Classification of prostate magnetic resonance spectra using Support Vector Machine

Sébastien Parfait, Paul Michael P.M. Walker, Gilles Créhange, Xavier Tizon, Johel Miteran

Classification of prostate magnetic resonance spectra using support vector machine

S. Parfait[a,*], P. M. Walker[a,b], G. Créhange[a,c], X. Tizon[d], J. Mitéran[a,**]

[a] *Laboratoire Electronique Informatique et Image (LE2I), UMR 5158 CNRS, Faculté Mirande 21000 Dijon, France*
[b] *Department of Magnetic Resonance Spectroscopy, University Hospital of Dijon, 2 Boulevard Maréchal de Lattre de Tassigny, 21033 Dijon, France*
[c] *Department of Radiotherapy, Anti-Cancer Centre Georges François Leclerc, rue du Professeur Marion, 21000 Dijon, France*
[d] *Oncodesign SA, 20 rue Jean Mazen B.P. 27627, 21076 Dijon, France*

Abstract

Prostate cancer is the most common cancer in men over 50 years of age and it has been shown that nuclear magnetic resonance spectra are sensitive enough to distinguish normal and cancer. In this paper, we propose a classification technique of spectra from magnetic resonance spectroscopy. We studied automatic classification with and without quantification of metabolite signals. The dataset is composed of 22 patient datasets with a biopsy-proven cancer, from which we extracted 2464 spectra from the whole prostate and of which 1062 were localised in the peripheral zone. The spectra were manually classed into 3 different categories by a spectroscopist with 4 years experience in clinical spectroscopy of prostate cancer: undetermined, healthy and pathologic. We used different preprocessing methods (module, phase correction only, phase correction and baseline correction) as input for Support Vector Machine and for Multilayer Perceptron, and we compared the results with those from the expert. If we class only healthy and pathologic spectra we reach a total error rate of 4.51%. However, if we class all spectra (undetermined, healthy and pathologic) the total error rate rises to 11.49%. We have shown in this paper that the best results are obtained using the

*Corresponding author
**Principal corresponding author
Email addresses: sebastien.parfait@u-bourgogne.fr (S. Parfait), johel.miteran@u-bourgogne.fr (J. Mitéran)

pre-processed spectra without quantification as input for the classifiers and we confirm that Support Vector Machine are more efficient than Multilayer Perceptron in processing high dimensional data.

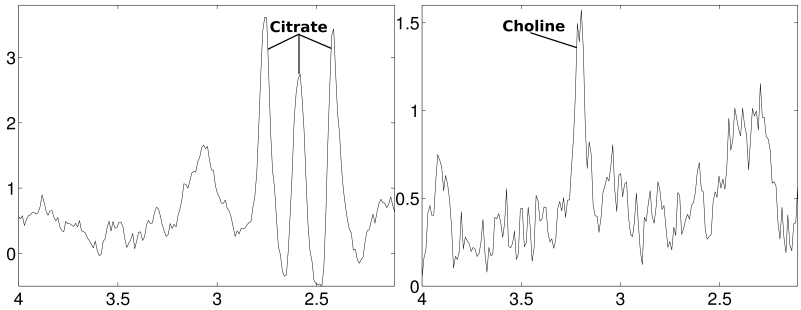
Keywords: SVM, Prostate Cancer, Magnetic resonance spectroscopy

1. Introduction

In men with an elevated level of serum prostate specific antigen (PSA), the diagnosis of prostate cancer usually requires histological confirmation on a biopsy sample that is obtained invasively during transrectal ultrasound imaging. Magnetic Resonance Spectroscopy (MRS) has demonstrated a high level of ability to differenciate prostate tissue from Cancer [1] and it is now feasible without endorectal coil [2, 3]at higher magnetic fields ($\geq$ 3T) one more time for patient comfort. Three-dimensional proton (3D 1H) Magnetic Resonance Spectroscopic Imaging (MRSI) of the entire gland displays relative concentrations of metabolites, particularly citrate and choline (Fig. 1). Normal prostate tissue contains high levels of citrate (higher in the peripheral zone than in the central zone and transition zone) whereas, in the presence of prostate cancer, the citrate level is diminished or undetectable [4] and choline level is apparently increased (Fig. 2). We can notice that prostate cancer is a focal tumour and if the examination is done early enough only a few part of the prostate is occuped by cancer, the rest was healthy.

Even with several years of practice, the decision as to whether a spectrum is healthy or pathologic can be quite difficult and time consuming. Therefore we decided to implement an automatic classification to reduce subjective human invention and to speed up diagnosis.

Although the literature covering prostate cancer depiction (automatically or not) is quite abundant, few papers concern in vivo human MRS with automatic classification. Tiwari et al. produced several articles on this subject [5, 6]. In the more recent paper, they used linear (Principal component analysis, z-score) and non-linear (Locally linear embedding, Graph embedding) dimensionality reduction algorithms on a multicenter MRS database (from ACRIN trial) with a hierarchical clustering to first determine which spectra were in prostate and secondly identify prostate cancer spectra. They reached a sensitivity of 97.66% and a specificity of 98.87% in first experience and sensitivity of 81.36% and a specificity of 64.71% in prostate cancer detection. Although for each prostate examination they had, at their disposal, biopsies



(a) (b)

Figure 1: Typical spectra from (a) normal prostate showing high citrate level (triplet centered around 2.65 ppm) and (b) cancer showing high choline level, at 3.2 ppm (on the ordinate arbitrate unit, on the abscissa chemical shift in ppm).

it was complicated to match biopsy and spectroscopy location, so they chose to define a statistical ground truth.

Matulewicz et al. [7] used MRS data acquired with an endo-rectal coil from patients before radical prostatectomy. A spectroscopist labeled each voxel as healthy or tumour based on established rules. Tumor-classed voxels were confirmed on the basis of histopathologic maps with sextant precision. A Partial Least Square algorithm with Orthogonal Signal Correction filtering has been tested for prediction. They gave prominence to a shorter time of analysis compared with a visual inspection by a spectroscopist, but the accuracy is not comparable to the one of a spectroscopist.

Kelm et al. [8] have compared two approaches : subspace methods on spectral patterns versus quantification of some metabolites in spectra. The datasets were classified with linear and non linear classifiers (Support Vector Machine, Gaussian processes, random forests) and they obtained their best results with a non-linear classifier on magnitude spectra.

Valenzuela et al. [9] proposed an automatic procedure to extract several features from light microscope images and automatically classified them. They reach a classification rate of 85.7% in cancer and 91.4% in benign prostate. This work aimed at help the pathologist in classifying slides.

Jung et al. [10] tried to investigate the accuracy and interobserver variability for MRSI of the prostate. Data were acquired with an endorectal coil from 22 patients. They used histopathologic tumor maps to decide whether

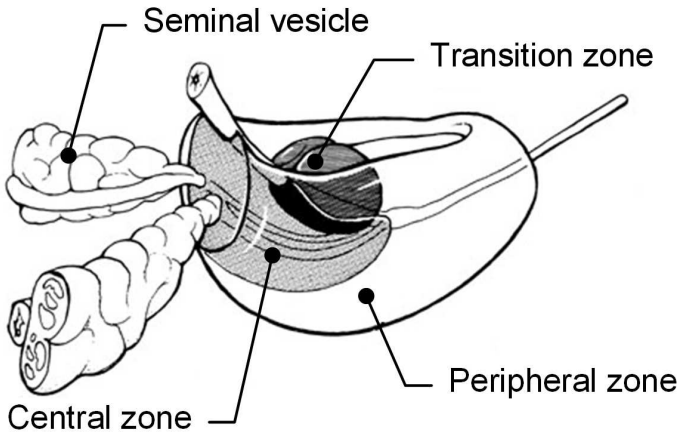


Figure 2: Prostate anatomy. Prostate is composed of 3 zones (central, peripheral and transition).(Extract from Greene DR, Shabsigh R, Scardino P T.Urologic ultrasonography. In: Walsh P C, Retik A B, Stamey T A et al. eds. Campbell's Urology, 6th edn.Philadephia: WB Saunders, 1992; 342-393)

a voxel was benign or malignant. Two spectroscopists scored each voxel on a standardized five-point scale only by analyzing spectra visually. This score was then compared to a histopathologic reference, firstly by considering scores 4 and 5 as pathological, secondly by considering 3 to 5 as pathological and interreader agreement was evaluated by using K-statistics. A good agreement score, close to k=0.80 was reached. When spectra were considered as malignant for a score of 4-5, sensitivity close to 70% and specificity a little under 90%. When a score of 3 to 5 was considered as malignant, sensitivity was over 90% but specificity was only around 70%.

A major study on the classification of pathological tissues based on Magnetic Resonance Spectroscopy (MRS) is the INTERPRET study [11] : Several european MRS centers collaborated to create a consequential database of brain spectra. The aim was then to separate different types of tumor using different classification techniques. Among all data published, the results of Devos et al. [12] show that 1) Support Vector Machine (SVM) are more efficient in high dimensional space than a linear technique 2)L2-normalisation method is very simple and gave good results 3) the best results were obtained with L2-normalised magnitude spectra without baseline correction and simply using peak integration or PCA for dimensionality reduction.

In a clinical context and depending on the chosen acquisition parameters

and the actual size of the gland itself, 3D-MRSI has the potential to provide data from several hundred voxels located entirely within the prostate. The sheer volume of information renders the task of visually or manually classifying data from individual patients virtually impossible. From the standpoint of the spectroscopist and the clinician, a relatively rapid, robust method towards voxel classification is necessary if MRS is to be seriously considered as a clinical tool in prostate cancer detection, localisation and follow-up after treatment.

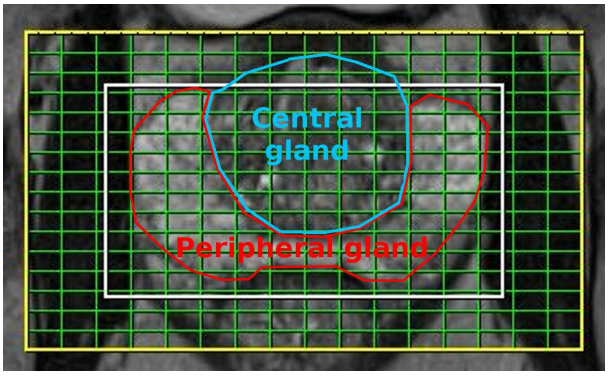
It was our aim, therefore, to develop an automatic classification scheme based on MRS data in order to assist prostate cancer localisation. To this end, we have chosen the support vector machine (SVM), a machine learning technique originating from statistical theory [8] and often used for the classification of images. The SVM has been widely used in pattern recognition applications due to its computational efficiency and good generalization performance even in the case of non linearly separable classes and in case of non-uniform distribution. Moreover, some studies [12, 13, 14] have shown that SVM were generally more efficient on whole data than data having undergone dimension reduction techniques. We chose to compare this method with a Multilayer Perceptron (MLP) that have already proved it qualities in medical applications [15, 16].

Thus, this paper presents the application of SVM and MLP to the automatic classification and localisation of prostate cancer, comparing several pre-processing steps of spectra, with and without estimations of metabolite signals. The study subjects, acquisition techniques and the dataset are presented in section 2. In section 3, the preprocessing methods studied for feature extraction are described, and a brief overview of SVM is presented. Section 4 contains experiments and results while a conclusion is presented in section 5.

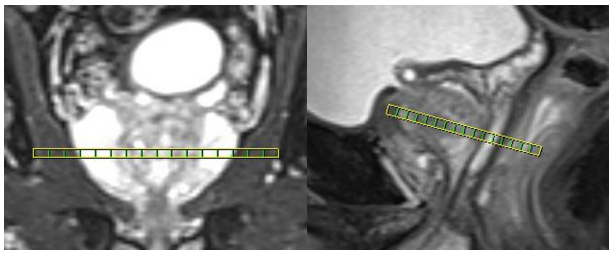
2. Acquisition techniques and materials

2.1. Study subjects

All data were acquired in the Department of Magnetic Resonance Spectroscopy of our University Hospital. This is a cross-sectional study on 22 patients (56-81 years old, mean age=67.9 $\pm$ 6.7, median=67.5) with a proven prostate cancer (PSA level = 4-61 ng/mL mean=11.3 $\pm$ 11.9 median=8.3 and positive biopsies). They all received a complete pelvis Magnetic Resonance



(a) Transverse



(b) coronal (c) sagittal

Figure 3: Prostate T_2 MRI with spectroscopic grid. Central and peripheral zone are surrounded. The white box is the VOI where tuning of the device is optimized. On sagittal and coronal slices we can see the projection of the spectroscopic grid

Imaging (MRI) and MRS examination. Every patient underwent transrectal ultrasound-guided biopsy at least 6 weeks before MRI.

2.2. MRI and MRS imaging techniques

Images acquired using a three-dimensional T_2-weighted fast spin-echo (TR/TE/ETL: 3000 ms/143 ms/109, slice thickness: 1.5 mm) sequence and orientated perpendicular to the prostate peripheral zone (PZ) - rectal wall axis were used to position the 3D spectroscopic grid. The nominal matrix and field-of-view (FOV) of the 3D T_2-weighted fast spin-echo images were 320 x 256 and 280 x 240 mm^2, respectively, thereby affording sub-millimetric pixel resolution within the imaging plane. Once the grid positioned correctly on the prostate (figure 3), 3D ^{1}H MRSI data were acquired by using a water and lipid suppressed double-spin-echo point-resolved spectroscopic (PRESS)

sequence optimized for the quantitative detection of both choline and citrate. Water and lipid suppression were achieved by using a dual-band spectral spatial pulse technique [17]. Data sets were acquired as 16 x 12 x 8 (interpolated to 16 x 16 x 8) phase-encoded spectral arrays (2048 voxels) with a nominal spectral resolution of 0.28-0.36 cm^3 before interpolation, TR/TE: 720/140 ms and a 13-minute acquisition time. A spectral bandwidth of 1250 Hz was used over 512 complex data points.

2.3. Dataset

From the 22 examinations, we only retained those voxels located entirely within the prostate and more specifically within the peripheral zone, where the majority of prostate cancer lesions are found. Finally, 1062 spectra from the peripheral zone were retained for further analysis.

Based on the spectra and the results from the MRI examination, a spectroscopist with more than 4 years experience in prostate spectroscopy visually classed the spectra into three categories: Undetermined, Healthy and Pathologic (Fig. 4) according to decision rules based on choline, creatine, polyamines and citrate level [18]. The notion of an "Undetermined" spectrum had to be taken into account in the context of an automatic classification, because a number of spectra, from the spectroscopist point of view, did not contain any significant information neither on healthy tissue nor on cancer. This is mainly due to low metabolic concentrations within the voxel. Such would be the case, for example, in the presence of prostatitis, an inflammatory condition of the gland. The final number of spectra available for each class is presented in table 1. Note that the proportion of spectra of each class is expected to be representative of prior probabilities of these classes.

Undetermined	Healthy	Pathologic	Total
286	636	140	1062

Table 1: Number of spectra per class

3. Preprocessing and classification methods

3.1. Preprocessing of nuclear magnetic resonance spectra

The Nuclear Magnetic Resonance (NMR) device provides a signal called the free induction decay (FID). To obtain the spectrum we have to initially perform the following simple operations using Matlab :

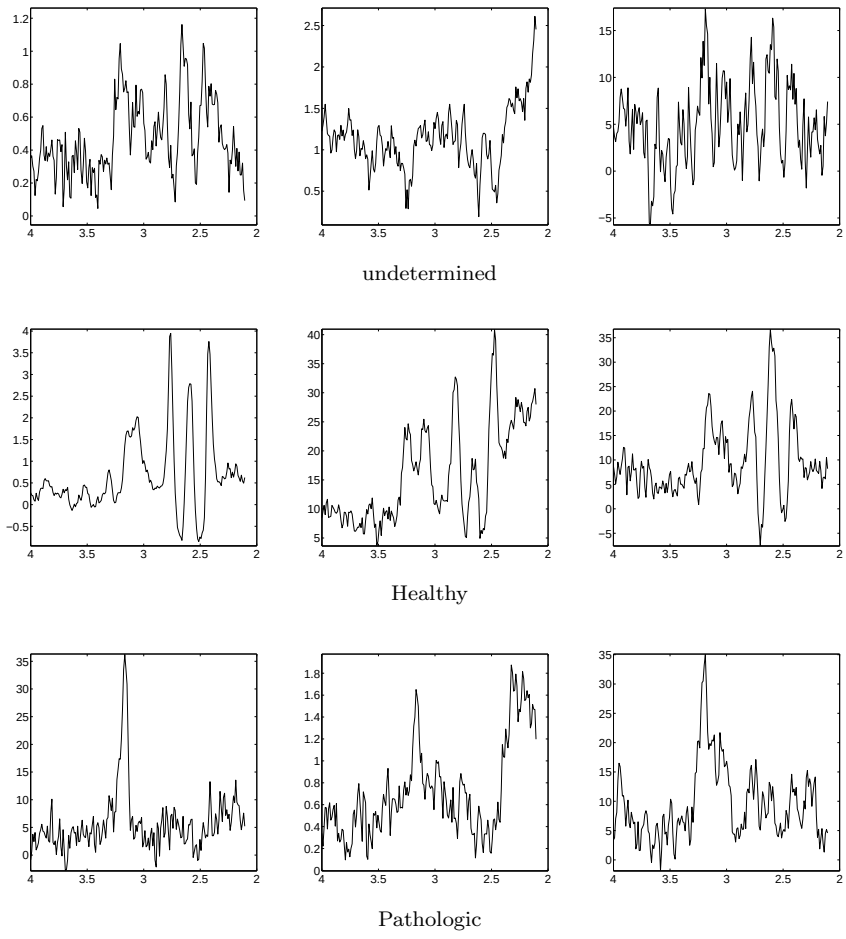


Figure 4: Samples of 3 undetermined, 3 healthy and 3 pathologic spectra from prostate (on the ordinate arbitrate unit, on the abscissa chemical shift in ppm).

- Zero-filling of the free induction decays obtained from the NMR device to double the number of complex points from 512 to 1024,
- FFT to obtain spectra,
- normalization of spectra with respect to the volume of the voxel.

However, it is necessary to introduce more sophisticated pre-processing steps in order to analyse spectra efficiently. The NMR spectrum is by nature a complex signal, with real and imaginary parts. Although, only the real part is usually analysed by the spectroscopist, this signal can be difficult to process notably due to the presence of interfering background signals leading to baseline distortion and to phase correction problems.

The discrete Fourier transform vectors are written as functions of frequency ν. Thus, the discrete spectrum data points $Y(\nu_k)$ are modeled as:

$$\hat{Y}(\nu_k) = S(\nu_k).\exp(-i\Phi(\nu_k)) \quad with \quad S(\nu_k) = \rho.M(\nu_k) + B(\nu_k) + \epsilon(\nu_k) \quad (1)$$

where $\exp(-i\Phi(\nu_k))$ is a phase correction factor, $M(\nu_k)$ is the signal of interest, ρ is a normalization factor, $B(\nu_k)$ the baseline component and $\epsilon(\nu_k)$ is an additive noise. The aim of the preprocessing steps is to estimate the signal $M(\nu_k)$ from the signal $\hat{Y}(\nu_k)$. We have then to correct the phase, remove the noise, remove the baseline and determine the normalization factor ρ.

As our acquisition techniques present a good signal to noise ratio, the noise is just treated by a simple Gaussian filter.

3.1.1. Phase correction

Phase correction aims to find the coefficient $\Phi(\nu_k)$ from equation 1. We have employed the algorithm ACME, proposed by Chen et al. in [19] to solve the problem of phase correction. This method is based on the minimization of an entropy function to find the best spectral form. Indeed, zero-order and first-order phase corrections are required for Fourier transformed NMR spectra :

- A zero-order phase misadjustment arises from the phase difference between the reference phase and the receiver detector phase. This correction is frequency-independent,
- A first-order phase misadjustment arises from the time delay between excitation and detection, flip-angle variation across the spectrum and

phase shifts from the filter employed to reduce noise outside the spectral bandwidth [1-3]. This correction is frequency-dependent.

The phase corrected spectra is

$$S(\nu_k) = \hat{Y}(\nu_k).\exp(i\Phi(\nu_k))$$

If n is the number of samples within the spectra, $\Phi(\nu_k)$ is the total phase angle correction given by :

$$\Phi(\nu_k) = \varphi_0 + \frac{\nu_k}{n}.\varphi_1$$

with φ_0 and φ_1 the value of the zero order and first order phase correction.

The ACME algorithm tries to find the best parameters φ_0 and φ_1 that minimize an objective function E which possesses a Shannon type information entropy measurement and which is the addition of two terms. The first, denoted H is built from the normalized derivative of the real part of the signal. The second, denoted P is a penalty function to ensure nonnegative bands in the spectra.

If we denote $R_{\hat{Y}}(\nu_k)$ and $I_{\hat{Y}}(\nu_k)$ the real and the imaginary part of $\hat{Y}(\nu_k)$, H is given by

$$H = \sum_{i=1}^{n-3} -D_i.ln(D_i)$$

with

$$D_i = \frac{dR_{\hat{Y}}(\nu_i)}{\sum_{k=1}^{n-3} dR_{\hat{Y}}(\nu_k)}$$

and

$$dR_{\hat{Y}}(\nu_i) = \left| \frac{R_{\hat{Y}}(\nu_i + 2) - R_{\hat{Y}}(\nu_i)}{2} \right|$$

P, the penalty function is given by :

$$P = \gamma \sum_{i=1}^{n} F(R_{\hat{Y}}(\nu_i)).(R_{\hat{Y}}(\nu_i))^2$$

where γ is a penalty factor which should be set appropriately to balance the contributions of the entropy and penalty parts. The function F is defined as

$$F(x) = \begin{cases} 0, x \geqslant 0 \\ 1, x < 0 \end{cases}$$

The function $E = H + P$ is minimized by the simplex method. Phase correction on a small set of spectra allow us to choose a value of 1000 for γ because it permit to obtain spectra near hand made phase correction.

3.1.2. Baseline correction

The presence of macromolecules or lipids within the acquisition voxel may give rise to very wide peaks in the MR spectra. These peaks are often of little interest in the context of prostate cancer and may distort the spectrum by artificially elevating the baseline of the signal. Several methods exist to detect and correct this problem: some very simple and some more complex using the wavelet transform or a Bayesian framework [20, 21, 22, 23]. We compared two methods, the first one based on the work of Lieber et al. [21] and the second one described in [12]. The method of Devos consist in multiplying the NMR temporal signal with an exponential decreasing function before applying FFT to obtain a baseline. It is possible to tune a parameter of the exponential function (β) in order to optimize the baseline estimation. The method of Lieber is based on a low-pass filter (convolution with a Gaussian function) applied iteratively as follows:

- Initialise $S'(\nu_k) = S(\nu_k)$
- For each iteration

  Compute $S''(\nu_k) = (S' * G)(\nu_k)$ where G is a Gaussian function

  If $S(\nu_k) \geq S''(\nu_k)$ then $S'(\nu_k) = S''(\nu_k)$
- Output : $B(\nu_k) = S''(\nu_k)$

We compared the two algorithms using a set of 100 simulated spectra built using the Gaussian decomposition of real spectra, added to a simulated baseline composed of two Gaussian functions. For each method, we computed the quadratic error between the theoretical baseline and the estimated baseline. For the Lieber method, we tuned the number of iteration from 1 to 1000, and for the Devos algorithm, we tuned the β parameter from 0 to 10. Results are depicted Fig.5 a and b. Regarding the quadratic error,

the method of Lieber outperforms the method of Devos. We choose then this method and we fixed T=200 iterations for next experiments (after 200 iterations, the error rate decrease is no longer significant). An example of baseline estimation using the algorithm of Lieber is presented in Fig.5 c and d.

3.1.3. Normalisation

Due to different machine settings for different patients (position and dimension of the VOI, magnetic and radiofrequency field homogeneities ...), the detected NMR signal may be very different from one patient to another. Signal normalisation is therefore required for the estimation of the parameter ρ of the eq. 1. As the human prostate contains about 70% of NMR-visible water, a spectrum without prior water suppression will contain a largely dominant water peak. Two techniques are usually used to normalize spectra [12] :

- estimation of the concentration of water and use of this value as a normalisation factor,
- T_2-normalisation : $\rho = \sqrt{\sum_k (S(\nu_k))^2}$

As in the Devos study [12], we observed that the most efficient normalisation method for classification problem is T_2-normalization. We used this technique in the next experiments.

3.1.4. Relative concentration estimation of metabolites

A common way to analyze MRS data is to estimate the relative concentration of metabolites detected in the spectrum. LCModel is a software initially developed for brain spectra processing, by S. Provencher [24] which analyzes the in vivo spectrum as a linear combination of a basis set of complete model spectra of metabolite solutions in vitro (eq. 2).

$$\hat{Y}(\nu_k) = \exp(-i\Phi(\nu_k)) \left[\sum_{j=1}^{N_B} \beta_j B_j(\nu_k) + \sum_{l=1}^{N_M} C_l \sum_{n=-N_s}^{N_s} \Gamma_n . M_l(\nu_{k-n}; \xi_l, \delta_l) \right] \quad (2)$$

with the constraints

$$C_l \geqslant 0, \quad \xi_l \geqslant 0, \quad \sum_{n=-N_S}^{N_S} \Gamma_n = 1$$

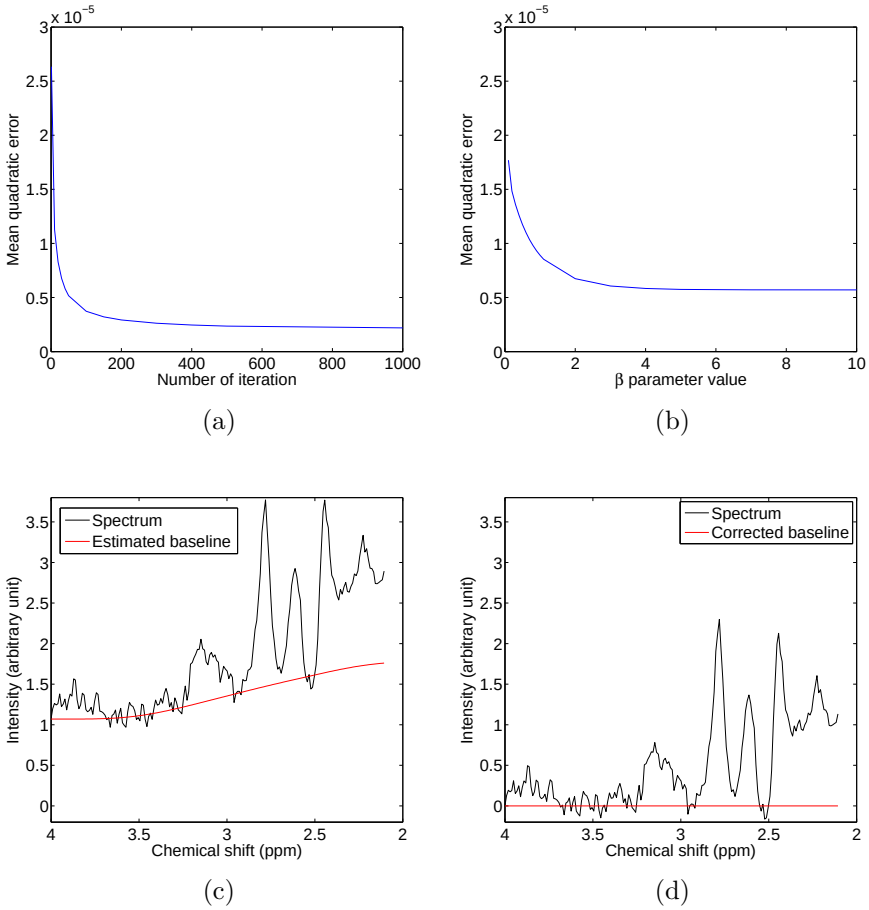


Figure 5: Quadratic error between simulated and estimated baseline with Lieber-based baseline correction (a) and Devos baseline correction (b) Spectrum from prostate before (c) and after Lieber-based baseline correction (d).

and where

- The baseline is represented by $\sum_{j=1}^{N_B} \beta_j B_j(\nu_k)$, a sum of N_B cubic B-splines, $B_j(\upsilon)$, with equally spaced knots.
- C_l are the concentrations of each metabolite.
- $M_l(\nu_{k-n}; \xi_l, \delta_l)$ modelizes the spectrum of the metabolite l, ξ_l is a parameter to account the shorter T_2 echo times in vivo and δ_l is a shifting parameters to account small errors in referencing spectra.
- Γ_n is a normalization factor.

The algorithm tries to minimize the difference between this model and the real spectra (more information are available in [24]).

Since LCModel was first developed to estimate brain spectra, the software has been modified to cater for prostate spectra - the basis sets were simulated rather than generated from in vitro data (see Fig. 6). The estimated concentration values can be used as an input vector of a classification method, or simply thresholded such as in [25]. The results will be compared with the human based ground truth established in 2.3.

3.2. Classification

3.2.1. Support Vector Machine

SVM is a universal learning machine developed by Vladimir Vapnik [26] in 1979. A review of the basic principles follows, considering a 2-class problem (whatever the number of classes, it can be reduced, by a "one-against-others" method, to a 2-class problem).

The SVM performs a mapping of the input vectors from the input space (initial feature space) R^d into a high dimensional feature space Q; the mapping is determined by a kernel function K. It finds a linear decision rule in the feature space Q in the form of an optimal separating boundary, which leaves the widest margin between the decision boundary and the input vector mapped into Q. This boundary is found by solving the following constrained quadratic programming problem:

Maximize:

$$W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

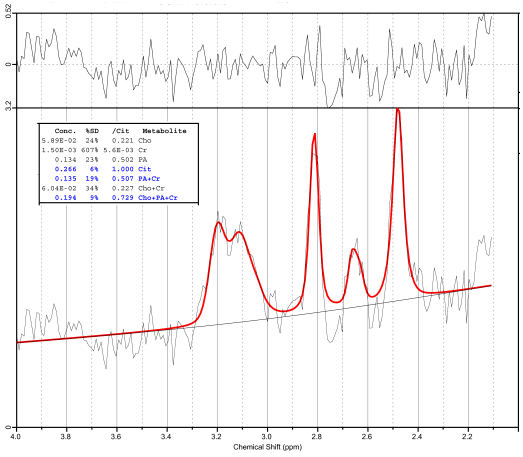


Figure 6: Example of spectrum processed by LCmodel. Initial spectrum is in black in the lower part. The simulated spectrum (red) and the baseline (black) are overpainted. In upper part there is the residual. In the box are the concentrations estimated by LCModel with a confidence index (Cramer Rao lower bounds) and the ratio with citrate.

Under the constraints

$$\sum_{i=1}^{n} \alpha_i y_i = 0$$

and $0 \leq \alpha_i \leq T$ for i=1, 2, ..., n where $x_i \in R_d$ are the training sample set vectors, and $y_i \in \{-1, +1\}$ the corresponding class label. T is a constant needed for non separable classes. $K(u, v)$ is an inner product in the feature space Q which may be defined as a kernel function in the input space. The condition required is that the kernel $K(u, v)$ be a symmetric function which satisfies the following general positive constraint:

$$\int_{R_d} K(u, v)g(u)g(v)\, du\, dv > 0$$

Which is valid for all $g \neq 0$ for which $\int g^2(u)du < \infty$ (Mercer's theorem).

The choice of the kernel $K(u, v)$ determines the structure of the feature space Q. A kernel that satisfies eq. 3.2.1 may be presented in the form:

$$K(u, v) = \sum_{k} a_k \Theta_k(u)\Theta_k(v)$$

Where a_k are positive scalars and the functions Θ_k represent a basis in the space Q. A Radial Basis Function SVM (RBF) was used in this study:

$$K(x, y) = \exp\left(\frac{-\|x - y\|^2}{2\sigma^2}\right)$$

The separating plane is constructed from those input vectors, for which $\alpha_i \neq 0$. These vectors $s_i, i = 1, ..., N_v$ are called *support vectors* and reside on the boundary margin. Mapping the separating plane back into the input space R_d, gives a separating surface which forms the following nonlinear decision rules:

$$C(x) = Sgn\left(\sum_{i=1}^{N_v} y_i \alpha_i \cdot K(s_i, x) + b\right)$$

All results with SVM were obtained with a home made software based on the LIBSVM library [27].

3.2.2. MultiLayer Perceptron

Perceptron is a family of Neural Networks, imagined by Rosenblatt [28]. We used a commercial software : NeuroSolutions version 5.07 developed by NeuroDimension, Inc.. Perceptron were composed with one hidden layer. The number of neurals was automatically tuned by a genetic algorithm to obtain the best learning rates.

4. Experiments and results

4.1. Experiments

We aimed to find the best combination of signal preprocessing and classification methods, and so for each preprocessing method, (raw spectra, spectra after phase correction by ACME, spectra after phase correction by ACME and baseline correction) we looked at the classification total error rates with each classification method (SVM, MLP).

As described in the section 3.1.4, LCModel is used for the estimation of relative concentration of metabolites (choline, polyamine, creatine, citrate). The preprocessing steps are included in the software (phase and baseline correction, frequency shift). From these relative concentrations, two experiments were carried out:

- Bayes classification rule, based on the ratio choline/citrate. This ratio was used in a previous study [25] for the evaluation of cancer in the prostate.

- SVM based classification, using the four concentration values as an input vector.

We also evaluated the performance of classification methods without estimating the relative concentration of metabolites. For these experiments, the spectra are used directly as an input vector of the SVM classifier after the following preprocessing (normalized by a T_2-Normalization described in 3.1.3):

- use of the module of the spectra,

- correction of the spectra by the ACME algorithm restricted to the chemical shift range 2.0 - 4.0 ppm, in which the main metabolites are present. That represent 192 data points. Several experiments carried out using larger ranges did not allow to obtain a lower classification error.

- correction of the spectra by the ACME algorithm and baseline correction.

The preprocessing result, in the case of a restricted range of analysis, is a vector of dimension 192 used as an input of the classification method. Indeed, the potential advantage of SVM is to deal correctly with high dimensional input vectors with a good generalization power. For each SVM based experiment, the error rate is measured using a 5-fold crossvalidation. We used a RBF Kernel, which is often used in the literature because it depends on only one parameter, which can be easily be tuned to obtain an optimum classification error rate and usually gives the best results [12, 29, 13, 30].

We first evaluated the ability of the methods to separate healthy and pathological spectra. Secondly, we evaluated the methods using all the classes, including "Undetermined" spectra.

4.2. Results

4.2.1. Two class experiments

The results of the two class experiments (misclassification rate, sensitivity and specificity) are presented in the table 2. The first row of this

	Error rate(%)			Sensitivity (%)	Specificity (%)
	Healthy	Pathologic	Total		
Relative concentrations					
LCModel + **Bayes**	9.15	33.33	12.89	66.42	90.88
LCModel + **SVM**	1.73	37.14	8.12	62.85	98.27
LCModel + **MLP**	1.60	42.90	8.97	57.14	98.42
Whole spectra with **SVM**					
Module	1.26	20.71	4.77	79.28	98.74
ACME	2.36	15.71	4.77	84.28	97.64
ACME + Baseline Correction	1.89	16.43	4.51	83.57	98.11
Whole spectra with **MLP**					
Module	0.80	85.70	3.23	14.28	99.21
ACME	3.30	20.71	6.44	79.28	96.69
ACME+ Baseline Correction	2.99	19.29	5.93	80.71	97.01

Table 2: Healthy vs Pathologic misclassification rate, sensitivity and specificity (%)

table presents the results of optimum classification obtained applying a simple Bayes based classification rule (single threshold applied to the ratio choline/citrate). The high total error rate (12.89%) shows that this ratio does not separate the classes Healthy and Pathologic efficiently. The second and the third rows present the result of classification using the relative concentration estimated by LCModel as an input vector of the SVM based classification and the MLP. The total error rate is similar (between 8% and 9%) for both classification methods. It is lower (8.12%) than for the previous case, but higher than for all other experiments. This is partially due to the fact that LCModel fails to evaluate correctly the concentration of some metabolites. Indeed, the software was initially developed for the spectroscopic analysis of the brain, for which it is known as a very efficient method. Thus, although the detection of the limited number of prostate metabolites was satisfactory when high concentrations of citrate were present, serious errors were observed in pathological tissues when citrate is severely depleted. In such cases, the minimization function did not converge to the right solution, a large frequency shift was observed, and as a result, the dominating choline peak was often confused with other metabolites. In addition, residual lipid resonances were erroneously quantified as citrate thereby overestimating relative citrate concentration. Errors in peak detection also hindered correct baseline detection (Fig. 7d): when LCModel did not cater for the presence of the residual lipids (ca 2.0-2.4 ppm), the baseline was artificially high, thereby underestimating citrate concentration. Although there is a total error rate of 3.23% using module as input of MLP, this result must be compared with the

Pathologic error rate of 85.70%: we want to detect pathologic spectrum and so this result is not particularly interesting (sensitivity is only 14.28%). So the best result (error rate : 4.51%) is obtained using a correction of the spectra by the ACME algorithm and baseline correction combined with SVM. In that case, the sensitivity is 83.57% and the specificity is 98.11%. Even if the difference between the use of the module and the ACME+baseline correction is not really significant, when comparing the global classification error, nevertheless the ACME+baseline based method is more efficient (sensitivity is higher) when considering the misclassification rate of pathologic spectra (error rate of 16.43% vs 20,71%). SVM outperforms here the MLP, regarding the sensitivity and the global misclassification rate.

4.2.2. Three class experiments

	Undetermined	Healthy	Pathologic	Total
Relative concentrations				
LCModel + **SVM**	52.45	7.08	47.14	24.58
LCModel + MLP	57.34	9.75	52.86	28.25
Whole spectra with **SVM**				
Module	20.28	6.76	27.86	13.18
ACME	23.42	7.39	30.00	14.69
ACME + Baseline correction	17.83	5.66	25.00	11.49
Whole spectra with **MLP**				
Module	24.13	6.45	40.00	15.63
ACME	27.62	9.28	24.29	16.20
ACME + Baseline correction	19.93	9.12	27.86	14.50

Table 3: 3 Classes experiments: misclassification rates(%)

		Predicted class		
		Undetermined	Healthy	Pathologic
Actual class	Undetermined	82.17	11.54	6.29
	Healthy	4.72	94.34	0.94
	Pathologic	14.29	10.71	75.00

Table 4: Confusion matrix for SVM with ACME/Baseline Correction

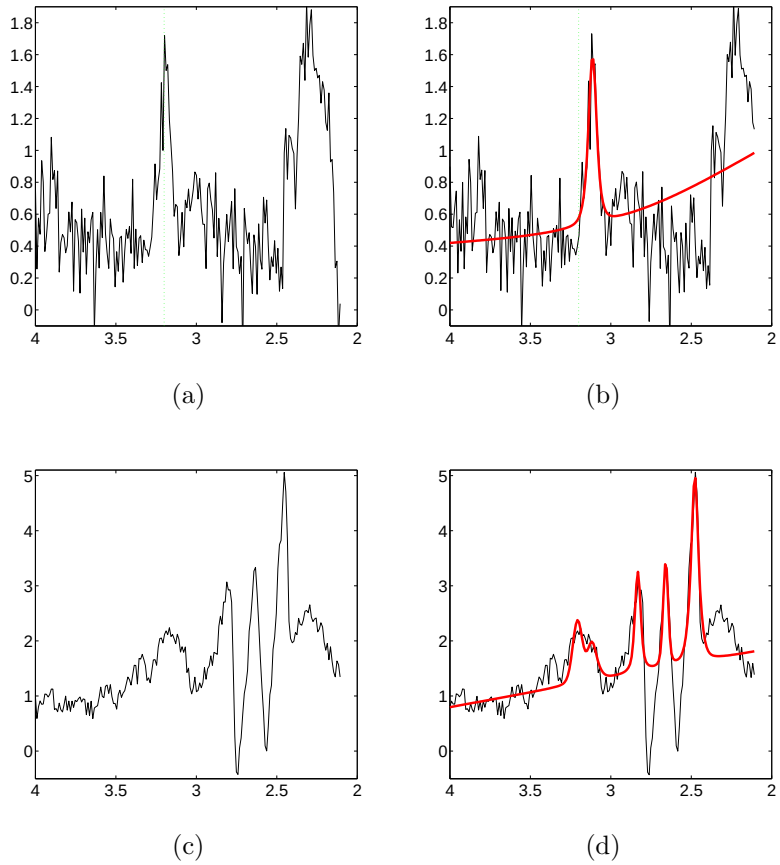


Figure 7: LCModel typical errors (a) spectrum with position of Choline (b) Spectrum corrected by LCModel : choline peak fitted as creatine (c) Spectrum before LCModel (d) Spectrum fitted by LCModel: complex form of citrate renders fit complicated. (on the ordinate arbitrate unit, on the abscissa chemical shift in ppm)

These experiments confirm the previous results concerning the problem of LCModel-SVM method to separate the classes efficiently (the total error is 24.58% with SVM and 28.25% with MLP). Classification errors are presented in table 3. For the other studies, we also obtained the best results using ACME algorithm and baseline correction (total error is 11.49%). The confusion matrix obtained for this optimum result is presented in table 4. We observed that SVM distinguished only with difficulty the "Pathology" class spectra from the "Undetermined" class spectra. Indeed, in the three-class analysis, 14.29% of "Pathology" voxels were classed in the "Undetermined" category.

5. Conclusion

If magnetic resonance spectroscopy is to be seriously considered as a clinical tool in the diagnosis and localisation of prostate cancer, a robust method for voxel classification is an absolute necessity. Indeed, the volume of information provided by 3D 1H MRSI renders the task of visually or manually classifying data from individual patients a very arduous and time consuming task. The aim of this work was then to study the feasibility of such a fast and automatic classification tool. We built a data set of 1062 spectra from 22 patients, allowing to evaluate experimentally several pre-processing steps (estimation of concentration of metabolites, direct use of spectra samples with or without phase and baseline correction), and classification methods (SVM and MLP).

The main conclusion from these experiments is that it is possible to automatically classify spectra and then to depict prostate cancer by using MR spectroscopy, with a total misclassification rate of 4.15%, a sensitivity of 83.57% and specificity of 98.11%. We have proven experimentally that it is more efficient to use the entire spectrum as an input vector of the classifier rather than use specific peak integrals or metabolite ratios, commonly used in previously published papers. We have also observed that SVM and MLP are equally efficient with a small advantage to the SVM mainly regarding the sensitivity.

The importance of appropriate baseline and phase corrections is also clearly illustrated: in all cases, application of the ACME routine with a baseline correction gave the best results in terms of mean misclassification rate, although the gain was relatively modest with respect to the module based approach alone. This is explained by the fact that both methods tend to

minimize "under-the-baseline" signals. Our results confirm that good baseline and phase correction are needed to correctly class spectra and encourage us to evaluate or develop in the next future more sophisticated baseline detection algorithms.

Our method allows also to speed up the global analysis process of one prostate, from around 10mn for an examination including estimation of metabolite concentrations, to less than 1mn using direct classification (this is an estimation but it does not take into account the selection of voxels within the prostate). This acceleration will allow us to investigate, for example, new processes taking into account interaction between neighbouring voxels, and not just classifying voxels independently.

In our work, the selection of voxels within prostate was performed manually on the basis of the positioning of the spectroscopic grid over the T_2-weighted images. This is particularly time consuming and would benefit from a semi-automatic or fully automatic treatment. Further work is necessary, but one could readily imagine methods based on the thresholding of water peak linewidths and water peak intensities rather than sophisticated algorithms applied directly on the spectra containing the metabolites of interest. Further work is also necessary to improve the classification performance of "Undetermined" spectra by introducing some new classification features, using a classification tree based on SVM, or by mixing MRS and MRI information.

This study is part of the Pharmimage® project. It is funded by the Conseil Régional de Bourgogne, European social fund and Oncodesign SA. We thank M. Provencher for the improvements he brought to his software to cater for prostate spectra.

References

[1] J. Kurhanewicz, D. B. Vigneron, H. Hricak, P. Narayan, P. Carroll, S. J. Nelson, Three-dimensional h-1 mr spectroscopic imaging of the in situ human prostate with high (0.24-0.7-cm3) spatial resolution, Radiology 198 (3) (1996) 795–805.

[2] T. W. Scheenen, S. W. Heijmink, S. A. Roell, C. A. Hulsbergen-Van de Kaa, B. C. Knipscheer, J. A. Witjes, J. O. Barentsz, A. Heerschap, Three-dimensional proton mr spectroscopy of human prostate at 3 t without endorectal coil: Feasibility, Radiology 245 (2) (2007) 507–16.

[3] P. M. Walker, G. Créhange, S. Parfait, D. Ben Salem, A. Cochet, C. Boichot, P. Maingon, F. Brunotte, Relationship between human prostate cancer adc measurements and 3d-mrsi at 3t without endorectal coils, in: ESMRMB annual congress, Vol. 21 suppl. 1, Springer, Valence, 2008.

[4] F. G. Claus, H. Hricak, R. R. Hattery, Pretreatment evaluation of prostate cancer: role of mr imaging and 1h mr spectroscopy, Radiographics 24 Suppl 1 (2004) S167–80.

[5] P. Tiwari, A. Madabhushi, M. Rosen, A hierarchical unsupervised spectral clustering scheme for detection of prostate cancer from magnetic resonance spectroscopy (mrs), Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv 10 (Pt 2) (2007) 278–86.

[6] P. Tiwari, M. Rosen, A. Madabhushi, A hierarchical spectral clustering and nonlinear dimensionality reduction scheme for detection of prostate cancer from magnetic resonance spectroscopy (mrs), Medical Physics 36 (9) (2009) 3927–3939.

[7] L. Matulewicz, K. Zakian, A. Shukla-Dave, J. Jansen, Y. Mazaheri, H. Hricak, J. Koutcher, A pattern recognition model for automatic classification of 1h mrsi voxels in the prostate (2009).

[8] B. M. Kelm, B. H. Menze, C. M. Zechmann, K. T. Baudendistel, F. A. Hamprecht, Automated estimation of tumor probability in prostate magnetic resonance spectroscopic imaging: Pattern recognition vs quantification, Magnetic Resonance in Medicine 57 (1) (2007) 150–159.

[9] O. Valenzuela, I. Rojas, F. Rojas, L. Marquez, Automatic classification of prostate cancer using pseudo-gaussian radial basis function neural network, in: European Symposium on Artificial Neural Networks - ESANN'2005, Bruges, Belgium, 2005.

[10] J. A. Jung, F. V. Coakley, D. B. Vigneron, M. G. Swanson, A. Qayyum, V. Weinberg, K. D. Jones, P. R. Carroll, J. Kurhanewicz, Prostate depiction at endorectal mr spectroscopic imaging: Investigation of a standardized evaluation system, Radiology 233 (3) (2004) 701–708.

[11] INTERPRET, International network for pattern recognition of tumours using magnetic resonance, http://gabrmn.uab.es/interpret/ (2002).

[12] A. Devos, L. Lukas, J. A. K. Suykens, L. Vanhamme, A. R. Tate, F. A. Howe, C. Majos, A. Moreno-Torres, M. van der Graaf, C. Arus, S. Van Huffel, Classification of brain tumours using short echo time 1h mr spectra, Journal of Magnetic Resonance 170 (1) (2004) 164–175.

[13] L. Lukas, A. Devos, J. A. K. Suykens, L. Vanhamme, F. A. Howe, C. Majs, A. Moreno-Torres, M. Van Der Graaf, A. R. Tate, C. Ars, S. Van Huffel, Brain tumor classification based on long echo proton mrs signals, Artificial Intelligence In Medicine 31 (1) (2004) 73–89.

[14] L. Journaux, J.-C. Simon, M. Destain, F. Cointault, J. Miteran, A. Piron, Plant leaf roughness analysis by texture classification with generalized fourier descriptors in a dimensionality reduction context, Precision Agriculture (2010) 1–16.

[15] J. Kaartinen, S. Mierisova, J. M. E. Oja, J.-P. Usenius, R. A. Kauppinen, Y. Hiltunen, Automated quantification of human brain metabolites by artificial neural network analysis from in vivo single-voxel1h nmr spectra, Journal of Magnetic Resonance 134 (1) (1998) 176–179.

[16] I. J. Bakken, D. Axelson, K. A. Kvistad, E. Brodtkorb, B. Muller, J. Aasly, I. S. Gribbestad, Applications of neural network analyses to in vivo 1h magnetic resonance spectroscopy of epilepsy patients, Epilepsy Research 35 (3) (1999) 245–252.

[17] A. A. Schricker, J. M. Pauly, J. Kurhanewicz, M. G. Swanson, D. B. Vigneron, Dualband spectral-spatial rf pulses for prostate mr spectroscopic imaging, Magnetic Resonance in Medicine 46 (6) (2001) 1079–1087.

[18] A. Shukla-Dave, H. Hricak, C. Moskowitz, N. Ishill, O. Akin, K. Kuroiwa, J. Spector, M. Kumar, V. E. Reuter, J. A. Koutcher, K. L. Zakian, Detection of prostate cancer with mr spectroscopic imaging: an expanded paradigm incorporating polyamines, Radiology 245 (2) (2007) 499–506.

[19] L. Chen, Z. Weng, L. Goh, M. Garland, An efficient algorithm for automatic phase correction of nmr spectra based on entropy minimization, Journal of Magnetic Resonance 158 (1-2) (2002) 164–168.

[20] R. Fischer, K. M. Hanson, V. Dose, W. von der Linden, Background estimation in experimental spectra, Physical Review E 61 (2) (2000) 1152.

[21] C. A. Lieber, A. Mahadevan-Jansen, Automated method for subtraction of fluorescence from biological raman spectra, Appl Spectrosc 57 (11) (2003) 1363–7.

[22] V. Mazet, D. Brie, J. Idier, Baseline spectrum estimation using half-quadratic minimization (September 2004).

[23] H.-W. Tan, S. D. Brown, Wavelet analysis applied to removing non-constant, varying spectroscopic background in multivariate calibration, Vol. 16, Wiley, Chichester, ROYAUME-UNI, 2002, journal of chemometrics.

[24] S. W. Provencher, Estimation of metabolite concentrations from localized in vivo proton nmr spectra, Magnetic Resonance in Medicine 30 (6) (1993) 672–9.

[25] G. Créhange, S. Parfait, M. Liegard, P. Maingon, D. Ben Salem, A. Cochet, M. Funes de la Vega, L. Cormier, F. Bonnetain, C. Mirjolet, F. Brunotte, P. M. Walker, Tumor volume and metabolism of prostate cancer determined by proton magnetic resonance spectroscopic imaging at 3t without endorectal coil reveal potential clinical implications in the context of radiation oncology, Int J Radiat Oncol Biol Phys.

[26] V. Vapnik (Ed.), The Nature of Statistical Learning Theory, springer-verlag Edition, Springer-Verlag, 1995.

[27] C.-C. Chang, C.-J. Lin, Libsvm : a library for support vector machines, software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm (2001).

[28] F. Rosenblatt (Ed.), Principles of Neurodynamics : perceptrons and the theory of brain mechanisms, Spartan Books, Washington DC, 1962.

[29] S. Chaplot, L. M. Patnaik, N. R. Jagannathan, Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network, Biomedical Signal Processing and Control 1 (1) (2006) 86–92.

[30] F. Smach, J. Miteran, M. Atri, J. Dubois, M. Abid, J.-P. Gauthier, An fpga-based accelerator for fourier descriptors computing for color object recognition using svm, Journal of Real-Time Image Processing 2 (4) (2007) 249–258.