

Journal of Biomedical Optics

SPIEDigitalLibrary.org/jbo

Hyperspectral imaging based method for fast characterization of kidney stone types

Francisco Blanco
Montserrat López-Mesas
Silvia Serranti
Giuseppe Bonifazi
Josef Havel
Manuel Valiente

Hyperspectral imaging based method for fast characterization of kidney stone types

Francisco Blanco,^a Montserrat López-Mesas,^a Silvia Serranti,^b Giuseppe Bonifazi,^b Josef Havel,^{c,d,e} and Manuel Valiente^a

^aUniversitat Autònoma de Barcelona, Centre Grup de Tècniques de Separació en Química (GTS), Unitat de Química Analítica, Departament de Química, 08193 Bellaterra, Spain

^bSapienza–Università di Roma, Dipartimento di Ingegneria Chimica Materiali Ambiente, 00184 Roma, Italy

^cMasaryk University, Department of Chemistry, Faculty of Science, Kamenice 5/A14, 625 00 Brno, Czech Republic

^dMasaryk University, Department of Physical Electronics, Faculty of Science, Kotlářská 2, 611 37 Brno, Czech Republic

^eMasaryk University, R&D center for low-cost plasma and nanotechnology surface modifications, Kotlářská 2, 611 37 Brno, Czech Republic

Abstract. The formation of kidney stones is a common and highly studied disease, which causes intense pain and presents a high recidivism. In order to find the causes of this problem, the characterization of the main compounds is of great importance. In this sense, the analysis of the composition and structure of the stone can give key information about the urine parameters during the crystal growth. But the usual methods employed are slow, analyst dependent and the information obtained is poor. In the present work, the near infrared (NIR)-hyperspectral imaging technique was used for the analysis of 215 samples of kidney stones, including the main types usually found and their mixtures. The NIR reflectance spectra of the analyzed stones showed significant differences that were used for their classification. To do so, a method was created by the use of artificial neural networks, which showed a probability higher than 90% for right classification of the stones. The promising results, robust methodology, and the fast analytical process, without the need of an expert assistance, lead to an easy implementation at the clinical laboratories, offering the urologist a rapid diagnosis that shall contribute to minimize urolithiasis recidivism. © 2012 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.JBO.17.7.076027]

Keywords: kidney stone; hyperspectral imaging; artificial neural networks; renal calculi characterization.

Paper 11631 received Oct. 28, 2011; revised manuscript received May 30, 2012; accepted for publication Jun. 26, 2012; published online Jul. 25, 2012.

1 Introduction

Urolithiasis, the formation of calculi within the urinary tract, is a rather common disease, which affects approximately 10% to 12% of the population in developed countries.¹ In this field, much effort has been done,² basically addressing the reduction of the recidivism rates and the increasing of the quality of life of patients, while cutting the medical costs for related treatments and surgery.

A careful study of the structure of the expelled kidney stone can give key information on the stone formation process and the urine conditions during the crystal growth.³ Thus, by specifically analyzing the core and outer shells of the sample, it is possible to determine the material that was serving as the precipitation nucleus and the substances that were precipitating afterwards. This information will provide a robust diagnosis that could lead to the appropriate treatment for every patient.

So far, several methodologies have been developed for the classification of kidney stones. The most extended ones are the examination of kidney stones by stereoscopic microscopy^{3,4} and infrared (IR) analysis.^{5,6} The former has an advantage in which it allows the identification of different substances on the whole area of the sample; so it is possible to fully determine the morphological characteristics of the stone. However, this methodology is laborious and, more significantly, it is strongly dependent on the operator. Concerning IR methodologies, they

require grinding the sample losing so the possibility of any spatial analysis.

In addition, the use of other techniques, namely near infrared (NIR),⁷ scanning electron microscopy-energy-dispersive X-ray spectroscopy (SEM-EDS),⁸ and even X-ray diffraction,⁹ has also been assessed. NIR spectroscopy showed a good performance for the determination of the composition of renal calculi and quantification for mixtures. Nevertheless those works still require some sample pretreatment which makes the analysis impractical to be applied in hospital facilities.

1.1 Hyperspectral Imaging Technique

Hyperspectral imaging (HSI) or chemical imaging technique is based on the utilization of an integrated hardware and software architecture able to measure a spectrum for each pixel of the acquired image, being then possible to characterize the whole surface of the sample,^{10,11} characteristics really useful for the objectives of the present work. It is desirable to analyze samples with flat surface because it always yields a better spectrum as the signal-to-noise (S/N) ratio increases due to better recording of the reflected radiation from the sample to the detector.

The acquired information is contained in a three-dimensional (3-D) dataset, characterized by two spatial dimensions and one spectral dimension, the so-called “hypercube”. According to the different wavelength and the spectral sensitivity of the device, several physical-chemical characteristics of a sample can be investigated and analyzed. For these reasons, HSI techniques

Address all correspondence to: Manuel Valiente, Universitat Autònoma de Barcelona, Centre Grup de Tècniques de Separació en Química (GTS), Unitat de Química Analítica, Departament de Química, 08193 Bellaterra, Spain. Tel: 34 93 581 2903; Fax: 34 93 581 1985; E-mail: Manuel.Valiente@uab.cat

represent an attractive solution for characterization, classification, and quality control of different products in many different fields, such as pharmaceuticals,^{12,13} medicine,^{14,15} food inspection,^{16,17} artworks,¹⁸ and materials recycling.^{19–22}

There are two conventional ways to construct the hypercube; the first one known as the “staring imager” configuration and the second known as “pushbroom” acquisition.

The “staring imager” configuration keeps the image field of view fixed, obtaining images one wavelength after another.²³ Hypercubes obtained using this configuration thus consist of a three-dimensional stack of images (one image for each wavelength examined), stored in band sequential (BSQ) format. Wavelength in the “staring imager” configuration is typically moderated using a tuneable filter being the acousto-optic tuneable filters (AOTFs) or the liquid crystal tuneable filters (LCTFs), the most predominantly employed. AOTFs have been used in the construction of commercially available NIR-CI systems. The main advantages of AOTFs are the good transmission efficiency, fast scan times, and large spectral range. On the other hand, LCTFs show greater promise in filtering of Raman images due to the superior spectral bandpass and image quality.

The “pushbroom” configuration is based on the acquisition of simultaneous spectral measurements from a series of adjacent spatial positions, which require relative movement between the object and the detector.²⁴ Some devices produce hyperspectral images based on a point step and acquire mode: spectra are obtained at single points on a sample, then the sample is moved and another spectrum taken. Hypercubes obtained using this configuration are stored in the band interleaved by pixel (BIP) format. Advances in detector technology have reduced the time required to acquire hypercubes. Line mapping instruments record the spectrum of each pixel in a line of sample which is simultaneously recorded by an array detector; the resultant hypercube is stored in the band interleaved by line (BIL) format.

In this study, a “pushbroom” configuration was adopted, based on the utilization of a device of the ImSpector™ series spectrometers, developed by SpecIm™ (Finland). The spectrographs are constituted by optics based on volume type holographic transmission grating.¹⁰ The grating is used in patented prism-grating-prism construction (PGP element), characterized by a high diffraction efficiency, good spectral linearity, and nearly free geometrical aberrations due to the on-axis operation principle.

For handling of the huge amount of data available from HSI, chemometric techniques are required. Principal component analysis (PCA) is usually used for screening the raw data before the application of a classification technique such as artificial neural network (ANN). ANNs show an ever-increasing number of applications in many fields, since they can satisfactorily solve complex analytical problems.^{25–27} Although simple in structure, the vast number of interconnections inside the ANN structure show an interesting potential for calculations.

The aim of the present study is to evaluate the application of HSI in the NIR field (1000 to 1700 nm) for the characterization and classification of renal calculi with the help of ANN, as this would be helpful for medical diagnosis, improving the conventional characterization done so far. In fact, HSI seems to be an interesting possibility to implement in hospital analyses where kidney stone samples have to be analyzed. This technique has already been used for testing a resolution method by analyzing two examples of kidney stones²⁸ and it has also been proved to

be useful for other medical applications such as surgery monitoring.^{29,30} However, HSI has not been deeply used for the study of kidney stones, once removed from the human body, or for the classification of the different types of kidney stones. On the other hand, ANNs have already been used in medicine with good results, even in urology,³¹ where the main applications have been related to cancer diagnostics and other illnesses, although they have never been used for the classification of kidney stones.

2 Materials and Methods

2.1 Sample Preparation

Two hundred and fifteen samples were selected from a library of more than 1400 renal calculi. Samples were collected at the urology service of the Hospital, Universitari de Bellvitge, Barcelona (Spain). All kidney stones were obtained either by surgical removal or by natural expulsion. After collection, the stones were thoroughly rinsed with water and ethanol. Once cleaned, the stones were stored in individual clean vials showing no decomposition or damage in the structure during periods longer than a year.

Eleven types of different kidney stone components, including their mixtures, were considered. Firstly, seven main types of kidney stones were considered as formed by the pure compounds: uric acid anhydrous (AUA), brushite (BRU), calcium oxalate dihydrate (COD), calcium oxalate monohydrate (COM), cystine (CYS), hydroxyapatite (HAP), and struvite (STR). Secondly, mixtures of the former ones, namely: uric acid dihydrate (AUD), mixed calcium oxalate and hydroxyapatite (MXL), mixed calcium oxalate and hydroxyapatite (MXD), and COD transformed into COM (TRA) were analyzed. The selection criterion was based on variability appearance, for each type of stone.

For HSI technique, it is desirable to get a flat surface for the analysis, so all the samples were cut with a surgical knife, although it was also possible to measure and correctly classify round parts of stones. The inner part of the whole sample was used for the analysis, since the equipment allows the measurement of the entire stone. Besides, the use of the complete stone helped on the characterization of all its parts, contributing to the correct classification of each sample. Many samples showed a heterogeneous surface. In those cases, both the external part of the stone and the core were analyzed.

2.2 Conventional Methodologies

Samples were firstly analyzed by means of stereoscopic microscopy. The analyses were carried out as described in the reported bibliography.^{3,4} For those samples that were not well characterized by this method, an SEM-EDS analysis was performed.⁸ For this purpose, two different SEM equipments were used: JEOL JSM-6300 Scanning Electron Microscope (Japan), coupled to an Oxford Instruments Link ISIS-200 (UK) X-Ray Dispersive Energy Spectrometer (Univ. Autònoma de Barcelona); and a HITACHI S2500 (Japan) Scanning Electron Microscope, coupled to a KeveX 8000 (USA) X-Ray Dispersive Energy Spectrometer (Univ. “La Sapienza” of Rome). The structure of the samples was analyzed, and the EDS analysis was performed on some parts of the stone to confirm the elemental composition of the sample. Some representative results are shown in Figs. 1 and 2.

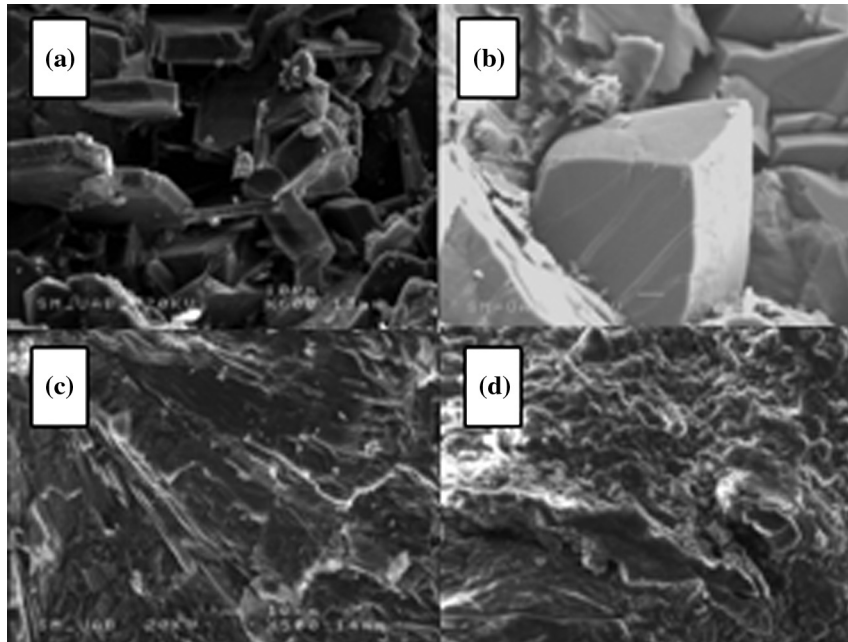


Fig. 1 SEM images of the internal part of some kidney stones. (a) AUD, (b) COD, (c) COM, and (d) HAP.

The results coming out from these conventional methodologies were used as references for the results obtained from HSI data.

2.3 Hyperspectral Imaging Device and Architecture Set-Up

The HSI system used (Univ. “La Sapienza” of Rome) was constituted by the following components: optics, spectrograph, camera, translation unit, energizing source, and control unit (Fig. 3). The core of the system was a spectral camera NIR (Specim, Finland) embedding an ImSpector™ N17E imaging spectrograph working in the wavelength region 1000 to 1700 nm and a Te-cooled InGaAs photodiode array camera (spatial resolution: rms spot radius <math><15\ \mu\text{m}</math>; spectral resolution: 7 nm; 121 wavelengths measured).

The spectrometer was coupled to a 50 mm lens. The resolution of the image width was 320 pixels, while the number of frames, that is, the resolution of the image in the Y axes varied

from 200 to 350 pixels, depending on the number of samples measured at a time.

The energizing source was constituted by a diffused light cylinder with aluminium internal coating, embedding five halogen lamps that produce an intense and continuous spectrum signal, which was optimized for the NIR region.

The spectral camera was hosted in a laboratory platform equipped with an adjustable speed (from 0 to 50 mm/s) conveyor belt (width = 26 cm and length = 160 cm). Spectra acquisition can be carried out continuously or at specific time intervals. The device was fully controlled by a PC unit equipped with the Spectral Scanner™ v.2.3 acquisition/preprocessing software (DV srl, Italy).

The spectra of the samples were measured fixing them to a plastic holder with the inner side upwards, so the core was visible to the detector.

The stones previously classified as pure compounds, by means of conventional methodologies, were used for creating a library of compounds, which could be later applied for the classification of unknown samples.

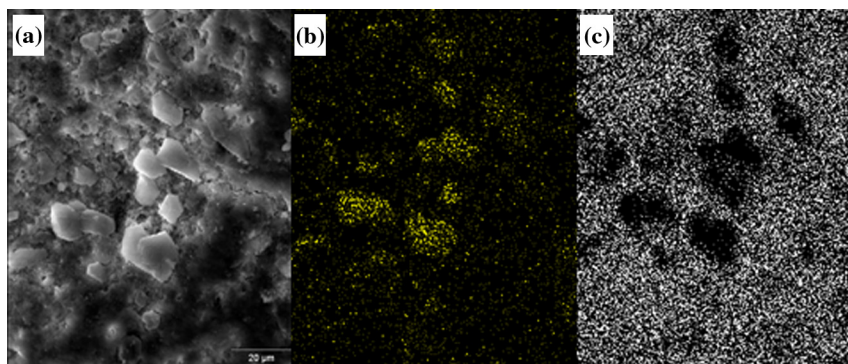


Fig. 2 Mapping experiments on the surface of a kidney stone. (a) SEM image of some a struvite kidney stone; (b) mapping for Mg; and (c) mapping for Ca.



Fig. 3 HSI based architecture employed for the acquisition of the kidney stones spectra.

2.4 Spectra Handling

The acquisition of the spectral signatures was carried out after a preliminary calibration performed in two steps: i) black image acquisition and ii) measurements of “white reference image” using a standard white ceramic tile. After the calibration phase, the spectral image was acquired and the reflectance (R) was computed according to the following equation:

$$R = \frac{r_s - r_b}{r_w - r_b},$$

where r_s is the reflectance measured for the sample, r_b is the reflectance measured for the black (background noise), and r_w is the reflectance of the standard white (100% reflectance).

The analysis of the stones was performed on five randomly selected region of interests (ROIs) on the surface of the sample in order to avoid the selection of any part of the stone that might contain organic matter. Then, all the regions in a sample were considered as a single sample for data treatment. Although the ROIs analysis loses the information for every single pixel, it has two main advantages. First, it is a way of smoothing the data, since the surface of the sample might not be homogeneous considering adjacent pixels due to organic matter which could have been trapped into the structure of the stone during the crystallization process. Second, the amount of information that is going to be handled is much smaller than in the analysis of the full pixel matrix.

Additionally, an image analysis for each individual pixel was performed using a PCA treatment to show the capabilities of this technique for handling entire images on the classification process.

2.5 Data Treatment

For the analysis of the data, three different methods were used: factor analysis (FA), PCA, and ANN.

Although FA and PCA share the same main goal, that is, the reduction of the number of variables of the system by finding latent connections between real variables, they have a basic difference that allows obtaining different information.³²

FA is based on correlation between variables, and the number of factors considered descriptors of the system is selected according to the associated eigenvalue.^{33,34} Several criteria have been suggested in order to decide this number of factors. Kaiser’s rule, which states that only those factors with an associated eigenvalue greater than 1 are representative for the system,³² has been applied in this study. All other factors stand for linear combinations of real variables, describing only noise.

On the other hand, PCA is based on the variance of the data.³⁵ Therefore, the main principal component (PC) for the system follows the direction in which the data have a bigger variance.

Due to these properties, FA was used to check how many different components can be distinguished in all the samples, while PCA was useful for the variable selection.

With regard to ANNs, they are considered a sophisticated and powerful computational tool which solves difficult analytical problems by learning from real cases. An ANN is a computational model formed from a certain number of single units, artificial neurones or nodes, connected with coefficients (weights), w_{ij} , which constitute the neural structure. Despite many different structures that ANNs can take, already described in the literature,^{25,36} the structure used in this study was constituted by three layers: inputs, one hidden layer, and outputs. The input layer receives the information about the system (the nodes of this layer are simple distributive nodes, which do not alter the input value at all). The hidden layer processes the information initiated at the input, while the output layer is the observable response or behavior. The algorithm chosen for the learning process was the backpropagation. The optimization of the ANN was carried out by minimizing the root mean square (RMS) error,³⁷ when modifying the number of nodes in the hidden layer.

2.6 Software

HSI data has been treated using the following software: STATISTICA (Tulsa, OK);³⁸ The Unscrambler v 9.1 (Camo Process, Oslo, Norway);³⁹ MATLAB v 7.0 (MA, USA).⁴⁰ PLS toolbox by Eigen-vector Research, Inc. and TRAJAN v 3.0 (Horncastle, UK).⁴¹

3 Results

3.1 Factor Analysis

The software STATISTICA was used to perform FA on the data. The data used for this analysis was the spectra corresponding to the seven types of different components, since they represent the whole variety of substances studied.

Figure 4 shows how the weight of the Eigenvalue associated to each factor decreases for less important factors, being the seventh factor the last one having a value greater than 1. According to Kaiser’s rule, it can be seen that FA distinguishes the main species forming renal calculi.

Taking into account these results, a PCA analysis was carried out in order to create a model that is able to classify the different types of kidney stones analyzed.

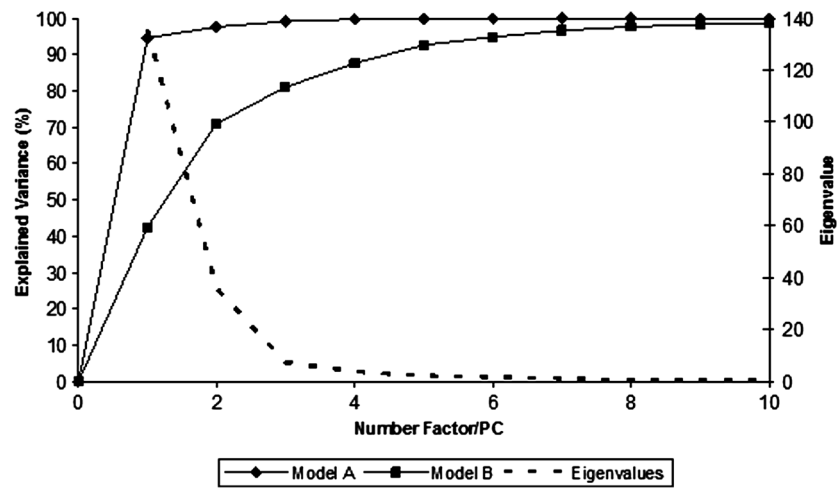


Fig. 4 Eigenvalues for the seven main groups of renal calculi and explained variance for the PCA model. Model A: reflectance data, Model B: first derivative.

3.2 Principal Components Analysis

PCA was used for screening the spectral data. It was useful for selecting variables before performing the classification, which was made using ANNs.

3.2.1 PCA on each type of renal calculi: identification of outliers

To determine the existence of outliers within the raw data, a PCA analysis was performed on the seven types of kidney stones, including each type of different samples of the same chemical composition. The data was cleaned in order to have a more reliable database.

The elimination of any point from the original samples is a crucial step, since it decreases the variability of the system. Moreover, there is no way to be completely positive whether a sample is correctly considered as an outlier. Hence, the criterion for the determination of outliers was based on the variability of the spectra. Due to the biological nature of the samples, the structure of the crystals that form the kidney stone might contain organic matter. Moreover, the surface measured may not be completely regular (even after cutting with a surgical knife), thus having a reflectance value differing from the average of the group.

The software ‘The Unscrambler’ was used to perform PCA analysis on 5 ROIs for every type of each group of samples considering each ROI as a different sample in terms of the software application. In this very first step, the whole range of variables was taken into account for the calculations, that is, 97 variables. By organizing the data, it was possible to check if any of the regions or even a whole sample was really different from the rest.

A point was considered an outlier if it was located outside the Hotelling T^2 ellipse of the scores plot. Nonetheless, a sample was not considered if many regions of it were in such a position. In this way, the over-fitting of the data was avoided.

3.2.2 PCA on the seven main groups of renal calculi

After cleaning the data, a PCA was performed on the seven main groups of renal calculi in order to create a model able to correctly classify the kidney stones studied. This model was

done taking all variables into calculation. The efficiency was checked by means of cross-validation and all data was mean centered.

Firstly, the model was directly calculated from the acquired reflectance spectra. Figure 4 Model A shows that, in this case, the two first PCs stand for the 95% of the variance of the model, having the rest of PCs a really low value.

In order to get a wider distribution of the variance of the data, the first derivative of the raw spectra was calculated. This step requires a previous smoothing of the data, which was done by means of the Savitzky-Golay algorithm, using a 5-point window. In this case, a much distributed explanation of the variance is seen; namely, up to 7 PCs are needed to explain the 96% of the variance for the model (Fig. 4 Model B).

Consequently, it can be concluded that first derivative magnifies the slight differences existing between the seven groups of renal calculi. For this reason, the first derivative of the spectra was used for further treatment of the data.

3.2.3 Variable selection

Multivariate methods can certainly give a higher amount of information than univariate ones, though many variables from the measured range might not give valuable data, but noise. In this sense, we can take advantage profit from the fact that one of the basic statements of PCA is the reduction of number of variables, by taking only the information contained in the independent variables. Actually, not all the regions of the NIR spectrum give relevant information to the model. Then, by choosing those wavelengths with the greatest classification capabilities, the noise of the system is reduced thereby increasing its precision and simplifying the calculations.

To achieve this goal, the selection of variables is based on their loading value. In a PCA analysis, the loadings stand for the weight that every real variable has on each of the principal components. The wavelengths to be taken into further calculations are those that have a highest value for the loading for each PC. The representation of the loadings of every wavelength for each PC (Fig. 5) is a clear way to see the most important wavelengths, since it can be seen which regions of the spectrum are mostly influencing for a given PC. On the other hand, the regions of the spectrum that have a bigger weight on a given

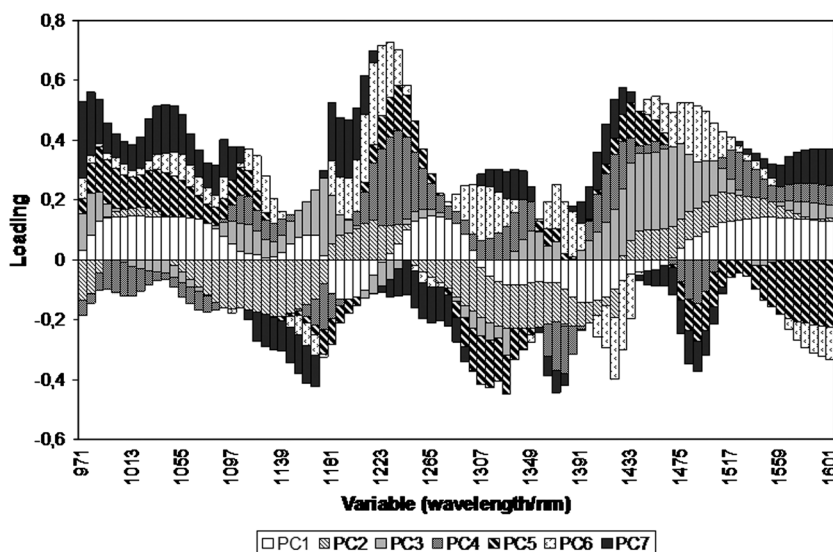


Fig. 5 Loadings of each variable for the seven first PCs.

PC should correspond to the vibration wavelength of the compound classified by this PC.

Following the above mentioned path, it was possible to define the variables defining every compound from the information in Fig. 5 and the representation of the scores values, see results in Table 1.

3.3 ANN: Optimization of the Neural Network

The software 'TRAJAN' was used for creating the ANN model, using the first derivative of the reflectance spectra measured.

The raw data of the spectra contain all the measured wavelengths, so in order to decrease the number of variables in the system 50 wavelengths were selected according to the values of the loadings for each wavelength, as shown in Fig. 5. In this way, a matrix of 140 samples (seven main types of kidney stones) and 50 variables was used for the analysis.

The ANN to be optimized was defined as: (inputs, number of nodes in the hidden layer, outputs), where inputs stand for the number of samples used to create the model, and outputs, the number of types of kidney stones, that is (140, n, 7). The value for the outputs was a single variable, which could take up to seven different nominal values.

The algorithm used for the optimization of the structure was the backpropagation. After the optimization of the ANN, it was seen that the optimum number of nodes in the hidden layer was four, and the value after which a non decrease in the RMS error²⁴ was observed. To avoid an overfitting of the model, no more nodes were used.

4 Discussion

4.1 PCA and Variable Selection

It has been seen that PCA allows for the determination of the main variables that define the system. For instance, the scores and loadings for PC2 can be also analyzed (see Fig. 6). In this case, AUA and CYS have opposite scores sign, meaning that PC2 is clearly differentiating between these two components. PC2 is well defined with positive loading values for wavelengths from the ranges 1188 TO 1230 nm and 1440 to 1542 nm. All these vibrations are associated to C–H bond vibrations. This association perfectly fits with the structure of cystine, since this is the only component of the studied set which has C–H bonds. On the other hand, wavelengths 971 to 978 nm and 1083 to 1167 nm have negative loading values for PC2. In

Table 1 Characteristic NIR vibrations for each type of renal calculi.

Type of kidney stone	Characteristic PC	Δ with highest loading (nm)	Associated NIR vibrations
CYS	PC1 < 0 ... PC2 > 0	1188–1230, 1440–1542	CH, CH ₂ , CH ₃
AUA	PC1 < 0 ... PC2 < 0	971–978, 1083–1167	Ar-OH
COD COM	PC3 > 0 ... PC2 > 0 PC3 < 0 ... PC4 > 0	1426–1475	H ₂ O
BRU	PC3 < 0 ... PC4 < 0	1223–1244	CH
HAP STR		Not well defined	

this case, the associated bond vibrations are the Ar-OH, which can be related only to AUA.

Additionally to this case, some other data defining the rest of compounds can be observed in Table 1. Clearly, the vibration differentiating COM and COD is the band for water.

It can be appreciated that the NIR bands that define BRU appear in the range of C-H vibrations. This fact might be explained due to especially high amount of organic matter contained in BRU calculi, since no C-H bonds are found in brushite structure.

The main drawback the PCA has shown was the difficulty to distinguish HAP and STR results which can be clearly seen in Fig. 6, where these two compounds appear as two overlapped clusters.

Though the only exception of HAP and STR renal calculi, the PCA analysis has shown to be useful for the classification of the components of kidney stones.

4.2 Image Analysis

A different PCA analysis was also performed on the seven groups of kidney stones, taking directly the hyperspectral cube and using the software MATLAB.

One sample from each type of kidney stone was analyzed, getting a classification pixel-by-pixel. As in the previous PCA, the data used was the first derivative of the spectra and the data were again mean centered.

Some samples considered pure compounds (as characterized by means of stereoscopic and SEM-EDS microscopies) were used to create a model for later classification of unknown samples.

The identification of the different compounds was done by interpretation of the colors obtained when creating reconstructed RGB images from the hyperspectral cube. When using this kind

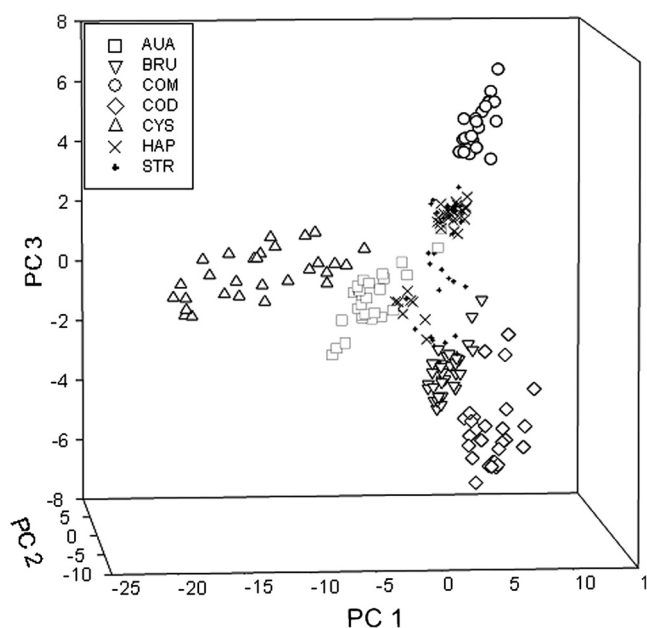


Fig. 6 3D representation of scores for PC1, PC2, and PC3.

of representation, the scores values for each of the 3 PCs used for the plot will determine the color of the pixel.

Figure 7 shows different colors for each kind of kidney stones, providing a simple way to distinguish a type among the others, except for STR and HAP. Indeed, the results obtained analyzing the hyperspectral cube completely matched with those obtained by taking ROIs from the samples.

Moreover, the software allows the selection of a group of pixels from the scores representation. Therefore, it is possible to identify which area of a sample or group of samples contain a given compound, when the scores values of each one are known.

An example is shown in Fig. 8. Three different areas are highlighted: COD, COM, and the interface, namely TRA.

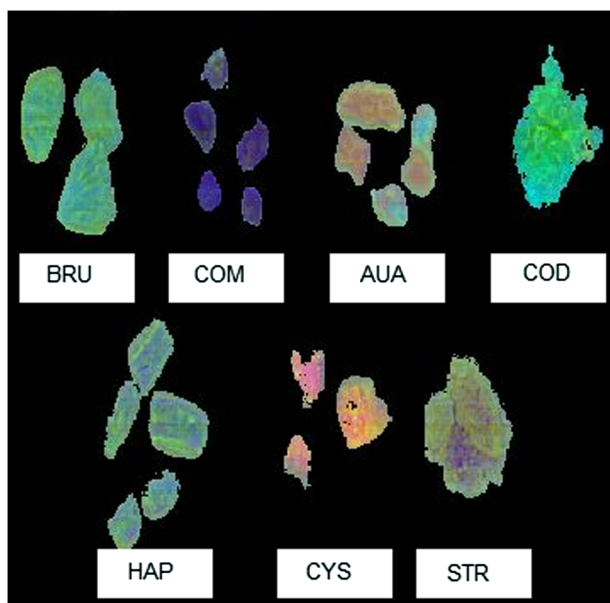


Fig. 7 RGB image for the seven main groups of renal calculi.

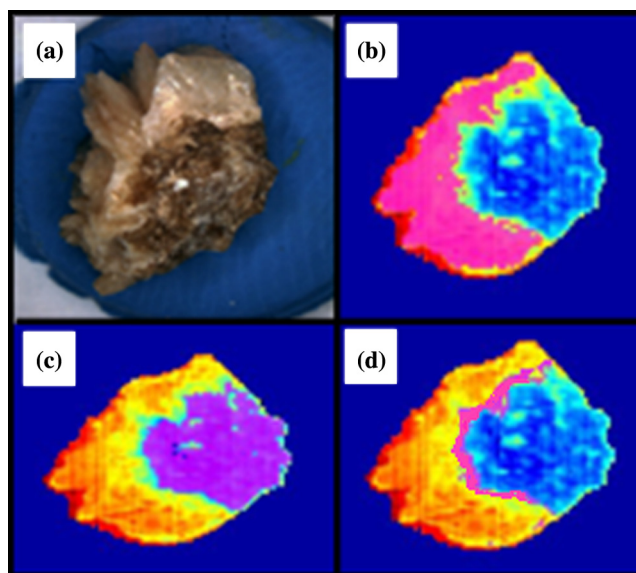


Fig. 8 (a) Image of the real stone; highlighted: (b) COD, (c) COM, and (d) TRA.

This last compound is the result of the slow change that the COD (kinetic derivative) undergoes into COM (more stable).

4.3 ANN Performance

The model was applied to the seven groups of main compounds forming kidney stones, using the previous conditions. The results, checked with cross-validation leave-one-out, show up to 100% of accuracy for the classification.

The next step was to further reduce the number of variables in the system, so that the foretold advantages of the variable reduction become clearer. With only 30 wavelengths selected, the accuracy for the classification remained at 100%.

When even less variables are taken for the calculations, the percentage for right classification dropped down to 95% at most. Consequently, the final choice was the model including 30 wavelengths.

However, due to the importance of a precise classification regarding the diagnostic for the patient, the really interesting objective is the classification of 11 groups of different types of kidney stones, that is, the seven main compounds and their mixtures.

Thus, an optimization of a new ANN model, capable for the classification of this much more complex data was required. For this purpose, 50 wavelengths were introduced as variables, being the structure of the optimized ANN (215, 13, 11). In this case, the rate of well-classified samples reached 94.4%. The reader is referred to Table 2 for the accurate classification of the set of 215 samples of kidney stones used for this study showing the classification obtained by the conventional technique and that developed in the present work.

Table 2 Comparison between the results obtained using conventional techniques (stereoscopic microscopy and SEM) and the new developed methodology (HSI-ANN). Note: For the classification of the samples by HSI, a library of substances was defined with some stones considered as pure compounds. The items highlighted are the erroneously classified.

Sample	Classification conventional tech.	Classification HSI
1	COM	COM
2	COM	COM
3	TRA	TRA
4	MXD	MXD
5	AUA	AUA
6	COM	COM
7	STR	STR
8	TRA	TRA
9	HAP	HAP
10	COM	COM
11	MXL	MXL

Table 2 (Continued).

Sample	Classification conventional tech.	Classification HSI
12	COM	COM
13	COD	COD
14	AUA	AUA
15	AUA	AUA
16	STR	STR
17	COD	COD
18	STR	STR
19	HAP	HAP
20	MXD	MXD
21	COM	COM
22	COM	COM
23	COM	COM
24	MXD	MXD
25	AUA	AUA
26	AUA	AUA
27	COD	COD
28	AUA	AUA
29	COM	COM
30	HAP	HAP
31	MXD	MXD
32	MXL	MXL
33	TRA	COM
34	STR	STR
35	MXD	MXD
36	COM	COM
37	COM	COM
38	AUA	AUA
39	AUD	AUD
40	STR	STR
41	TRA	TRA
42	TRA	TRA
43	COM	COM
44	COM	COM
45	HAP	HAP

Table 2 (Continued).

Sample	Classification conventional tech.	Classification HSI
46	COD	COM
47	MXD	MXD
48	MXL	MXL
49	TRA	TRA
50	TRA	TRA
51	COM	COM
52	AUA	AUA
53	AUA	AUA
54	COD	COD
55	COD	COD
56	COM	COM
57	AUA	AUA
58	HAP	HAP
59	BRU	BRU
60	HAP	HAP
61	TRA	TRA
62	COM	COM
63	AUA	AUA
64	COM	COM
65	COM	COM
66	STR	STR
67	COD	COD
68	COD	COD
69	MXL	MXL
70	HAP	HAP
71	COM	COM
72	TRA	COD
73	AUA	AUA
74	AUA	AUA
75	STR	STR
76	STR	STR
77	TRA	TRA
78	COM	COM
79	COM	COM

Table 2 (Continued).

Sample	Classification conventional tech.	Classification HSI
80	MXD	MXD
81	AUA	AUA
82	COD	COD
83	HAP	HAP
84	COD	COD
85	COD	COD
86	COM	COM
87	AUD	AUD
88	MXL	MXL
89	TRA	TRA
90	TRA	COM
91	TRA	TRA
92	COM	COM
93	COD	COD
94	COM	COM
95	STR	STR
96	MXD	MXD
97	HAP	HAP
98	BRU	BRU
99	AUD	AUD
100	COD	COD
101	COD	COM
102	AUA	AUA
103	HAP	HAP
104	HAP	HAP
105	TRA	TRA
106	COM	COM
107	COM	COM
108	COM	COM
109	COD	COM
110	CYS	CYS
111	STR	STR
112	STR	STR
113	TRA	TRA

Table 2 (Continued).

Sample	Classification conventional tech.	Classification HSI
114	HAP	HAP
115	BRU	BRU
116	MXL	MXD
117	TRA	TRA
118	TRA	TRA
119	COD	COD
120	COD	COD
121	HAP	HAP
122	COM	COM
123	MXD	MXD
124	AUD	AUD
125	BRU	BRU
126	AUA	AUA
127	AUA	AUA
128	TRA	TRA
129	CYS	CYS
130	MXL	MXL
131	AUA	AUA
132	TRA	TRA
133	COM	COM
134	STR	STR
135	AUD	AUD
136	COD	COD
137	AUA	AUA
138	MXL	MXL
139	COM	COM
140	AUD	AUD
141	AUA	AUA
142	BRU	BRU
143	TRA	COM
144	COD	COD
145	COD	COD
146	COD	COD
147	COM	COM

Table 2 (Continued).

Sample	Classification conventional tech.	Classification HSI
148	COM	COM
149	HAP	HAP
150	CYS	CYS
151	AUD	AUD
152	AUA	AUA
153	TRA	TRA
154	TRA	TRA
155	AUD	AUD
156	MXL	MXL
157	BRU	BRU
158	COD	COD
159	MXD	MXD
160	MXD	MXD
161	COM	COM
162	STR	STR
163	COM	COD
164	TRA	TRA
165	COD	COD
166	HAP	HAP
167	CYS	CYS
168	CYS	CYS
169	AUD	AUD
170	STR	STR
171	AUD	AUD
172	TRA	TRA
173	COM	COM
174	COM	COM
175	AUA	AUA
176	AUD	AUD
177	MXL	MXL
178	BRU	BRU
179	CYS	CYS
180	AUD	AUD
181	COD	COD

Table 2 (Continued).

Sample	Classification conventional tech.	Classification HSI
182	COD	COD
183	STR	STR
184	MXD	MXL
185	AUD	AUD
186	HAP	HAP
187	TRA	TRA
188	HAP	HAP
189	AUD	AUD
190	COM	TRA
191	AUD	AUD
192	STR	STR
193	STR	STR
194	COD	COD
195	CYS	CYS
196	BRU	BRU
197	BRU	BRU
198	TRA	TRA
199	TRA	TRA
200	TRA	TRA
201	COM	COM
202	CYS	CYS
203	AUD	AUD
204	COD	COD
205	AUA	AUA
206	AUA	AUA
207	MXD	MXL
208	CYS	CYS
209	CYS	CYS
210	BRU	BRU
211	STR	STR
212	STR	STR
213	HAP	HAP
214	AUD	AUD
215	AUA	AUA

As seen in this Table, the error for the last model, including all types of kidney stones, is only slightly higher than that for the simpler model. Nevertheless, the second model offers a great advantage, since it is able to deal with any of the kidney stone types usually found. The prediction differences between the newly developed methodology and the conventional techniques involve basically calcium oxalate kidney stones. Considering the similar nature of those types of stones and the fact that their composition often includes mixtures of lithogenic substances, observed differences appear to be logical. Information collected in Table 3 summarizes the results discussed here that represent an efficient classification of kidney stones from a very simple methodology.

In this work, the suitability of the NIR-HSI technique with the use of ANNs for the characterization of kidney stones has been demonstrated. The regions of the NIR reflectance spectra that have the strongest power for the classification of the different components have been identified. The use of ANNs for the hyperspectral data treatment has proved to produce similar results to those obtained from conventional techniques, including samples containing compound mixtures. However, it is important to remark that the conventional methodology requires trained operators, whereas ANNs perform the classification independently from the operator's knowledge.

The implementation of this developed methodology as a routine analysis in medical practice might be achievable for a clinical laboratory, and it would be simple to perform. The method requires little training for its use, and no special knowledge about chemistry is needed. Once the software has been installed in the computer, the fast measurements allow a quick and easy classification of the stones. This application would allow obtaining a faster and much robust diagnosis, increasing the quality of treatment to each specific patient.

Table 3 Percentage of samples correctly classified by the developed HSI-ANN methodology.

Type of kidney stone	No. of samples	No. samples correctly classified	% of correct classification
COM	39	37	94.9
COD	27	24	88.9
TRA	27	23	85.2
HAP	18	18	100
STR	19	18	100
MXL	10	9	90
MXD	13	11	84.6
BRU	25	25	100
AUA	17	17	100
AUD	10	10	100
CYS	10	10	100
		Total	94.4

Acknowledgments

The Spanish Ministry for Science and Innovation (MICINN) is acknowledged for financial support to the present study (Project CTQ2009-07432 (subprogram PPQ)). The Spanish-Italian Program for Integrated Actions is acknowledged to allow joint experiments with respective scientists (Grant of Reference IT2009-0024). Prof. Josef Havel acknowledges the support of MICINN to his sabbatical stay at UAB (SAB2009-0005), the J. H. Support from Ministry of Education, Youth and Sports of the Czech Republic (Projects MSM, 0021622411, 0021627501, and CZ.1.05/2.1.00/03.0086), and the Czech Science Foundation (Projects No. 104/08/0229, 202/07/1669).

References

- O. W. Moe, "Kidney stones: pathophysiology and medical management," *Lancet* **367**(9507), 333–344 (2006).
- C. Pak, "Kidney stones," *Lancet* **351**(9118), 1797–1801 (1998).
- F. Grases et al., "Simple classification of renal calculi closely related to their micromorphology and etiology," *Clin. Chim. Acta* **322**(1–2), 29–36 (2002).
- A. Nayir, "Determination of urinary calculi by binocular stereoscopic microscopy," *Pediatr. Nephrol.* **17**(6), 425–432 (2002).
- E. V. Wilson, M. J. Bushiri, and V. K. Vaidyan, "Characterization and FTIR spectral studies of human urinary stones from Southern India," *Spectrochim. Acta A* **77**(2), 442–445 (2010).
- C. A. Lehmann, G. L. McClure, and I. Smolens, "Identification of renal calculi by computerized infrared spectroscopy," *Clin. Chim. Acta* **173**(2), 107–116 (1988).
- E. Peuchant et al., "Discriminant analysis of urinary calculi by near-infrared reflectance spectroscopy," *Clin. Chim. Acta* **205**(1–2), 19–30 (1992).
- D. R. Basavaraj et al., "The Role of Urinary Kidney Stone Inhibitors and Promoters in the Pathogenesis of Calcium Containing Renal Stones," *EAU-EBU Update Ser.* **5**(3), 126–136 (2007).
- A. I. Ancharov et al., "Model experiment of in vivo synchrotron X-ray diffraction of human kidney stones," *Nucl. Instrum. Meth. A* **575**(1–2), 221–224 (2007).
- T. Hyvarinen, E. Herrala, and A. Dall'Ava, "Direct sight imaging spectrograph: a unique add-on component brings spectral imaging to industrial applications," *Proc. SPIE* **3302**, 165–356 (1998).
- P. Geladi, H. Grahn, and J. Burger, "Multivariate images, hyperspectral imaging: background and equipment," in *Techniques and Applications of Hyperspectral Image Analysis*, H. Grahn and P. Geladi, Eds., pp. 1–15, John Wiley & Sons, West Sussex, England (2007).
- J. M. Amigo et al., "Study of pharmaceutical samples by NIR chemical-image and multivariate analysis" *TRAC-Trend Anal. Chem.* **27**(8), 696–713 (2008).
- A. A. Gowen et al., "Recent applications of chemical Imaging to pharmaceutical process monitoring and quality control," *Eur. J. Pharm. Biopharm.* **69**(1), 10–22 (2008).
- R. Jolivot, P. Vabres, and F. Marzani, "Reconstruction of hyperspectral cutaneous data from an artificial neural network-based multispectral imaging system," *Comput. Med. Imaging Graph.* **35**(2), 85–88 (2011).
- Z. Liu et al., "Classification of hyperspectral medical tongue images for tongue diagnosis," *Comput. Med. Imaging Graph.* **31**(8), 672–678 (2007).
- A. A. Gowen et al., "Hyperspectral imaging—an emerging process analytical tool for food quality and safety control," *Trends Food Sci. Tech.* **18**(12), 590–598 (2007).
- A. Del Fiore et al., "Early detection of toxigenic Fungi on maize by hyperspectral imaging analysis," *Int. J. Food Microbiol.* **144**(1), 64–71 (2010).
- M. Kubik, "Hyperspectral imaging: a new technique for the non-invasive study of artworks," in *Physical Techniques in the Study of Art, Archaeology and Cultural Heritage 2*, D. Creagh and D. Bradley, Eds., pp. 199–259, Elsevier, Amsterdam (2007).
- G. Bonifazi and S. Serranti, "Imaging spectroscopy based strategies for ceramic glass contaminants removal in glass recycling," *Waste Manage.* **26**(6), 627–639 (2006).
- G. Bonifazi and S. Serranti, "Hyperspectral imaging based procedures applied to bottom ash characterization," *Proc. SPIE* **6755**, 67550B (2007).
- G. Bonifazi et al., "Innovative sensing technologies applied to post-consumer polyolefins recovery," *Metal Int.* **14**(2), 5–10 (2009).
- S. Serranti, A. Gargiulo, and G. Bonifazi, "Characterization of post-consumer polyolefin wastes by hyperspectral imaging for quality control in recycling processes," *Waste Manage.* **31**(11), 2217–2227 (2011).
- E. Lewis et al., "Near-infrared chemical imaging as a process analytical tool," in *Process Analytical Technology*, K. Bakeev, Ed., p. 187, Blackwell Publishing, Oxford (2005).
- K. C. Lawrence et al., "Calibration of a pushbroom hyperspectral imaging system for agricultural inspection," *T ASABE* **46**(2), 513–521 (2003).
- G. Hanrahan, "Computational neural networks driving complex analytical problem solving," *Anal. Chem.* **82**(11), 4307–4313 (2010).
- P. J. Lisboa, E. C. Ifeachor, and P. S. Szczepaniak, *Artificial Neural Networks in Biomedicine*, Springer, London (2000).
- S. Agatonovic-Kustrin and R. J. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," *Pharmaceut. Biomed.* **22**(5), 717–727 (2000).
- S. Piqueras et al., "Resolution and segmentation of hyperspectral biomedical images by multivariate curve resolution-alternating least squares," *Anal. Chim. Acta* **705**(1–2), 182–192 (2011).
- M. S. Holzer et al., "Assessment of renal oxygenation during partial nephrectomy using hyperspectral imaging," *J. Urology* **186**(2), 400–404 (2011).
- L. Jafari-Saraf and I. L. Gordon, "Hyperspectral imaging and ankle: brachial indices in peripheral arterial disease," *Ann. Vasc. Surg.* **24**(6), 741–746 (2010).
- J. T. Wei et al., "Understanding artificial neural networks and exploring their potential applications for the practicing urologist," *Urology* **52**(2), 161–172 (1998).
- J. E. Arruda et al., "A guide for applying principal-components analysis and confirmatory factor analysis to quantitative electroencephalogram data," *Int. J. Psychophysiol.* **23**(1–2), 63–81 (1996).
- C. J. Kessler et al., "Factor analysis of trends in Texas acidic deposition," *Atmos. Environ.* **26**(6), 1137–1146 (1992).
- W. F. Velicer and D. N. Jackson, "Component analysis versus common factor analysis: some issues in selecting an appropriate procedure," *Multivariate Behav. Res.* **25**(1), 1–28 (1990).
- R. G. Brereton, *Applied Chemometrics for Scientists*, Wiley, England (2007).
- D. F. Brougham et al., "Artificial neural networks for classification in metabolomic studies of whole cells using 1 h nuclear magnetic resonance," *J. Biomed. Biotechnol.* (2011).
- J. Havel, P. Lubal, and M. Farková, "Evaluation of chemical equilibria with the use of artificial neural networks," *Polyhedron* **21**(14–15), 1375–1384 (2002).
- Statsoft, "Big Data Analytics, Enterprise Analytics, Data Mining Software, Statistical Analysis, Predictive Analytics", 2012, <http://www.statsoft.com/> (July 2012).
- Camo, "Leading Multivariate Data Analysis & Design of Experiments Software", 2012, <http://www.camo.com/rt/Products/Unscrambler/unscrambler.html> (July 2012).
- MathWorks, "MATLAB and Simulink for Technical Computing", 2012, <http://www.mathworks.com/index.html> (July 2012).
- Trajan Software, "Trajan Software Ltd", 2012, <http://www.trajan-software.demon.co.uk/> (July 2012).