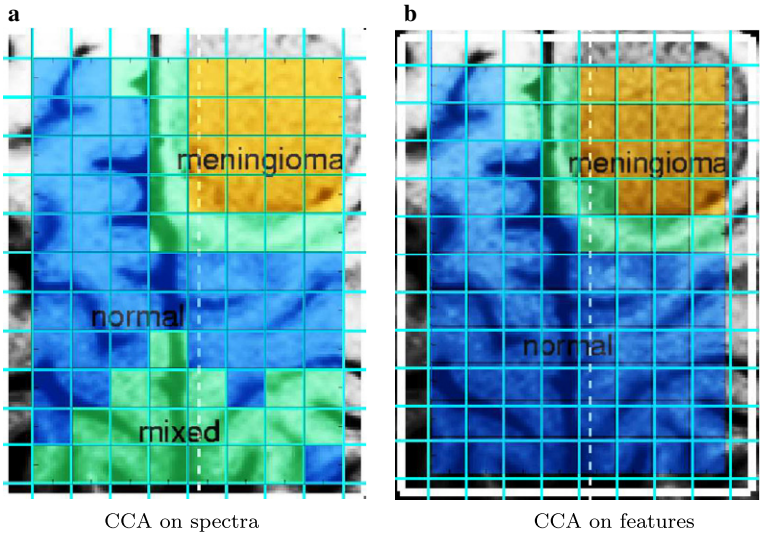

Fig. 9. The result of CCA after the tumor typing step (a) applied on spectra and (b) applied on features for a patient with a meningioma tumor.

tissue types in a new MRSI dataset. In this study, 'unexpected' tissue types are e.g. edema, necrosis, …. During the on-going European eTumour project [27], a large database of all possibly expected tissue types (adults & children) is being constructed. With this database, optimal spectral models can be generated. We expect that our CCA-based method with these spectral models will in the future reliably assign MRSI voxels to the correct histopathological class.

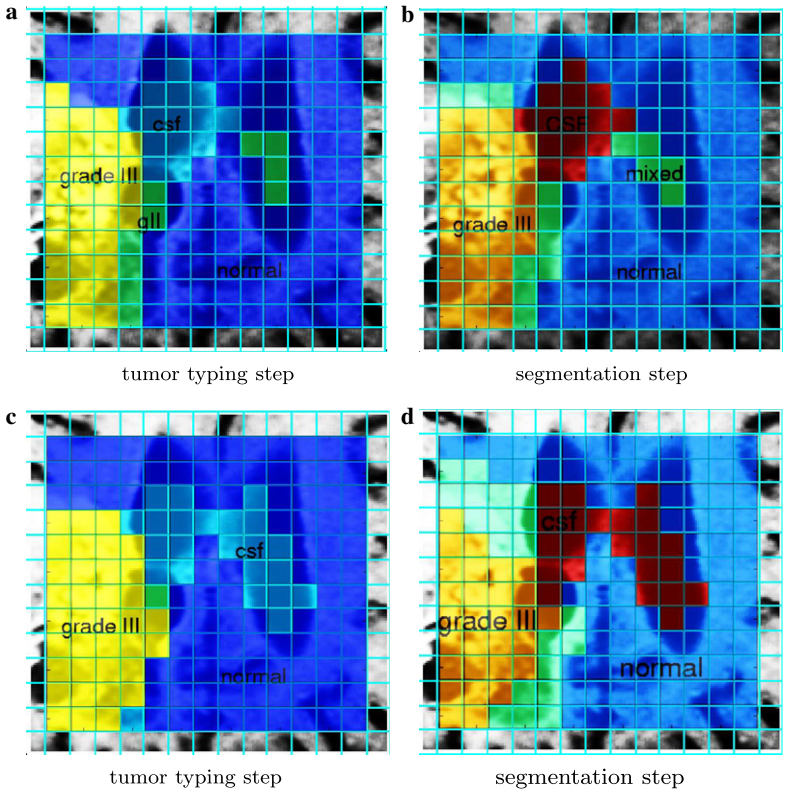

Fig. 10. The result of CCA on spectra (a) after the tumor typing step and (b) after the segmentation step and the result of features in (c) and (d).

5. Conclusion

CCA is a technique that quantifies the relationship between two multivariate vectors. In [18] CCA was successfully applied to segment long echo-time prostate MRSI. In this paper, a CCA-based method is introduced to translate short echo-time brain MRS images into easy-to-understand nosologic images. In order to distinguish in a reliable way between different brain tumor types, a two-steps method is presented. The first step is a tumor typing step and in the second step the heterogeneity of the detected tumor is shown. In addition, we introduce a CCA-based method on feature sets. This feature sets integrate information of MRSI and MRI. We also show that the CCA-based method on such feature sets outperforms the CCA-based method on the whole spectra in *in vivo* cases.

Acknowledgments

We thank the whole Department of Analytical Chemistry at the University of Nijmegen for the preprocessing of the *in vivo* data. This research is funded by a PhD grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). Research supported by Research Council KUL: GOA-AMBioRICS, CoE EF/05/006 Optimization in Engineering, 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 (Non-invasive brain oxygenation), G.0321.06 (Tensors/Spectral Analysis), research communities (ICCoS, ANMMM); IWT: PhD Grants; Belgian Federal Science Policy Office IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling'); EU:INTERPRET(FP5), BIOPATTERN (FP6-2002-IST 508803), ETUMOUR (FP6-2002-LIFESCIHEALTH 503094), Healthagents (IST200427214).

References

[1] M. Bernstein, A.G. Parrent, Complications of CT-guided stereotactic biopsy of intra-axial brain lesions, J. Neurosurg. 81 (1994) 165–168.
[2] A.R. Tate, C. Majós, A. Moreno, F.A. Howe, J.R. Griffiths, C. Arus, Automated classification of short echo time spectra in *in vivo* brain spectra: a multicenter study, Magn. Res. Med 49 (2003) 29–36.
[3] M.C. Preul, Z. Caramanos, R. Leblanc, J.G. Villemure, D.L. Arnold, Using pattern analysis of *in vivo* proton MRSI data to improve the diagnosis and surgical management of patients with brain tumors, NMR Biomed. 11 (1998) 192–200.
[4] J. Lindon, E. Holmes, J. Nicholson, Pattern recognition methods and applications in biomedical magnetic resonance, Prog. Nucl. Magn. Res. Spectrosc. 39 (2001) 1–40.
[5] R. Duda, P. Hart, D. Stork, Pattern Classification, second ed., John Wiley, New York, 2001.
[6] A. Devos, A.W. Simonetti, M. van der Graaf, L. Lukas, J.A. Suykens, L. Vanhamme, L.M. Buydens, A. Heerschap, S. Van Huffel, The use of multivariate MR imaging intensities versus metabolic data from MR spectroscopic imaging for brain tumour classification, J. Magn. Res. 173 (2005) 218–282.
[7] B. Wilken, P. Dechent, J. Herms, C. Maxton, E. Markakis, F. Hanefeld, J. Frahm, Quantitative proton magnetic resonance spectroscopy of focal brain lesions, J. Pediatr. Neurol. 23 (2000) 22–31.

[8] S.J. Nelson, Minireview: multivoxel magnetic resonance spectroscopy of brain tumors, Mol. Cancer Therap. 2 (2003) 497–507.
[9] A. Simonetti, W. Melsen, M. van der Graaf, G. Postma, A. Heerschap, L. Buydens, A chemometric approach for brain tumor classification using magnetic resonance imaging and spectroscopy, Anal. Chem. 75 (2003) 5352–5361.
[10] F. Szabo De Edelenyi, C. Rubin, F. Estve, G. Sylvie, M. Dcorps, V. Lefournier, J.-F. Le Bas, C. Rmy, A new approach for analyzing proton magnetic resonance spectroscopic images of brain tumors: nosologic images, Nat. Med. 6 (2000) 1287–1298.
[11] M. Castillo, J.K. Smith, L. Kwock, Correlation of myo-inositol levels and grading of cerebral astrocytomas, Am. J. Neuroradiol. 21 (2000) 1645–1649.
[12] S.J. Barton, F.A. Howe, A.M. Romlins, S.A. Cudlip, J.K. Nicholson, B.A. Bell, J.R. Griffiths, Comparison of ^{1}H MRS of human brain tumours with ^{1}H HR-MAS spectroscopy of intact biopsy samples in vitro, Magn. Res. Mater. Phys. Biol. Med. 8 (1999) 121–128.
[13] G. Fan, B. Sun, Z. Wu, Q. Guo, Y. Guo, In vivo single-voxel proton MR spectroscopy in the differentiation of high-grade gliomas and solitary metastases, Clin. Radiol. 59 (2004) 77–85.
[14] C. Majos, M. Julia-Sape, J. Alonso, M. Serrallonga, C. Aguilera, J.J. Acebes, C. Arus, J. Gili, Brain tumor classification by proton MR spectroscopy: comparison of diagnostic accuracy at short and long TE, Am. J. Neuroradiol. 25 (2004) 1696–1704.
[15] A. Simonetti, W. Melsen, F. Szabo de Edelenyi, J. van Asten, A. Heerschap, L. Buydens, Combination of feature-reduced MR spectroscopic and MR imaging data for improved brain tumor classification, NMR Biomed. 18 (2005) 34–43.
[16] H. Hotelling, Relations between two sets of variates, Biometrika 28 (1936) 321–371.
[17] O. Friman, Adaptive analysis of functional MRI Data, PhD thesis, Linköping University, Sweden, 2003.
[18] T. Laudadio, P. Pels, L. De Lathauwer, P. Van Hecke, S. Van Huffel, Tissue segmentation and classification of MRSI data using Canonical Correlation Analysis, Magn. Res. Med. 54 (2005) 1280–1284.
[19] B. Mirkin, Clustering for data mining, Chapmann & Hall/CRC, Florida, USA, 2005.
[20] J.-B. Poullet, A.W. Simonetti, D.M. Sima, B. De Neuter, L. Vanhamme, Ph. Lemmerling, S. Van Huffel, An open source short echo time MR quantitation software solution: AQSES, NMR Biomed. (2006) submitted.
[21] S. Das, P. Sen, Restricted canonical correlations, Linear Algebra Appl. 210 (1994) 29–47.
[22] M. Omladic, V. Omladic, More on restricted canonical correlations, Linear Algebra Appl. 321 (2000) 285–293.
[23] http://azizu.uab.es/INTERPRET/.
[24] U. Klose, In vivo proton spectroscopy in presence of eddy currents, Magn. Res. Med. 14 (1990) 26–30.
[25] T. Laudadio, N. Mastronardi, L. Vanhamme, P. Van Hecke, S. Van Huffel, Improved Lanczos algorithms for blackbox MRS data quantitation, J. Magn. Res. 157 (2002) 292–297.
[26] Z. Tong, T. Yamaki, K. Harada, K. Houkin, In vivo quantification of metabolites in normal brain and brain tumors by proton MR spectroscopy using water as an internal standard, Magn. Res. Imag. 22 (2004) 1017–1024.
[27] http://www.etumour.net/.