NMR IN BIOMEDICINE
NMR Biomed. (2007)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI:10.1002/nbm.1190

Fast nosological imaging using canonical correlation analysis of brain data obtained by two-dimensional turbo spectroscopic imaging

Teresa Laudadio,[1,2]* M. Carmen Martínez-Bisbal,[3,4] Bernardo Celda[3,4] and Sabine Van Huffel[1]

[1]Department of Electrical Engineering, Division ESAT-SCD, Katholieke Universiteit Leuven, Leuven-Heverlee, Belgium
[2]Istituto per le Applicazioni del Calcolo ''M. Picone'', National Research Council, (IAC-CNR), Bari, Italy
[3]Departamento de Química-Física, Facultad de Química, Universidad de Valencia, Valencia, Spain
[4]CIBER of Bioengineering, Biomaterials and Nanomedicine, ISC-III, Valencia, Spain

Received 24 January 2006; Revised 3 May 2007; Accepted 3 May 2007

ABSTRACT: A new fast and accurate tissue typing technique has recently been successfully applied to prostate MR spectroscopic imaging (MRSI) data. This technique is based on canonical correlation analysis (CCA), a statistical method able to simultaneously exploit the spectral and spatial information characterizing the MRSI data. Here, the performance of CCA is further investigated by using brain data obtained by two-dimensional turbo spectroscopic imaging (2DTSI) from patients affected by glioblastoma. The purpose of this study is to investigate the applicability of CCA when typing tissues of heterogeneous tumors. The performance of CCA is also compared with that of ordinary correlation analysis on simulated as well as *in vivo* data. The results show that CCA outperforms ordinary correlation analysis in terms of robustness and accuracy. Copyright © 2007 John Wiley & Sons, Ltd.

KEYWORDS: MR spectroscopic imaging; turbo spectroscopic imaging; nosological images; canonical correlation analysis; tissue segmentation; classification

INTRODUCTION

NMR is widely used in oncology as a non-invasive diagnostic tool to detect the presence of tumor regions in the human body. One application of NMR is MRI, which is applied in routine clinical practice to localize tumors and determine their size. MRI is able to provide an initial diagnosis, but its ability to delineate anatomical and pathological information is significantly improved by its combination with another NMR application, MRS. The latter provides information on the biochemical profile of tissues, thereby allowing clinicians and radiologists to identify in a non-invasive way the different tissue types characterizing the sample under investigation, and to study the biochemical changes underlying a pathological situation (1–3). There is a particular NMR application that provides spatial as well as biochemical information. This application is called magnetic resonance spectroscopic imaging (MRSI) and involves the simultaneous acquisition of signals from many volume elements.

**Correspondence to:* T. Laudadio, Department of Electrical Engineering, Division ESAT-SCD, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium.
E-mail: laudadio@esat.kuleuven.be

Abbreviations used: CCA, canonical correlation analysis; 2DTSI, two-dimensional turbo spectroscopic imaging; MRSI, magnetic resonance spectroscopic imaging; OCA, ordinary correlation analysis; PCA, principal component analysis; NAA, *N*-acetylaspartate.

MRSI data have recently been exploited in tissue segmentation and classification techniques. To this aim, many methods have been developed on the basis of a variety of theoretical principles. Some examples are thresholding techniques (4), region growing techniques (5), clustering techniques (6), Markov random field models (7), classifiers (8), and artificial neural networks (9). In particular, a new fast and accurate tissue typing technique has been successfully applied to prostate MRSI data (10). This technique is based on canonical correlation analysis (CCA), a statistical method able to simultaneously exploit the spectral and spatial information provided by the MRSI data. More precisely, CCA is used to quantify the relationship between two sets of variables, spectra of the measured data and signal subspaces modeling the spectra of characteristic tissue types, by means of correlation coefficients. These coefficients are then exploited to construct nosological images (11) in which all the detected tissue types are visualized.

Several important properties make CCA a valid alternative to other tissue typing techniques. In fact, the detection results provided by CCA are very accurate because, as already specified above, it exploits the spatial information characterizing the MRSI data in addition to the spectral one. Moreover, it allows the use of prior biochemical knowledge available when working with this type of NMR data. Finally, it is numerically efficient as it

Copyright © 2007 John Wiley & Sons, Ltd.

is able to process the given dataset in a few seconds, and sometimes, depending on the dimensions of the data matrix, in less than 1 s. The latter property is important because large amounts of data usually need to be processed, and, in a clinical environment, the results of the data analysis need to be available as soon as possible.

In this paper, the potential and limitations of the new technique are further investigated by using brain data instead of prostate data. Brain tumors differ significantly from prostate tumors as they are often very heterogeneous. The data considered in this study were obtained from patients with a glioblastoma tumor. This high-grade type of tumor is very heterogeneous, as different types of tissue can be present, such as high cellularity foci, necrotic areas, infiltration into normal appearing areas, and, sometimes, areas of lower glioma tumor grades, such as grade III. The data were acquired with two-dimensional turbo spectroscopic imaging (2DTSI) sequences (12,13), a technique that allows the acquisition of MRSI data with relatively high spatial resolution in a very short period of time ($\sim$6 min with a turbo factor of 3 and a matrix size of 24×24) compared with that required by standard chemical shift imaging. In this study, CCA is applied to 2DTSI data to assess its ability to detect and classify tissues even when a high degree of heterogeneity is present. The performance of the new technique is also compared with that of ordinary correlation (OCA). The results show that CCA outperforms OCA in terms of robustness and accuracy.

The paper is organized as follows. The Theory section is devoted to the application of CCA to MRSI data. In the Experimental section, the main biochemical features of brain tumors are described. The acquisition environment of the *in vivo* studies and the set up for the simulation studies are also defined. In the Results and Discussion section the results of the simulation and *in vivo* studies are reported and discussed, and the main conclusions are formulated. Finally, in the Appendix, the basic definitions and a reliable and efficient implementation of CCA are described.

THEORY

CCA applied to MRSI data

CCA represents the multivariate variant of OCA, which quantifies the relationship between two random variables x and y by means of the so-called correlation coefficient (14–16). The analytical definition and implementation of CCA are reported in the Appendix. In this section, the application of CCA to MRSI data is described.

During an MRSI experiment, a number of image slices were acquired. Each image can be considered to be a grid of volume elements, called voxels, each characterized by a time-domain signal of length N. The aim is to detect those voxels whose spectra correlate best with model tissue spectra, which are defined *a priori*. If OCA is applied to MRSI data, two univariate variables x and y are needed. Specifically, the x variable consists of the spectrum of the measured signal contained in each voxel, and the y variable consists of the model tissue spectrum. The correlation coefficient between x and y is computed and then assigned to the voxel under investigation. Once each voxel has been processed, a new grid of voxels, of the same size as the original image, is obtained, which contains correlation coefficients instead of MRSI signals. This new grid is called a correlation map. The difference between OCA and CCA consists mainly of a different choice of the variables x and y. To compute the correlation maps, it is possible to exploit the spatial information that characterizes the MRSI dataset by considering, as the variable x, a multivariate vector with components representing the spectrum that characterizes the voxel being considered as well as the spectra contained in the neighboring voxels. The variable y also consists of a multivariate vector. Its components represent the basis function of a signal subspace, which models the characteristic tissue spectrum that we are looking for, and its possible variations, which normally affect realistic MRSI data. In fact, in our application the aim is not limited to measuring the relationship between the measured spectra (x variable) and a model spectrum for each tissue type (possible y variable). In that way, we would not take into account all the possible realistic variations that normally affect spectra characteristic of the same tissue type. Instead, our aim is to take these variations into account as much as possible by considering as the y variable for each tissue type a set of spectra (subspace) rather than a single model spectrum, obtained as all the possible linear combinations of well-defined basis functions.

Once the x and y variables have been defined, CCA is applied voxel by voxel, and the largest canonical coefficient is assigned to the voxel under investigation, so that a correlation map is obtained, as in the OCA case. The nosological image, consisting of a single image in which all the detected tissue types are visualized as defined in Ref. (11), is then obtained by comparing, voxel by voxel, the canonical correlation coefficients characterizing the correlation maps obtained for all the different tissue types. The voxel under investigation is assigned to the tissue with the highest canonical correlation coefficient and labeled by the corresponding tissue color.

Figure 1 schematically shows the CCA approach when processing a 3×3 voxel region containing the real part of the spectrum x_5 along with its neighboring spectra. In this particular spatial model, called the '3×3 model', the variable x contains nine components, namely $x = [x_1,\ldots,x_9]^T$. Several approaches can be adopted to define the spatial model (x variable) and to model a proper signal subspace (y variable). An exhaustive overview is given in Refs (10) and (17). Here, for the sake of clarity, the procedure for defining some examples of spatial and subspace models is described.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/nbm

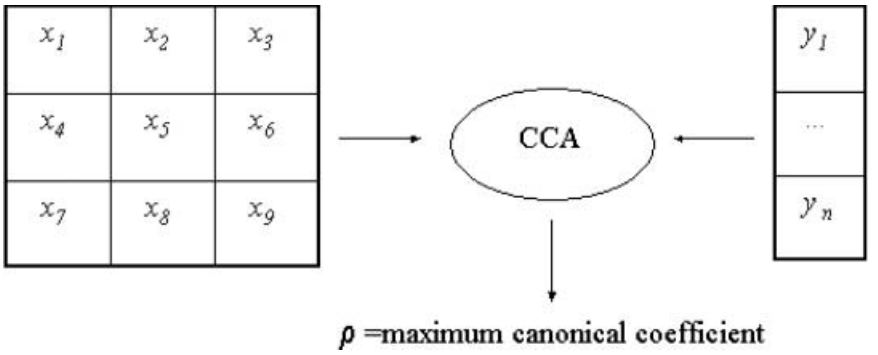


Figure 1. CCA applied to a 3 × 3 region of voxels in the MRS image and a set of *n* spectral basis functions.

Examples of spatial models

The single-voxel model: $x = [x_5]$
The 3×3 model (Fig. 1): $x = [x_1, \ldots, x_9]^T$
The symmetric 3×3 model: $x = [x_5, (x_1 + x_9)/2, (x_2 + x_8)/2, (x_3 + x_7)/2, (x_4 + x_6)/2]^T$
The symmetric 3×3 model without corner voxels: $x = [x_5, (x_2 + x_8)/2, (x_4 + x_6)/2]^T$
Observe that the single-voxel model represents the OCA.

Examples of subspace models

The Taylor subspace model. Compute the model spectrum of the tissue being investigated as the average of characteristic spectra $S_i(\omega_n)$, $i = 1, \ldots, M$, for that tissue type. Denote the model spectrum by $S(\omega_n)$, where ω_n represents the frequency and $n = 1, \ldots, N$.

Set the components of the variable y as:

$$\begin{cases} y_1(n) = S(\omega_n) \\ y_2(n) = \dfrac{S(\omega_{n+1}) - S(\omega_n)}{\Delta\omega} \end{cases} \quad (1)$$

where $\Delta\omega$ is the sampling frequency.

The PCA subspace model. Compute the model spectrum of the considered tissue as the average of characteristic spectra $S_i(\omega_n)$, $i = 1, \ldots, M$, for that tissue type. Denote the model spectrum by $S(\omega_n)$, where ω_n represents the frequency and $n = 1, \ldots, N$.

Arrange the given spectra $S_i(\omega_n)$, $i = 1, \ldots, M$, as rows into a matrix D.

Perform principal component analysis (PCA) on the matrix D and denote the first principal component as 1st PC.

Set the components of the variable y as:

$$\begin{cases} y_1(n) = S(\omega_n) \\ y_2(n) = 1^{st}PC \end{cases} \quad (2)$$

Concerning the choice of the y variable for the single-voxel approach, only one component was considered and set equal to the first component of the Taylor and PCA subspace models.

EXPERIMENTAL

Biochemical features of brain tumors

It is well known that brain tumors can be detected by studying alterations in metabolite concentrations (18–20). Although hundreds of neurochemical compounds are present in the human brain, only a few can be detected by ^{1}H MRS. The most important ones in brain tumor studies are: *N*-acetylaspartate (commonly known as NAA, 2.0 ppm), choline (3.2 ppm), creatine (3.0 ppm), lactate doublet (1.3 ppm), and lipids (0.9 and 1.3 ppm). Usually, in brain tumors, there are reduced concentrations of NAA and an increase in choline. Creatine can also be reduced in some tumors, and lipid resonances can appear very intense in the highest grade tumors. Specifically, NAA is a marker of neuronal/axonal density and viability, and its decrease is associated with loss of neurons or axonal injury. The resonance at 3.2 ppm is assigned to the trimethylammonium compounds, i.e. choline and its derivatives. Free choline is the minor component of this resonance, the major components being phosphorylcholine and glycerophosphorylcholine. In pathological conditions, the contribution of choline increases significantly, as it reflects membrane, myelin and lipid turnover. Creatine plays a crucial role in cell energetics and mainly originates from intracellular metabolite pools of creatine and phosphocreatine. Generally, the choline and creatine peaks show similar intensities in healthy tissue. The creatine peak is the sum of creatine and phosphocreatine with the same chemical shift of the CH_3 group. Lactate is an energy marker. It is usually not detectable in healthy brain tissue, and its concentration increases only under pathological conditions, which are usually characterized by reduced oxidative metabolism or increased glycolysis. Finally, lipids are also not detectable in normal brain tissue as they are bound to macromolecules in membranes and myelin. When a severe pathological condition affects the brain, lipids may become more

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/nbm

mobile, thereby yielding a detectable resonance. Biopsy studies also show that lipids correlate with necrosis, which is a histological characteristic of high-grade tumors (21). All these alterations can be detected by MRSI because, unlike conventional neuroimaging, it provides neurochemical and spatial information. In particular, MRSI enables the simultaneous analysis of different parts of a lesion and the surrounding and contralateral areas. This property is very important when studying brain tumors as they are often heterogeneous and may be characterized by the presence of viable tumor cells, cysts, zones of proliferation and necrosis, edema, and, in the case of highly infiltrative tumors such as gliomas, there may also be contributions from normal brain tissue. Furthermore, tumor growth is not well regulated because variations in cellular metabolism and cell density occur, and as a tumor progresses, it may contain cells of different grades. Finally, the ability to spatially map the biochemical information allows a more accurate pre-surgical diagnosis, such as that based on biopsy, and surgical approach.

Figure 2 shows the real part of two ^{1}H spectra measured in the brain of a patient affected by a glioblastoma tumor. The spectrum on the left-hand side is characteristic of tumor tissue, and that on the right-hand side is typical of healthy tissue.

Acquisition of 2DTSI brain data

The data were acquired at the Clínica Quirón, Servicio de Radiologia, Valencia, Spain, as a part of the routine clinical preoperative MRI and MRS protocol for the brain. All patients signed an informed consent form, in agreement with the guidelines of the university hospital ethical committee. Several days before the spectroscopy study, a full high-resolution MRI study with gadolinium injection was performed. A complete set of images was previously acquired with similar orientation. Moreover, before the acquisition of 2DTSI and in the same session, another set of MRI sequences (T_2, FLAIR turbo fluid-attenuated inversion recovery, and proton density) fully compatible with the previous MRI study in sagittal, coronal, and axial with the same field of view and orientation was acquired, and these axial images were the images that overlapped with the 2DTSI sequence. The MRI information for the different regions comes from both the previous images and the images obtained by spectroscopy, being completely equivalent, at least from the clinical diagnosis point of view. After imaging with turbo spin-echo T_2 and proton density weighting ($TR = 5154$ ms, $TE = 13$ and 115 ms), five MRS experiments were performed by 2DTSI in a 1.5 T Philips Gyroscan Intera NT (Philips Medical Systems, Best, The Netherlands) using a 90°–180°–180° pulse sequence. Signals from surrounding cortical and bone areas were suppressed with 12–16 saturation slabs perpendicular to the acquisition. Shimming and tuning were achieved with an automated procedure before acquisition. The water signal was suppressed with selective excitation. The multivoxel spectroscopic images consist of a grid of 24×24 voxels with a field of view of 230×230 mm and a slice thickness of 20 mm. Each volume unit dimension was $9.6 \times 9.6 \times 20$ mm (1.8 mL). The acquired signals are long-TE signals with $TE = 272$ ms and $TR = 2000$ ms. They consist of 256 complex data points. Data processing also included zero filling to 512 points, removal of the residual water resonance by using HLSVD-PRO (22) [improved variant of HLSVD (23)], and Fourier transformation.

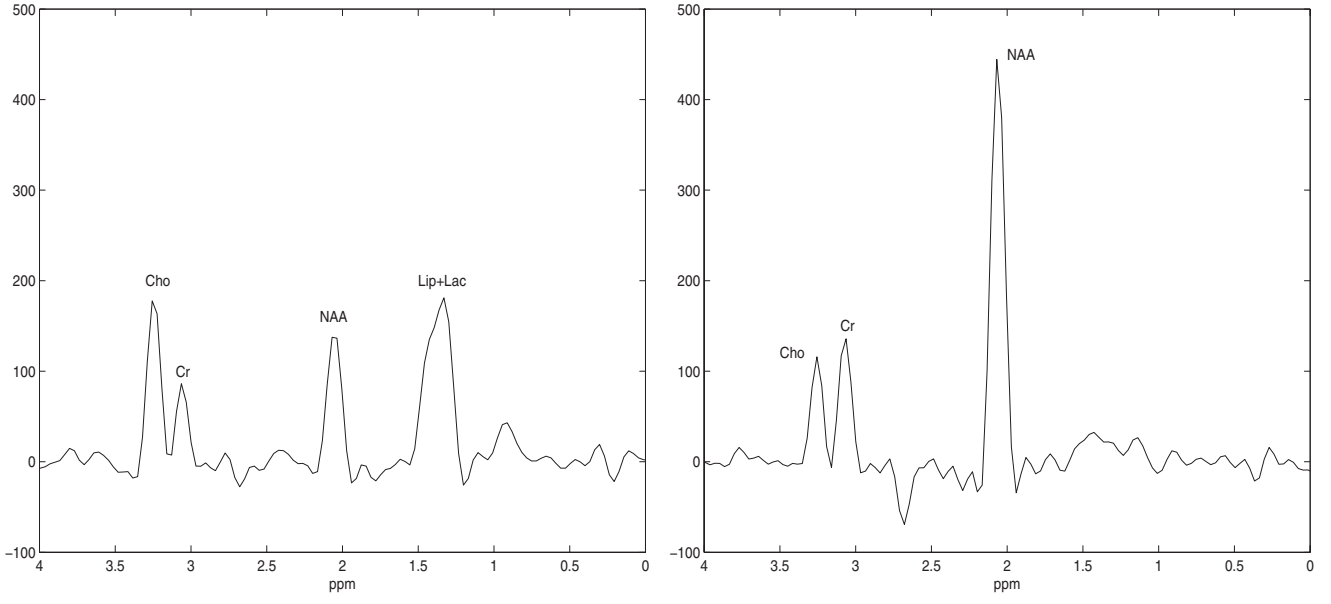


Figure 2. Real part of the ^{1}H spectrum of tumor tissue (left) and healthy tissue (right) at $TE = 272$ ms. Several metabolites are visible: choline (Cho), creatine (Cr), NAA, and lipids overlapping with lactate (Lip + Lac) for tumor tissue.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/nbm

A set of non-suppressed water spectra was also collected in the same sequence acquisition with a lower resolution (a grid of 12×12 voxels with size $19.2 \times 19.2 \times 20$ mm, 7.35 mL) as a reference in order to automatically correct the phase over the 2DTSI grid. The water resonance data in the reference set of spectra were used in the reconstruction for correction of chemical shift and amplitude differences due to magnetic field inhomogeneity. Spectra quality was controlled through the shape, the width at half height, and the intensity of the water resonances measured with the jMRUI software (24,25). These methodological control parameters remained constant for all ^{1}H MRS studies, indicating a stable magnetic field homogeneity (26).

CCA applied to simulated 2DTSI brain data

Figure 3 shows the real part of the spectra of five different tissue signals that are characteristic of tissue found in glioblastoma tumors. They represent necrotic tissue, mixed tissue containing necrotic and tumor cells (tumor + necrosis tissue), tumor tissue, mixed tissue containing tumor and healthy cells (mixed tissue), and healthy tissue. These signals were selected from the *in vivo* datasets and quantified using the jMRUI software (24,25). The parameter estimates were then used to construct the noiseless simulated signals. A simulated 2DTSI dataset was produced, consisting of a 16×16 grid of voxels, in which spectra of five different well-localized regions of tissue were inserted. The grid is shown in Fig. 5 (left image) and the tissue regions are indicated by different colors (see figure legend). To simulate a realistic 2DTSI dataset, variations were introduced on the parameter estimates. More precisely, spectra for each type of tissue were generated, randomly characterized by variations up to 15% for amplitudes, up to twice the given value for the damping factors, and up to a shift of 1.25 Hz for the frequencies. The real parts of the spectra were perturbed by adding Gaussian noise with standard deviation σ. OCA and CCA were then applied to detect the five tissue regions. The results obtained are described in the Results and Discussion section.

CCA applied to *in vivo* 2DTSI brain data

As already mentioned, the data were measured in the brain of five patients affected by glioblastoma. To apply CCA, model spectra of the possible tissue types that characterize brain affected by glioblastoma are needed. Such models are constructed by first selecting spectra from the datasets that, on the basis of spectroscopy, MRI and gadolinium-enhancement features, can be considered characteristic of tissues found in glioblastoma tumors. In the specific case of MRI, it is well accepted that, on T_2-weighted images, regions of edema are hyperintense, and the solid nucleus of the tumor is isointense with the gray matter, whereas necrotic areas are hypointense on T_1-weighted images with heterogeneous contrast enhancement and hyperintense on T_2 images (27). All the selected spectra corresponding to the same tissue type are then averaged to take into account possible variations that affect *in vivo* data. In the examples under

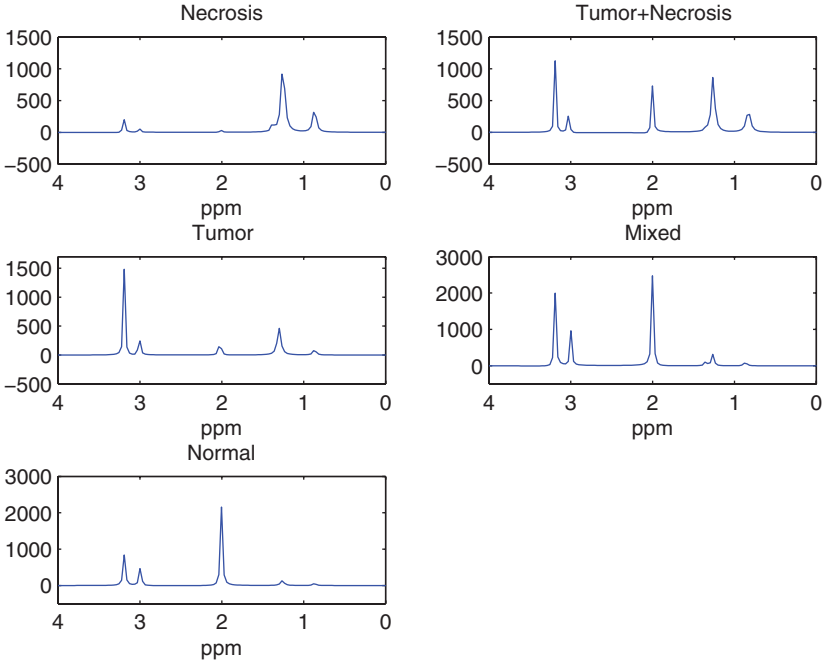


Figure 3. Real part of the spectra of the noiseless simulated 2DTSI data. The mixed tissue is a mixture of tumor and healthy tissue. This figure is available in colour online at www.interscience.wiley.com/journal/nbm

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/nbm

investigation, several spectra could be selected as representative of the following tissue types: tumor, edema, normal, infiltration, necrosis, and axonal damage. In addition, mixed tissue types could be recognized, namely tumor + necrosis tissue and mixed tissue containing tumor and healthy cells. Finally, some lipid spectra due to the presence of the skull were also considered as they sometimes affect the border regions of the processed area of interest. Table 1 shows the number of spectra selected for each example and for each tissue type.

Note that when a dataset is analyzed, e.g. the dataset related to Example 1, the spectra selected for the example under investigation do not contribute to the aforementioned averaging procedure, i.e. all the selected spectra for the same tissue type are averaged except those corresponding to Example 1. Thus, the model spectra used to analyze Example 1 are obtained as the average of the only selected spectra of Examples 2, 3, 4 and 5. Moreover, as only a few spectra could be selected for tumor + necrosis tissue from the given examples, more tumor + necrosis spectra were produced as the average of tumor spectra and necrosis spectra. The same procedure was applied to produce a higher number of mixed tissue spectra, obtained as the average of tumor and healthy spectra.

The spatial models defined in the Theory section were applied, and the best performance was obtained by applying the symmetric 3×3 model without corner voxels, namely $x = [x_5, (x_2 + x_8)/2, (x_4 + x_6)/2]^T$. Concerning the choice of the subspace model, the Taylor model was preferred over the PCA model because only a few characteristic spectra were available for some tissue types and therefore no useful matrix could be constructed on which to perform PCA.

The detection results are illustrated in the Results and Discussion section. Our aim is to compare the performance of OCA, which does not exploit any spatial information characterizing the 2DTSI data, with that of CCA. The nosological images obtained by both methods are displayed along with their correlation maps, which can be used for correct identification of the detected regions. Note that the black border voxels represent voxels that have been excluded from the processing as they are covered by the saturation slabs.

RESULTS AND DISCUSSION

Simulation studies

In the simulation studies, 400 simulation runs were performed for different noise standard deviation values. Before applying the proposed tissue segmentation method, five different y variables were defined, one for each brain tissue type. CCA was then applied between the noisy dataset and the above mentioned y variables, and five correlation maps were obtained. The nosological image was obtained by comparing, voxel by voxel, the five canonical correlation coefficients characterizing the different correlation maps. The considered voxel was then assigned to the tissue with highest canonical coefficient and labeled with the corresponding tissue color.

To compare the performance of CCA with that of OCA, the correlation coefficients between the original and the detected tissue regions were estimated for each simulation run. Figure 4 shows the mean value of these correlation coefficients, over the total number of

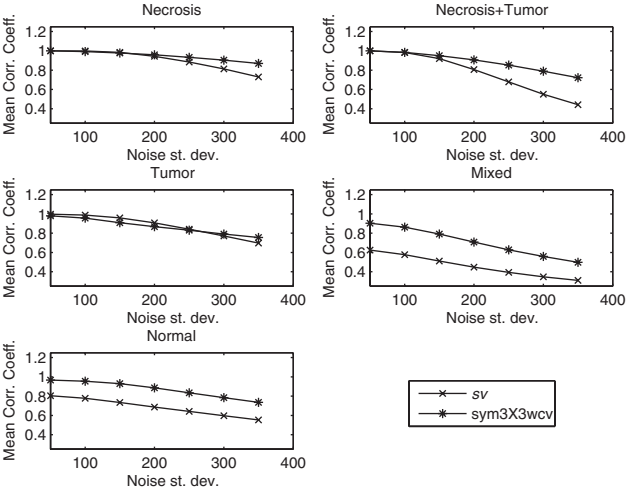


Figure 4. Mean of the correlation coefficients (Corr. Coeff.), over 400 simulation runs, between the detected regions and the original tissue regions as a function of the noise standard deviation (st. dev.), when different spatial models are considered. Subspace model: PCA. The mixed tissue is a mixture of tumor and healthy tissue.

Table 1. Tissue patterns and corresponding number of spectra selected for each dataset

Tissue	Example 1	Example 2	Example 3	Example 4	Example 5
Tumor	5	3	5	6	7
Edema	11	2	—	—	—
Normal	14	9	6	5	5
Infiltration	—	1	—	5	—
Tumor + necrosis	—	—	6	—	—
Necrosis	—	—	6	5	—
Axonal damage	—	—	—	—	5
Mixed	6	—	—	—	—
Lipids	—	—	3	—	—

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/nbm

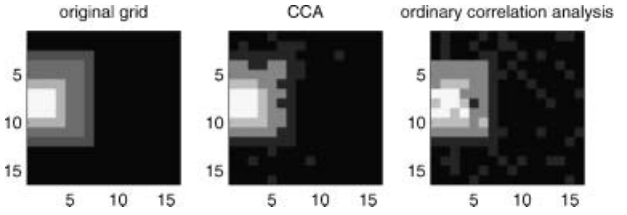


Figure 5. Left, Original grid. Middle, Nosological image obtained by the symmetric 3×3 model without corner voxel for a noise standard deviation equal to 200. Subspace model: PCA. Right, Nosological image obtained by the single-voxel model for a noise standard deviation equal to 200. White, Necrosis; light gray, tumor + necrosis; gray, tumor; dark gray, mixed tissue region; black, healthy tissue.

simulation runs and for each noise level. Different spatial models were applied, and the best performance was obtained by the symmetric 3×3 model without corner voxels (sym3 $\times$ 3wcv). In the figure, we show the results obtained by the latter, when adopting the PCA subspace model, and those obtained by the single-voxel approach (sv), namely OCA. The figure shows that the multivoxel approach significantly outperforms the single-voxel approach, especially when high noise levels are considered. In general, the performance of all models is sensitive to the noise level, to the number of voxels characterizing the original tissue region, and to the shape of the spectra. The quality of the performance for the mixed tissue regions is much poorer than that for the other tissue regions, because little spatial information can be gained and the concentrations of the different metabolites in the spectra are not as discriminating as in the other tissue spectra.

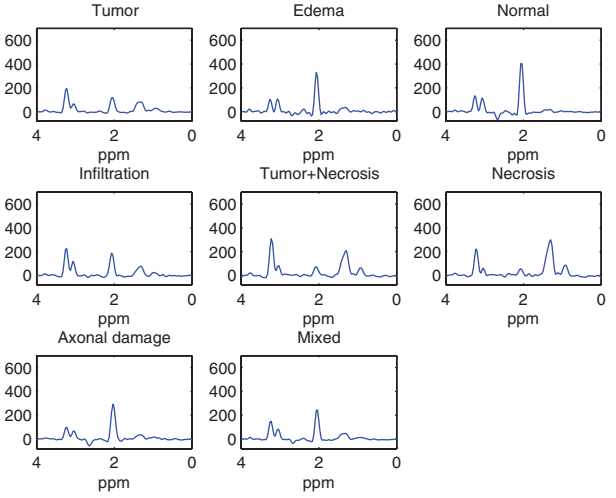


Figure 6. Real part of the tissue model spectra used in the Taylor subspace model of Example 3. This figure is available in colour online at www.interscience.wiley.com/journal/nbm

As a final result, in Fig. 5, the nosological images obtained by the symmetric 3×3 model without corner voxels (middle) and the single-voxel model (right), at high noise level with standard deviation equal to 200, are shown. The images clearly show that the multivoxel approach considerably outperforms the OCA approach. In fact, the tissue regions detected by CCA are well localized, despite their irregular edges and a few mixed voxels detected in the healthy part. On the other hand, the single-voxel approach detects the presence of tumor + necrosis, necrosis and mixed tissue voxels in parts of the grid where spectra of other types of tissue were originally inserted.

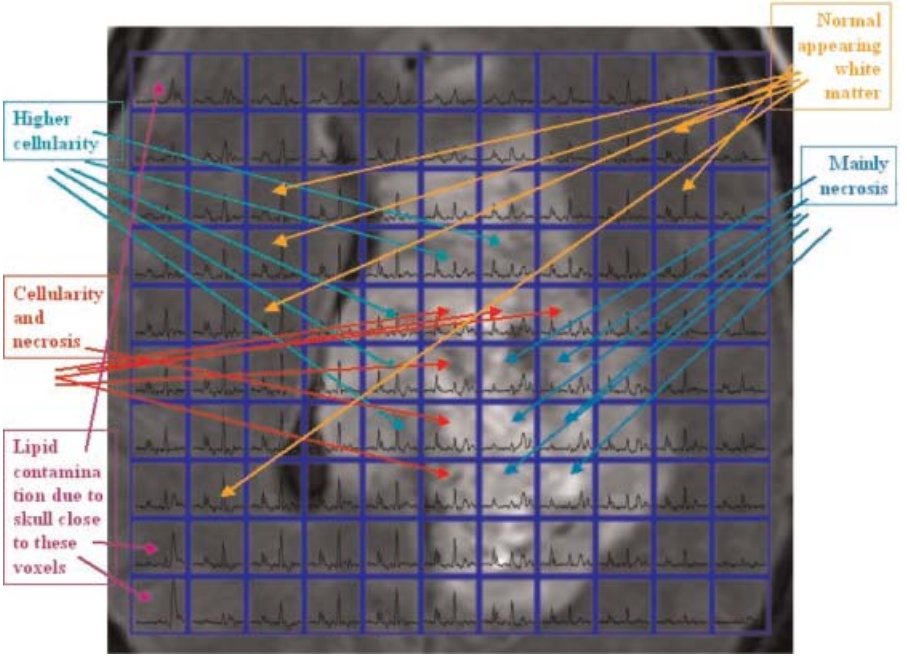


Figure 7. Area of interest of the T_2 image of the glioblastoma of Example 3. This figure is available in colour online at www.interscience.wiley.com/journal/nbm

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/nbm

In vivo studies

Detection results: Example 3. Figure 6 shows the real part of the tissue model spectra, on ppm scale, used to construct the subspace model of Example 3. These spectra clearly show that the tissue types considered are characterized by different contributions of the metabolites of interest. Figure 7 shows a screenshot provided by the Philips scanner in which the lesion can be observed. The slice of interest is the 13th slice, and the figure shows the area of interest of the image and the corresponding spectral map. Some spectra are also pointed out as characteristic of certain tissue types. The area of interest is processed, namely from row 8 to row 17 and from column 7 to column 16. Figure 8 shows the nosological image (top) and all the correlation maps (bottom), corresponding to the different tissue types, obtained by applying CCA (left) and OCA (right). The example under investigation is better characterized by CCA than OCA. In fact, on the basis of spectroscopy, CCA is able to effectively describe the different types of tissue that characterize the lesion. Moreover, OCA fails to detect edema tissue, which is assigned to brain regions where only normal tissue is present.

Concerning the computation times, both CCA and OCA are very fast, requiring less than 1 s to process the area of interest. However, OCA is slightly faster than CCA because the x and y variable only contain one component.

Finally, as a further evaluation of the performance of the two approaches, for each tissue region the average of the canonical correlation coefficients over the number of detected voxels was computed. Table 2 shows that the values obtained by CCA are considerably higher than those obtained by OCA.

Detection results: Example 4

Figure 9 shows a screenshot of the T_2 slice of interest along with the corresponding spectral map. The area of interest, from row 8 to row 17 and from column 7 to column 17, is processed. Figure 10 shows the nosological

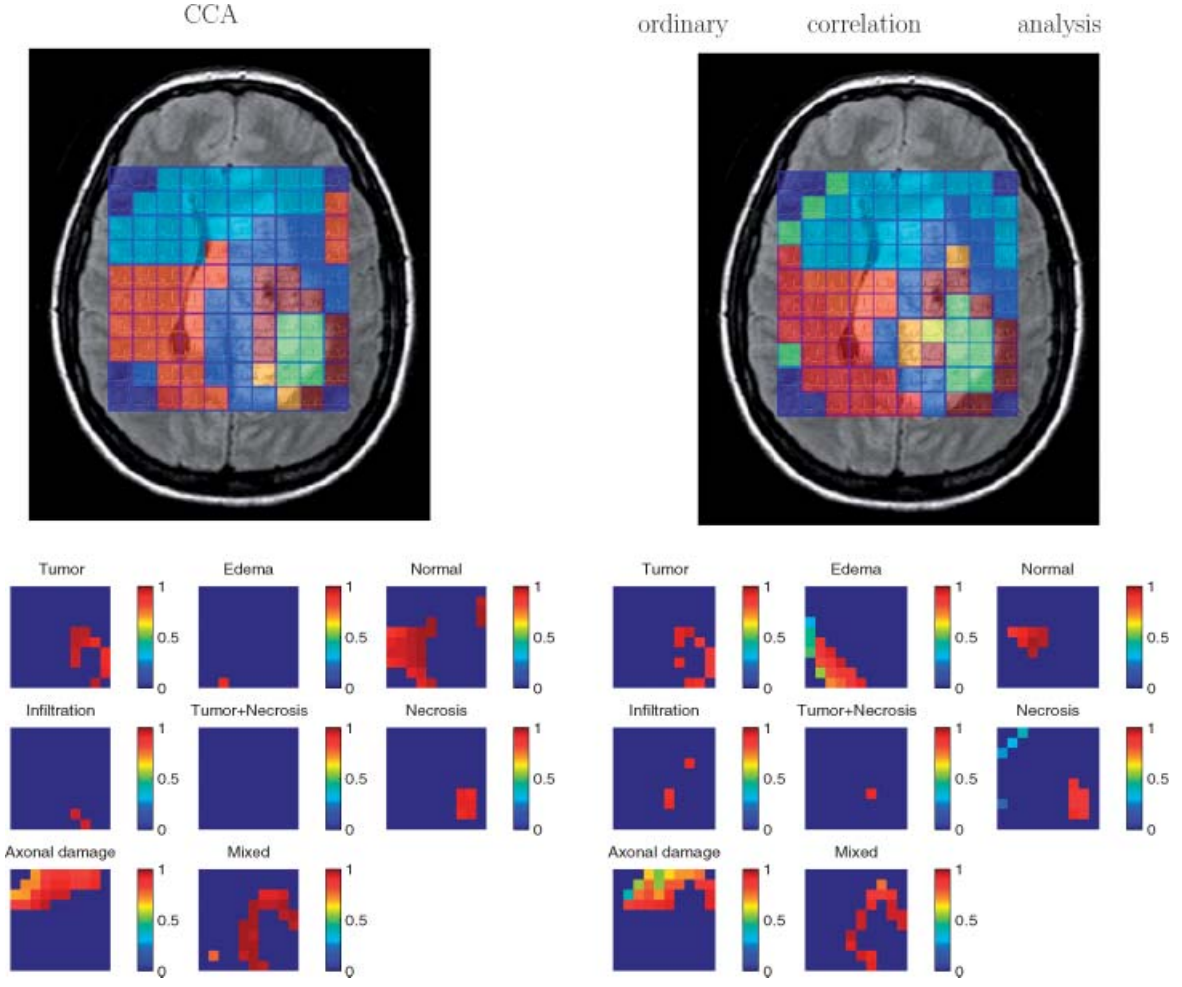


Figure 8. Example 3. Top, Nosological image obtained by CCA (left) and OCA (right). Bottom, Correlation maps corresponding to all the tissue types detected by CCA and OCA. CCA spatial model: symmetric 3×3 without corner voxels. CCA subspace model: Taylor.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/nbm

Table 2. *In vivo* examples: average canonical correlation coefficients, over the voxels belonging to the same tissue type, corresponding to the detected tissue region

	Example 3		Example 4	
Tissue	CCA	OCA	CCA	OCA
Tumor	0.9406	0.9005	0.7672	0.7243
Edema	0.9126	0.7453	ND[a]	0.7004
Normal	0.9517	0.9423	0.9256	0.8304
Infiltration	0.9469	0.9120	0.5484	0.5385
Tumor + necrosis	ND	0.9160	0.8592	0.7965
Necrosis	0.9255	0.6650	0.8706	0.6636
Axonal damage	0.8496	0.7640	ND	0.6056
Mixed	0.9582	0.9138	0.8952	0.8304
Lipids	ND	ND	ND	ND

[a]ND, Not detected (no voxels detected).

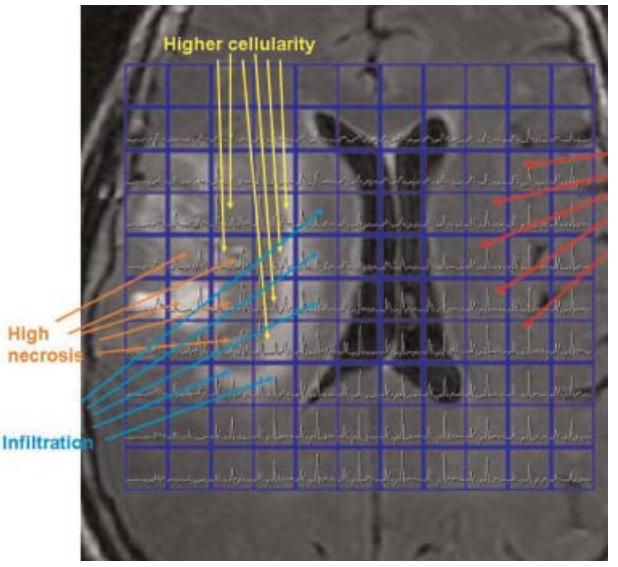


Figure 9. Area of interest of the T_2 image of the glioblastoma of Example 4.

image (top) and all the correlation maps (bottom) corresponding to the different tissue types obtained by applying CCA (left) and OCA (right). Similarly to Example 3, Example 4 is much better characterized by CCA than OCA. In fact, on the basis of spectroscopy, CCA is able to effectively describe the different types of tissue that characterize the lesion. Moreover, the correlation coefficients characterizing the correlation maps provided by OCA are appreciably lower than the canonical coefficients obtained by CCA. It is also of

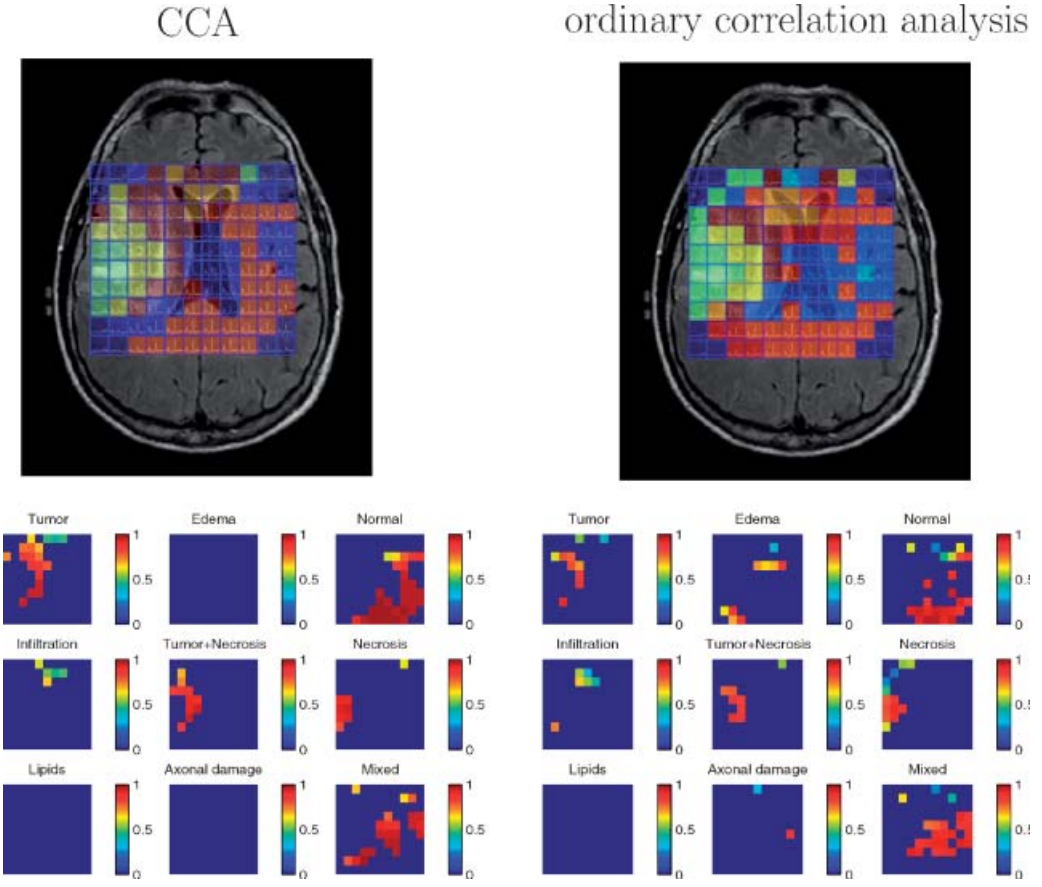


Figure 10. Example 4. Top, Nosological image obtained by CCA (left) and OCA (right). Bottom, Correlation maps corresponding to all the tissue types detected by CCA left) and OCA. CCA spatial model: symmetric 3×3 without corner voxels. CCA subspace model: Taylor.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/nbm

interest to observe that the mixed tissue detected by CCA actually corresponds to the ventricular and periventricular normal appearing areas, which means that both cerebral spinal fluid and tissue are present in the detected mixed tissue voxels.

Table 2 shows the average of the canonical correlation coefficients over the number of detected voxels, for each tissue region. Once again the values provided by CCA are higher than those ones provided by OCA.

Concerning the remaining examples, i.e. Examples 1, 2, and 5, the detection results provided by CCA are always superior to those provided by OCA. In particular, the latter is not able to effectively localize the edema as well as the normal tissue voxels. Generally, it detects broader areas of edema and smaller regions of normal tissue. Moreover, the edema voxels detected are usually located in the wrong position.

CONCLUSIONS

In this paper, the applicability of CCA as a tissue typing technique to 2DTSI brain data is shown. The biochemical features of brain tumors and the advantages of acquiring MRSI data by using 2DTSI sequences are highlighted. The data were obtained from patients affected by glioblastoma tumor, which is one of the most heterogeneous tumors. Our simulation and *in vivo* studies show that CCA is a fast and reliable tissue typing technique as it is able to efficiently characterize the different tissue regions of the brain, even when a high degree of heterogeneity is present. Moreover, the use of spatial information, characterizing 2DTSI data, makes the detection results of CCA better than those obtained by OCA. It is worth stressing that the performance of CCA is sensitive to the shape of the model spectra. The quality of the performance for the mixed tissue regions is much poorer than for the other tissue regions, because the difference in concentration of the different metabolites in the spectra is less pronounced than in the other tissue spectra. Quantifying the degree of uncertainty in the prediction is still an open problem and the subject of ongoing research.

Acknowledgements

This research was supported by Research Council KUL: GOA-AMBioRICS, CoE EF/05/006 Optimization in Engineering, IDO 05/010 EEG-fMRI, several PhD/postdoc and fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects, G.0321.06 (numerical tensor techniques for spectral analysis), research communities (ICCoS, ANMMM); IWT: PhD Grants, Belgian Federal Government: DWTC IUAP V-22 (2002–2006): Dynamical Systems and Control: Computation, Identification & Modelling; EU: BIOPATTERN (contract no. FP6-2002-IST 508803), ETUMOUR (contract no. FP6-2002-LIFESCIHEALTH 503094), HEALTHagents (contract no. FP6-2005-IST 027214). This work was also partially funded by National Grants from the Science and Education Ministry (SAF2004-06297 and SAF2004-20971-E) and Health Ministry (ISCIII2003-G03/185) of the Spanish Government and by ADIRM.

REFERENCES

1. Nelson S. Multivoxel magnetic resonance spectroscopy of brain tumors. *Mol. Cancer Ther.* 2003; **2**(5): 497–507.
2. Leclerc X, Huisman T, Sorensen A. The potential of proton magnetic resonance spectroscopy (^{1}H-MRS) in the diagnosis and management of patients with brain tumors. *Curr. Opin. Oncol.* 2002; **14**: 292–298.
3. Pickett B, Kurhanewicz J, Coakley F, Shinohara K, Fein B, Roach M. Use of MRI and spectroscopy in evaluation of external beam radiotherapy for prostate cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2004; **60**(4): 1047–1055.
4. Sahoo PK, Soltani S, Wong AKC. A survey of thresholding techniques. *Computer Vision, Graphics and Image Processing* 1988; **41**: 233–260.
5. Mangin JF, Frouin V, Bloch I, Regis J, Lopez-Krahe J. From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. *Journal of Mathematical Imaging and Vision* 1995; **5**: 297–318.
6. Coleman GB, Andrews HC. Image segmentation by clustering. *Proc IEEE* 1979; **5**: 773–785.
7. Held K, Kops ER, Krause BJ, Wells WM, Kikinis R, Muller-Gartner HW. Markov random field segmentation of brain MR images. *IEEE Trans Med Imaging* 1997; **16**(6): 878–886.
8. Bezdek JC, Hall LO, Clarke LP. Review of MR image segmentation techniques using pattern recognition. *Medical Physics* 1993; **20**: 1033–1048.
9. Reddick WE, Glass JO, Cook EN, Elkin TD, Deaton RJ. Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks. *IEEE Trans Med Imaging* 1997; **16**: 911–918.
10. Laudadio T, Pels P, De Lathauwer L, Van Hecke P, Van Huffel S. Tissue segmentation and classification of MRSI data using Canonical Correlation Analysis. *Magn. Reson. Med.* 2005; **54**(6): 1519–1529.
11. Szabo De Edelenyi F, Rubin C, Estève F, Grand S, Décorps M, Lefournier V, Le Bas JF, Rémy C. A new approach for analyzing proton magnetic resonance spectroscopic images (^{1}H MRSI) of brain tumors: nosologic images. *Nature Medicine* 2000; **6**(11): 1287–1289.
12. Duyn JH, Moonen CT. Fast proton spectroscopic imaging of human brain using multiple spin-echoes. *Magn. Reson. Med.* 1993; **30**(4): 409–414.
13. Martin AJ, Lui H, Hall WA, Truwit CL. Preliminary assessment of turbo spectroscopic imaging for targeting in brain biopsy. *AJNR Am. J. Neuroradiol.* 2001; **22**: 959–968.
14. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*, 3rd edition. Harpercollins College Publishers: New York, 1996.
15. Zha H. *The Singular Value Decomposition Theory, Algorithms and Applications*. PhD thesis, 1993, Pennsylvania State University, University Park, PA, USA.
16. Golub GH, Van Loan CF. *Matrix Computations*, 3rd ed. The Johns Hopkins Univ. Press: Baltimore, MD, 1996.
17. Friman O. *Adaptive Analysis of Functional MRI Data*. PhD thesis, 2003, Department of Biomedical Engineering, Linköpings University, Sweden.
18. Castillo M, Kwock L, Mukherji SK. Clinical applications of proton MR spectroscopy. *AJNR Am. J. Neuroradiol.* 1996; **17**: 1–15.
19. Castillo M, Kwock L. Proton MR spectroscopy of common brain tumors. *Neuroimaging Clin. N. Am.* 1998; **8**(4): 733–752.
20. Howe FA, Opstad KS. ^{1}H MR spectroscopy of brain tumours and masses. *NMR Biomed.* 2003; **16**(3): 123–131.

Copyright © 2007 John Wiley & Sons, Ltd.

21. Kuesel AC, Sutherland GR, Halliday W, Smith ICP. ^{1}H MRS of high grade astrocytomas: mobile lipid accumulation in necrotic tissue. *NMR Biomed.* 1994; **7**: 149–155.
22. Laudadio T, Mastronardi N, Vanhamme L, Van Hecke P, Van Huffel S. Improved Lanczos algorithms for blackbox MRS data quantitation. *J. Magn. Reson.* 2002; **157**: 292–297.
23. Pijnappel WWF, van den Boogaart A, de Beer R, van Ormondt D. SVD-based quantification of magnetic resonance signals. *J. Magn. Reson.* 1992; **97**: 122–134.
24. Naressi A, Couturier C, Devos JM, Janssen M, Mangeat C, de Beer R, Graveron-Demilly D. Java-based graphical user interface for the MRUI quantitation package. *MAGMA* 2001; **12**: 141–152.
25. Gadea M, Martínez-Bisbal MC, Marti-Bonmatí L, Espert R, Casanova B, Coret F, Celda B. Spectroscopic axonal damage of the right locus coeruleus relates to selective attention impairment in early stage relapsing-remitting multiple sclerosis. *Brain* 2004; **127**(1): 89–98.
26. Casanova B, Martínez-Bisbal MC, Valero C, Celda B, Marti-Bonmatí L, Pascual A, Landente L, Coret F. Evidence of wallerian degeneration in the normal appearing white matter in the early stages of relapsing-remitting multiple sclerosis. A ^{1}H-MRS study. *J. Neurol.* 2003; **250**: 22–28.
27. Edelman RR, Hesselink JR, Zlatkin MB (eds). *MRI: Clinical Magnetic Resonance Imaging*. WB Saunders Company: Philadelphia, 1996; **1**: 533–556.

APPENDIX

CCA: basic definitions and implementation

CCA represents the multivariate variant of OCA, which quantifies the relationship between two random variables x and y by means of the so-called correlation coefficient ρ:

$$\rho = \frac{Cov[x,y]}{\sqrt{V[x]V[y]}} \tag{3}$$

where Cov and V stand for covariance and variance, respectively. The correlation coefficient ρ is a scalar with value between -1 and 1 that measures the degree of linear dependence between x and y. For zero-mean variables, eqn (2) is replaced by:

$$\rho = \frac{E[xy]}{\sqrt{E[x^2]E[y^2]}} \tag{4}$$

where E stands for expected value. 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:

$$\begin{aligned} X &= \omega_{x_1}x_1 + \cdots + \omega_{x_m}x_m = \omega_x^T x \\ Y &= \omega_{y_1}y_1 + \cdots + \omega_{y_n}y_n = \omega_y^T y \end{aligned} \tag{5}$$

CCA computes the linear combination coefficients $\omega_x = [\omega_{x_1},\ldots,\omega_{x_m}]^T$ and $\omega_y = [\omega_{y_1},\ldots,\omega_{y_n}]^T$, called regression weights, so that the correlation between the new variables, X and Y, is maximum. The solution $\omega_x = \omega_y = 0$ is not allowed, and the new variables, X and Y, are called canonical variates. Several implementations of CCA are available in the literature. However, as shown in Ref. (10), the most reliable and fastest implementation is based on the interpretation of CCA in terms of principal angles between linear subspaces (15,16). For further details, the reader is referred to Ref. (10) and references therein. Here, an outline of the aforementioned implementation is provided for the sake of clarity.

Algorithm CCA (CCA by computing principal angles)

Given the 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 $t = 1,\ldots,N$

Step 1. Consider the matrices $\tilde{X}$ and $\tilde{Y}$, defined as follows

$$\begin{aligned} \tilde{X} &= \begin{bmatrix} x_1(1) & \cdots & x_m(1) \\ \vdots & & \vdots \\ x_1(N) & \cdots & x_m(N) \end{bmatrix}, \\ \tilde{Y} &= \begin{bmatrix} y_1(1) & \cdots & y_n(1) \\ \vdots & & \vdots \\ y_1(N) & \cdots & y_n(N) \end{bmatrix} \end{aligned} \tag{6}$$

Step 2. Compute the QR decompositions (16) of $\tilde{X}$ and $\tilde{Y}$

$$\begin{cases} \tilde{X} = Q_{\tilde{X}}R_{\tilde{X}} \\ \tilde{Y} = Q_{\tilde{Y}}R_{\tilde{Y}} \end{cases} \tag{7}$$

where $Q_{\tilde{X}}$ and $Q_{\tilde{Y}}$are orthogonal matrices and $R_{\tilde{X}}$ and $R_{\tilde{Y}}$ are upper triangular matrices.

Step 3. Compute the singular value decomposition (16) of $Q_{\tilde{X}}^T Q_{\tilde{Y}}$

$$Q_{\tilde{X}}^T Q_{\tilde{Y}} = USV^T \tag{8}$$

where S is a diagonal matrix and U and V are orthogonal matrices. The cosines of the principal angles are given by the diagonal elements of S.

Step 4. The diagonal elements of the matrix S represent the so-called canonical coefficients, and the corresponding regression weights are obtained as $\omega_{\tilde{X}} = R_{\tilde{X}}^{-1}U$ and $\omega_{\tilde{Y}} = R_{\tilde{Y}}^{-1}V$.

The computation of the principal angles yields the most robust implementation of CCA, as it is able to provide reliable results even when the matrices $\tilde{X}$ and $\tilde{Y}$ are rank-deficient (16).

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/nbm