Microchemical Journal 109 (2013) 122–127

Contents lists available at SciVerse ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

Classification of blue pen ink using infrared spectroscopy and linear discriminant analysis

Carolina Santos Silva [a], Flávia de Souza Lins Borba [a], Maria Fernanda Pimentel [b,*], Marcio José Coelho Pontes [c], Ricardo Saldanha Honorato [d], Celio Pasquini [e]

[a] Departamento de Química Fundamental, Universidade Federal de Pernambuco, Recife, PE, Brazil
[b] Departamento de Engenharia Química, Universidade Federal de Pernambuco, Recife, PE, Brazil
[c] Departamento de Química, Universidade Federal da Paraíba, João Pessoa, PB, Brazil
[d] Superintendência Regional em Pernambuco, Polícia Federal, Brazil
[e] Instituto de Química, Universidade Estadual de Campinas, Campinas, SP, Brazil

ARTICLE INFO

Article history:
Received 29 November 2011
Received in revised form 27 February 2012
Accepted 24 March 2012
Available online 30 March 2012

Keywords:
Infrared reflectance spectroscopy
Blue ink pen identification
Linear discriminant analysis
Successive projections algorithm
Genetic algorithm
Stepwise selection

ABSTRACT

Attenuated total reflectance (ATR) Fourier transform infrared (FTIR) spectroscopy associated to linear discriminant analysis (LDA) was employed to perform classification of blue pen ink according to types and brands, in a nondestructive way. To build a representative data set, blue pens of 3 types, namely ballpoint (5 brands), roller ball (2 brands) and gel (3 brands) were purchased from local dealers. Ten different pens, representing the best seller of each brand, were purchased, making a total of 100 pens. Circular areas were painted five times with each pen and spectra were taken in 2 different locations, using a Universal Attenuated Total Reflectance accessory (UATR), within the range of 4000 to 650 cm^{-1}. Three types of paper were employed: two brands of A4 sulfite paper (paper 1 and paper 2) and one recycled paper (paper 3). The genetic algorithm (GA), stepwise formulation (SW) and successive projections algorithm (SPA) were employed to select spectral variables employed in LDA. LDA models were built using the blue pen ink spectra obtained from paper 1. Three test sets were employed using the blue pen ink spectra obtained from papers 1, 2 and 3, in order to evaluate the influence of the paper on the predictions. The LDA models used to classify the pens according to their type (gel, rollerball and ballpoint) achieved a correct classification rate of 100% in the test set composed of blue pen ink spectra obtained from paper 1, using GA and SPA. Using SW, the rate achieved was 99.5%. For paper 2, SPA, GA and SW provided 100%, 97.3% and 93.8% of correct classification, respectively. For paper 3, SPA, GA and SW achieved a correct prediction rate of 100%, 100% and 94.9%, respectively. LDA models for classifications of pens according to their brand were 100% correct in their classification when the test set was composed of blue pen ink spectra obtained from papers 1 and 2. For the test set composed of blue pen ink spectra obtained from paper 3, LDA-SPA, LDA-GA and LDA-SW classified them correctly at 91.3%, 100% and 100%, respectively. The method developed was able to differentiate successfully all brands of pen used on each type of paper and could be a helpful tool for detection and confirmation of counterfeits in documents of legal importance.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Forensic document examination is a field in forensic science that studies the manipulation and falsification of documents, such as bank checks, medical certificates, working papers and passports. The evaluation of inks used to produce manuscripts can help in the identification of counterfeits [1]. Pen inks consist of complex systems, which require a combination of compounds to provide, not only the color, but also other characteristics associated to quality. For ink production, dyes, lubricants, resins, pigments, solvents, surfactants, emulsifiers and substances for pH control are usually employed.

The methods for ink analysis can be destructive or non-destructive. Using destructive methods, a portion of the ink has to be removed from the document using appropriate solvents. Different destructive analytical methods have been evaluated for the study of organic composition in pen inks, such as: UV–vis spectrometry [2–4]; Fourier transform infrared spectroscopy (FTIR) [4–6]; thin layer chromatography [4]; positive and negative-ion electrospray ionization mass spectrometry [7]; and high performance liquid chromatography (HPLC) [8]. Although these methods can identify some of the ink compounds, they destroy the document and cannot be applied in many forensic cases.

* Corresponding author at: Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Departamento de Engenharia Química. Rua Prof. Arthur de Sá S/N, Cidade Universitária, 50740-521, Recife, PE, Brazil. Tel.: +55 81 21268729; fax: +55 81 21267235.
E-mail addresses: mfernanda.pimentel@gmail.com, mfp@ufpe.br (M.F. Pimentel).

0026-265X/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.microc.2012.03.025

Non-destructive methods for forensic discrimination of pen inks have been reported mainly using Raman [5], UV-Visible, luminescence [9] and reflectance infrared [8] spectroscopies. Raman and reflectance infrared spectroscopy provides information about the chemical constituents of the inks, such as solvents, and resins, other than colorants.

Reflectance microspectrophotometry in the visible range is one of the standard non-destructive techniques used to discriminate among types of pen ink [10]. Forensic discrimination of ballpoint ink writing on paper is also routinely carried-out, non-destructively, using luminescence spectroscopy [3]. Zieba-Palus and Kunicki have shown that Raman spectrometry is valuable as a complementary technique to infrared spectrometry for analyzing the ink composition of 69 samples of blue and black inks originating from ballpoint pens and gel pens of various brands and manufacturers, commonly available in Poland [5]. Synchrotron-based FT-infrared micro-spectroscopy has been used by Wilkinson et al. for the comparative identification of writing inks [11]

Mass spectrometry has been also employed for non-destructive analysis of pen inks. Jones et al. reported on initial studies using mass spectrometry with an ion source called Direct Analysis in Real Time (DART) in situ for the non-destructive analysis of writing inks on paper without visible alteration [12]. Forty-three different black and blue ballpoint, black fluid, and black gel inks were examined. They have shown that mass spectrometry using DART produces spectra of sufficient quality and diversity to distinguish among all but the most similar of inks. Denman et al. [13] have shown that surface analysis by time-of-flight secondary ion mass spectrometry (TOF-SIMS) of blue ballpoint pen ink markings can non-destructively analyze organic and inorganic ink components directly with no interference from the paper substrate. Despite the fact that these techniques have been reported as non-destructive, a partial ink-extraction still occurs. Consequently, the documents are slightly modified.

Although research has been carried out to develop efficient analytical methods regarding the composition of inks, in most of the previously mentioned works, discrimination is carried-out by visual inspection to decide whether the two samples show the same spectra/chromatogram under the same experimental conditions [2]. Few papers describe the use of chemometric multivariate analysis to explore the analytical data efficiently and to enable unbiased decisions about the similarities among the ink samples.

Adam has shown that principal component analysis (PCA), followed by linear regression of the loadings, applied to a data set obtained from 10 black ballpoint pens using a standard luminescence spectrometer facilitates the separation of the ink luminescence from that of the paper and also allows a direct comparison between two ink lines drew on a document [9]. PCA has been applied to the UV–Vis spectra of inks obtained from black ballpoint pens available on the market in United Kingdom. For the complete set of 25 pens, interpretation of the loadings for the first few principal components showed that both the pen inks and the extracted (with ethanol) ink lines could be classified in an objective manner and in agreement with the results of parallel thin layer chromatography studies [3]. Thanasoulias et al. have employed cluster analysis (CA), PCA and discriminant analysis (DA) to UV–vis spectral data. Five commercially available brands of blue ball-point pen inks were used for the study. For each brand, each of 10 pens from the same batch were sampled by means of a stainless steel needle that was used to penetrate the wall of the plastic ink reservoir and transfer a small portion of the sample into the solvent (ethanol) [2]. This study showed a 100% correct classification of the training dataset between inks of different brands. The authors state that a more thorough study of the system could be made by using more ink samples, from different batches and that the models should also be tested on new samples, instead of items from the training dataset, to ensure a better assessment of the usefulness of the models. Using TOF-SIMS, organic and inorganic information taken from a total of 24 blue ballpoint pens (including replicates) were collected simultaneously and processed with PCA, identifying 41 out of 45 pairs (91%) of the pens analyzed [13]. Wang et al. developed a method for classifying blue ballpoint pen inks using FT-IR, that allowed a total of 108 samples to be divided into 35 subgroups using a pattern recognition system [6]. This method was found to be fast and reliable for bulk ink analysis, however, it has not yet been applied to identify inks on paper.

Kher et al. [8] applied PCA and linear discriminant analysis (LDA) to a data set composed of Micro-ATR infrared spectra obtained, non-destructively, from eight brands of six blue ballpoint pens each (totalizing 48 pens). HPLC analyses of the inks were also carried-out. On the IR data set, two approaches were used for data reduction before LDA analysis. The first approach consisted of using the PCA decomposed matrix while the second consisted of choosing a few discrete spectral features. The best result was obtained using the PCA scores which came up with a correct classification of only 62.5% of the pen ink samples. Applying LDA to the HPLC data set resulted in a correct classification of 97.9%.

The use of PCA scores or a few select wavenumbers, chosen after a priori consideration, as input variables for LDA, may cause loss of information relevant for classification. In order to overcome this problem, variable selection methods such as the genetic algorithm (GA) [14] and stepwise formulation (SW) [15] were employed to select the spectral variables for LDA in several classification problems. Recently, the successive projections algorithm (SPA) has been successfully applied to the classification of edible vegetable oils [16,17], diesel fuels [16], Brazilian soils [18], cigarettes [19], coffee [20] and diesel/biodiesel [21] samples.

The present paper proposes an analytical method based on infrared spectroscopy to classify blue inks from different types of pens (gel, rollerball and ballpoint pens), as well as inks from different brands for each type. For this purpose, the successive projections algorithm, genetic algorithm and stepwise formulation were employed to choose an appropriate subset of wavenumbers for a linear discriminant analysis model.

2. Materials and methods

2.1. Samples

Three different types of paper were employed: two brands of A4 print paper (paper 1 and paper 2) and one brand of recycled A4 paper (paper 3). To build a representative data set, blue pens of 3 types: ballpoint (5 brands), roller ball (2 brands) and gel (3 brands) were purchased commercially. Ten different pens, representing the best seller of each brand, were purchased: 5 from the same batch and 5 from different batches, making a total of 100 pens. Circular areas with a 1 cm radius were painted five times with each pen, on paper 1, and spectra were taken from 2 different locations of each. On paper 2 and 3, three pens of each brand were chosen from different batches and used to paint circular areas from which 5 spectra were taken.

2.2. Spectra acquisition

Spectra were acquired using *Spectrum400* (Perkin Elmer) equipment with a universal attenuated total reflectance (UATR) accessory, in the range 4000 to 650 cm^{-1}, with 4 cm^{-1} resolution and by averaging 16 scans. Spectra of each paper were also taken and subtracted from the ink pen spectrum. The integrity of the sample was completely preserved during spectra acquisition.

2.3. Classification models

LDA is a well-known technique for dimensionality reduction. In LDA, the dimensional embeddings are reduced in such a way that the

orientations of the projected data of classes on an arbitrary line or space are well-separated from each other. The criterion to determine these vectors is the maximization of the ratio of between-class variability and within-class variability in the training set. The number of training vectors must be larger than the number of feature vectors to be included in the LDA model. For high dimension data, such as in the IR spectral data set employed in this work, the use of appropriate variable selection procedures is usually required [22]. In this section, the three algorithms adopted for this purpose in the present paper (SPA, SW, and GA) is briefly described.

The successive projections algorithm (SPA) [23–25], originally proposed for spectral variable selection in multiple linear regression (MLR) models, was adapted by Pontes et al. [16] in the context of classification. For this, the best subset of variables is selected in order to minimize a cost function associated with the average risk G (Eq. (1)) of misclassification by LDA in a given validation set.

$$G = \frac{1}{Kv}\sum_{k=1}^{Kv} g_k \tag{1}$$

where g_k (defined in Eq. (2)) is the risk of misclassification of the k^{th} validation sample $\mathbf{x}_k$.

$$g_k = \frac{r^2(\mathbf{x}_k, \boldsymbol{\mu}_{Ik})}{\min_{Ij \neq Ik} r^2\left(\mathbf{x}_k, \boldsymbol{\mu}_{Ij}\right)} \tag{2}$$

where the numerator $r^2(\mathbf{x}_k,\boldsymbol{\mu}_{Ik})$ is the squared Mahalanobis distance between sample $\mathbf{x}_k$ (of class index Ik) and the sample mean $\boldsymbol{\mu}_{Ik}$ of its true class (both row vectors). The denominator in Eq. (2) corresponds to the squared Mahalanobis distance between sample $\mathbf{x}_k$ and the center of the closest wrong class.

The present work adopts the SW approach presented by Caneca et al. [15], in which the individual value of each spectral variable is evaluated according to its discriminability with respect to the classes under consideration. At each step, the spectral variable with the largest discriminability is selected, a *leave-one-out* cross-validation procedure is performed and the number of errors is noted. Before the next step, the variables that are highly correlated with those already selected are discarded to avoid collinearity problems. In this work, seven threshold values (0.10, 0.30, 0.50, 0.70, 0.80, 0.90, and 0.95) for the coefficient of multiple correlations were tested in order to decide which variables were to be discarded. The best threshold was selected on the basis of the number of classification errors in the validation set. More details of the stepwise formulation can be found in Ref. [15].

The genetic algorithm (GA) is a stochastic search technique that mimics the process of natural evolution [26,27]. This algorithm is routinely used to generate useful solutions to optimization and search problems. In the context of variable selection, the algorithm typically encodes subsets of variables in the form of strings of binary (0/1) values termed "chromosomes". Each gene in the chromosome is associated to one of the variables available for selection. When gene value is equal to 1, the corresponding variables (or wavenumbers) are included in the data set to produce the classification model. The evolution usually starts from a population of randomly generated individuals and evolves through generations. In each generation, the fitness of each individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. In this present work, a fitness value was defined for each chromosome as the inverse of the validation cost defined in Eq. (1), as in [16]. The GA routine was carried out over 100 generations with 200 chromosomes each. Mutation and crossover probabilities were set to 10% and 60%, respectively, previously suggested by Pontes et al. [16]. The whole process was repeated three times (starting from different random initial populations) and the best solution (in terms of the fitness value) resulting from the three realizations of the GA was selected.

Several pre-processing techniques were applied to the spectral data, such as Standard Normal Variate (SNV), Savitzky–Golay Smoothing, Savitzky–Golay first derivatives (3, 15 and 21 points windows, second order polynomial), normalization (range and maximum). A PCA was conducted after each pre-processing procedure to evaluate cluster separation according to each class. A PCA in each class was also performed in order to identify outliers. Spectra with high leverage and residue were excluded from the data set.

The Kennard Stone (KS) algorithm [28] was employed to select the samples which would compose the training (60% of samples), validation (20% of samples) and test (20% of samples) sets for paper 1. The training and validation samples were used in the modeling procedures (including variable selection for LDA), whereas the test samples were only used in the final evaluation and comparison of the classification models.

The classification models were built using only the registers made on paper 1. To evaluate the influence of the paper on predictions for the test set, spectra of three blue ink pens from each brand (selected randomly) were also obtained using different papers (papers 2 and 3). For each type of paper, new principal component analyses were performed to identify outliers. Spectra with high leverage and residue were excluded from the test set.

The final number of spectra used to compose the training, validation and test sets for each class is shown in Table 1.

Pre-processing and PCA were performed using The Unscrambler X.1. The LDA-SPA, LDA-GA, LDA-SW and KS algorithms were coded using Matlab® R2010a.

3. Results and discussion

3.1. Spectroscopic analysis

Fig. 1 shows the original average spectra for each of the ten classes of blue ink pen, obtained using paper 1. Spectra acquired with papers 1 and 2 are similar, but the spectra obtained using paper 3 differ, especially in the region between 1500 and 1300 cm^{-1}. Using ATR, the obtained spectra are composed by overlapping bands from the ink and mainly from the paper. [29]. Typical cellulose bands corresponding to : C–H and O–H stretching can be found between 3300 cm^{-1} and 2900 cm^{-1}; C–H angular bending is observed in the range 1500–1300 cm^{-1}; and C–O stretching is present at proximally 1030 cm^{-1}. Furthermore, the infrared spectra of print papers may contain spectral features produced by the presence of additives, such as calcium carbonate. The principal absorption peaks can be found around 1430, 875 and 712 cm^{-1} associated with this substance.

Table 1
Number (N) of training, validation and test samples in each class.

Class			Paper 1			Paper 2	Paper 3
Type	Brand	*N*	Training	Validation	Test	Test	Test
Ballpoint		499	299	100	100	74	68
	Bp_bc	100	60	20	20	15	15
	Bp_cp	100	60	20	20	14	15
	Bp_mt	99	59	20	20	15	15
	Bp_pe	100	60	20	20	15	12
	Bp_pi	100	60	20	20	15	11
Gel		291	175	58	58	45	41
	Gel_bc	100	60	20	20	15	15
	Gel_fc	91	55	18	18	15	13
	Gel_mo	100	60	20	20	15	13
Rollerball		193	115	39	39	27	29
	Rb_bc	95	57	19	19	15	14
	Rb_cp	98	60	19	19	12	15

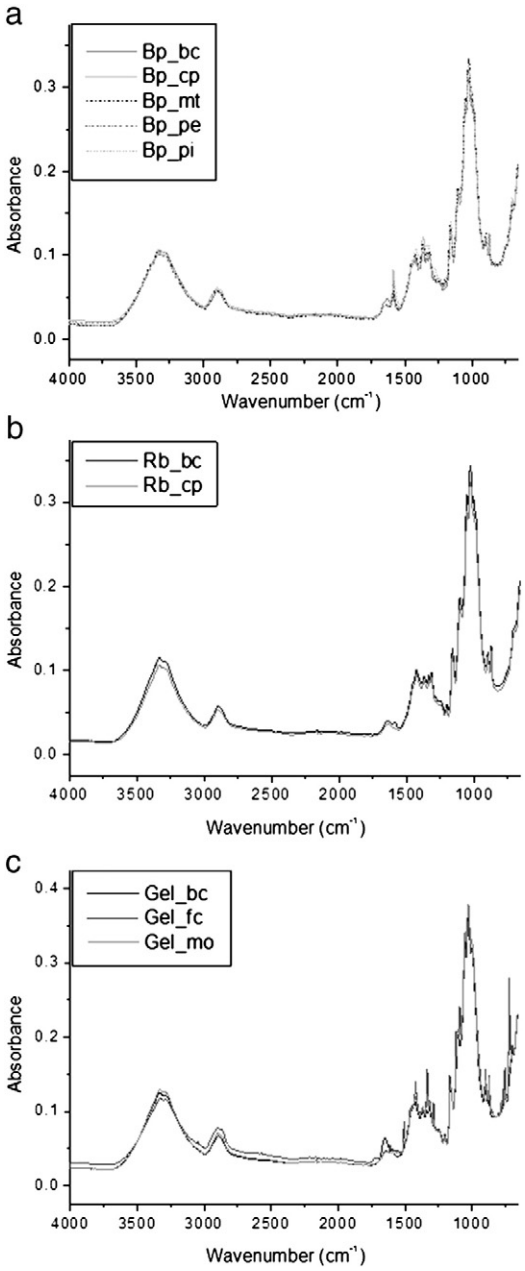

Fig. 1. Average spectra of the different brands of (a) ballpoint, (b) rollerball, (c) gel blue ink pen.

Other additives are present in minor quantities and should not produce a relevant effect of the paper spectra. The most important region for ink analysis is from 1800 to 650 cm^{-1}, where typical absorption bands can be found.

Among the pre-processing techniques tested, the one showing better separation of the classes employing PCA was the combination of standard normal variate, first Savitzy–Golay derivative (21 points window, second degree polynomial and range normalization). The score plot (PC1 × PC2 × PC3) for all blue ink pen spectra obtained, after preprocessing and paper subtraction (paper 1), is shown in Fig. 2. The first three components explain 73% of the total variance. There is an evident overlapping among the spectra from rollerball and gel ink pens.

3.2. LDA performed using blue ink on paper 1 spectra

For SPA selection procedure, 17 and 39 variables were incorporated in the model, when LDA was used to classify types and brands,

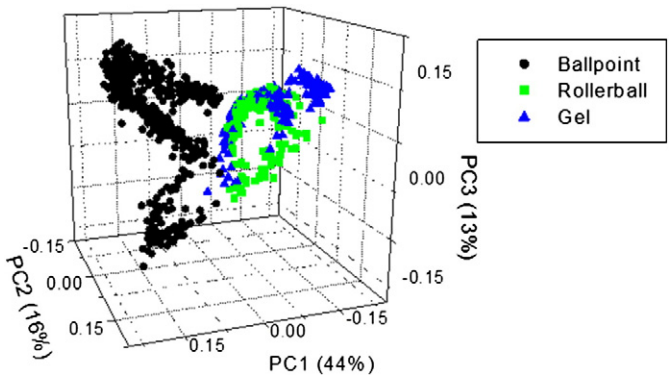

Fig. 2. Score plot for blue ink pen spectra after paper subtraction and pre-processing.

respectively. In the case of GA selection, 15 and 59 variables were selected (types and brand classes, respectively). The stepwise algorithm selected 62 variables (threshold of 0.70) for pen classification by type and 67 variables (threshold of 0.7) for classification by brands.

Fig. 3a and b show the variables selected by SPA, when the classification models were built for type and for brands, respectively. The variables selected by GA for type and brand classification models are presented in Fig. 3c and d, respectively. For SW, the variables are presented in Fig. 3e and f. As can be seen, SW algorithm selected many variables that are non-informative.

Table 2 shows the classification results for the test set for LDA-SPA, LDA-GA and LDA-SW models. The LDA produces a number of functions equal to the total number of classes of samples minus 1. Thus, the LDA model for types and brands generated respectively 2 and 9 discriminant functions. Although the LDA-SW model provided a rate of correct classification of 99.5% by type and 100% by brands, LDA-SPA and LDA-GA models built for ink pen classification by type, as well as for brands, provided 100% of correct classification for the test set. LDA-SPA model for classification by brand, however, uses fewer variables then GA. All these results are superior of those obtained by Kher et al. [8], using PCA and LDA which showed a correct classification rate of only 62.5% of the pen samples, using only one type of paper.

To illustrate, Fig. 4a shows the score plot of the first two discriminant functions for the LDA-SPA classification by type. Fig. 4b presents the first three discriminant functions for ink pen classification by brand. Obviously, 9 functions are necessary to discriminate ten brand classes. Nevertheless, Fig. 4b shows the scenery after the first three functions.

3.3. Predictions for the test set composed by blue ink spectra obtained on papers 2 and 3

For classification of blue ink pen by type, the LDA-SPA, LDA-GA and LDA-SW reached a rate of 100%, 97.3% and 93.8% respectively (Table 2). The samples that were misclassified belong to the rollerball class (brand *rb_cp)*, which could not be distinguished from samples of gel class. LDA-SPA, LDA-GA and LDA-SW models built for classification of the brand class, using the blue ink pen spectra obtained from paper 2, attained a correct prediction rate of 100% using the test set.

For the test set composed of blue ink pen spectra obtained on paper 3 (recycled paper), LDA-SPA, LDA-GA and LDA-SW models achieved a correct prediction rate of 100%, 100% and 94.9%, when the classification by type is performed. As for paper 2, the samples that were misclassified, using LDA-SW, belong to the rollerball class (brand *rb_cp)*, which could not be distinguished from samples of gel class. For brand classification, a correct classification rate of 91.3% was obtained for LDA-SPA. In this case, twelve errors were found, all

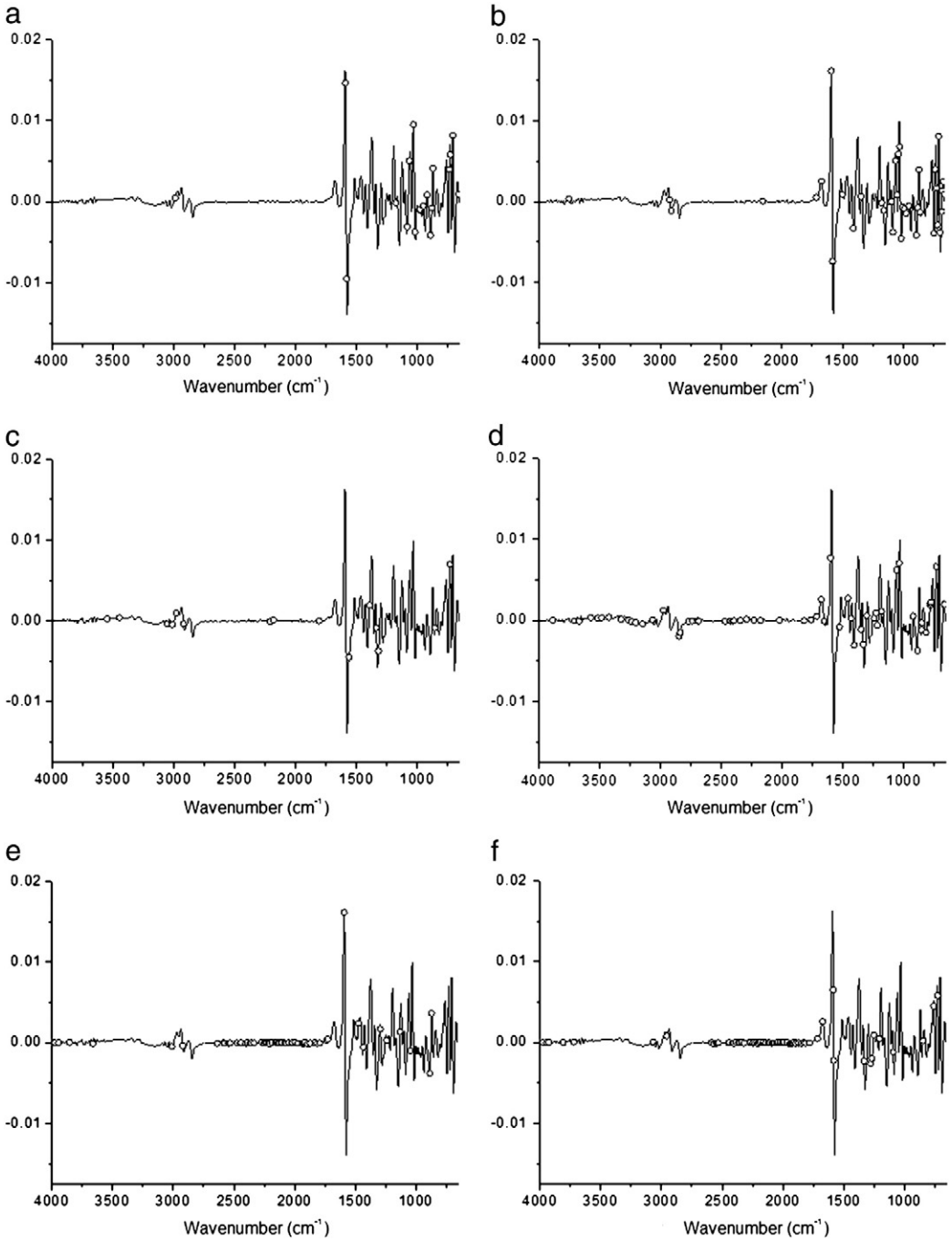

Fig. 3. Mean spectra (after paper spectra subtraction and pre-processing) with wavenumbers selected by SPA for (a) type and (b) brand LDA classification models; by GA (c) type and (d) brand LDA classification models; and by SW (e) type and (f) brand LDA classification models.

evolving *es_cp* ballpoint pens (12 *es_cp* ballpoint pen samples were classified as *es_pe* ballpoint pens). For both LDA-GA and LDA-SW models a correct brand classification rate of 100% was obtained.

Table 2
Classification results for the test set.

Percent of correct classification (%)		LDA-SPA	LDA-GA	LDA-SW
Paper 1	Types	100	100	99.5
	Brands	100	100	100
Paper 2	Types	100	97.3	93.8
	Brands	100	100	100
Paper 3	Types	100	100	94.9
	Brands	91.3	100	100

4. Conclusions

Infrared Spectroscopy associated with linear discriminant analysis, using SPA, GA and SW algorithms for variable selection were successful employed to classify blue ink pens by type (gel, rollerball and ballpoint) and by brands. Using an independent test set composed of ink spectra obtained on similar sulfite papers (paper 1 and 2), a correct classification rate of 100% was obtained by brand, independent of the variable selection algorithm. By type, LDA-GA and LDA-SW models did not show the same efficiency as for brand classification. The LDA-SW provided rates of 99.5% and 93.8% for papers 1 and 2, respectively. And LDA-GA model provided 100% and 97.3% for papers 1 and 2.

For the test set composed of ink spectra obtained on recycled paper (paper 3), LDA using the algorithms SPA, GA and SW provided a

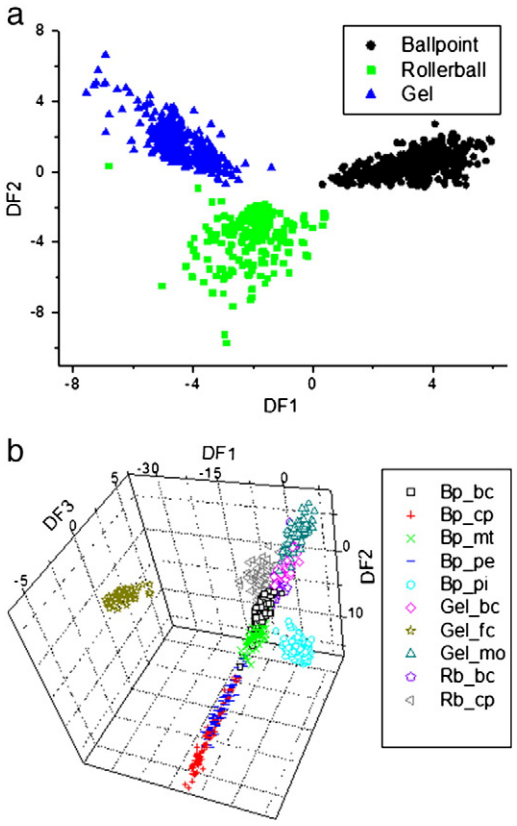

Fig. 4. LDA-SPA (a) DF2 × DF1 score plots for classification of ink pen by type (b) DF3 × DF2 × DF1 score plots for classification of ink pen by brand.

correct classification rate of 100%, 100% and 94.9%, when classification is performed for type of ink blue pen. Predictive testing performed for classification by brand showed a 100% correct classification for GA and SW and a 91.3% correct classification for LDA.

In general, LDA-SPA and LDA-GA showed a better performance than LDA-SW models.

The proposed method contributes to increasing the technological level of crime laboratories by allowing forensic experts a faster and unbiased interpretation of the evidence. It consists of an important tool for the detection and confirmation of counterfeiting in documents of legal importance, by comparing the inks used to produce the handwriting in different documents or sections of the same document.

Acknowledgments

This work was supported by FACEPE, FINEP, CNPq, CAPES and INCTAA.

References

[1] J.M.G. Góngora, I. Maguregui, R. Alonso, Analytical methods for dating modern writing instrument inks on paper, Magdalena Ezcurra, Forensic Sci. Int. 197 (2010) 1–20.
[2] N.C. Thanasoulias, N.A. Parisis, N.P. Evmiridis, Multivariate chemometrics for the forensic discrimination of blue ball-point pen inks based on their Vis spectra, Forensic Sci. Int. 138 (2003) 75–84.
[3] C.D. Adam, S.L. Sherratt, V.L. Zholobenko, Classification and individualization of black ballpoint pen inks using principal component analysis of UV–vis absorption spectra, Forensic Sci. Int. 174 (2008) 16–25.
[4] V. Causin, R. Casamassima, C. Marega, P. Maida, S. Schiavone, A. Marigo, A. Villari, The discrimination potential of ultraviolet–visible spectrophotometry, thin layer chromatography, and Fourier transform infrared spectroscopy for the forensic analysis of black and blue ballpoint inks, J. Forensic Sci. 53 (2008) 1468–1473.
[5] J. Zieba-Paulus, M. Kunicki, Application of micro-FTIR spectroscopy, Raman spectroscopy and XRF method examination of inks, Forensic Sci. Int. 158 (2006) 164–172.
[6] J. Wang, G. Luo, S. Sun, Z. Wang, Y. Wang, Systematic analysis of bulk blue ballpoint pen ink by FTIR spectrometry, J. Forensic Sci. 46 (2001) 1093–1097.
[7] L. Ng, P. Lafontaine, L. Brazeau, Ballpoint pen inks: characterization by positive and negative ion-electrospray ionization mass spectrometry for the forensic examination of writing inks, J. Forensic Sci. 47 (2002) 1238–1247.
[8] A. Kher, M. Mulholland, E. Green, B. Reedy, Forensic classification of ballpoint pen inks using high performance liquid chromatography and infrared spectroscopy with principal component analysis and linear discriminate analysis, Vib. Spectrosc. 40 (2006) 270–277.
[9] C.D. Adam, In situ luminescence spectroscopy with multivariate analysis for the discrimination of black ballpoint pen ink-lines on paper, Forensic Sci. Int. 182 (2008) 27–34.
[10] C. Rouxa, M. Novotnyb, I. Evansc, C. Lennardd, A study to investigate the evidential value of blue and black ballpoint pen inks in Australia, Forensic Sci. Int. 101 (1999) 167–176.
[11] T.J. Wilkinson, D.L. Perry, M.C. Martin, W.R. McKinney, A.A. Cantu, Use of synchrotron reflectance infrared spectromicroscopy as a rapid, direct, nondestructive method for the study of inks on paper, Appl. Spectrosc. 56 (2002) 800–803.
[12] R.W. Jones, R.B. Cody, J.F. McClelland, Differentiating writing inks using direct analysis in real time mass spectrometry, J. Forensic Sci. 51 (2006) 915–918.
[13] J.A. Denman, W.M. Skinner, K.P. Kirkbride, I.M. Kempson, Organic and inorganic discrimination of ballpoint pen inks by ToF-SIMS and multivariate statistics, Appl. Surf. Sci. 256 (2010) 2155–2163.
[14] L.H. Chiang, R.J. Pell, Genetic algorithms combined with discriminant analysis for key variable identification, J. Process. Control. 14 (2004) 143–155.
[15] A.R. Caneca, M.F. Pimentel, R.K.H. Galvão, C.E. Matta, F.R. Carvalho, I.M. Raimundo Jr., C. Pasquini, J.J.R. Rohwedder, Assessment of infrared spectroscopy and multivariate techniques for monitoring the service condition of diesel-engine lubricating oils, Talanta 70 (2006) 344–352.
[16] M.J.C. Pontes, R.K.H. Galvão, M.C.U. Araújo, P.N.T. Moreira, O.D.P. Neto, G.E. José, T.C.B. Saldanha, The successive projections algorithm for spectral variable selection in classification problems, Chemom. Intell. Lab. Syst. 78 (2005) 11–18.
[17] F.F. Gambarra-Neto, G. Marino, M.C.U. Araujo, R.K.H. Galvao, M.J.C. Pontes, E.P. Medeiros, R.S. Lima, Classification of edible vegetable oils using square wave voltammetry with multivariate data analysis, Talanta 77 (2009) 1660–1666.
[18] M.J.C. Pontes, J. Cortez, R.K.H. Galvão, C. Pasquini, M.C.U. Araújo, R.M. Coelho, M.K. Chiba, M.F. Abreu, B.E. Madari, Classification of Brazilian soils by using LIBS and variable selection in the wavelet domain, Anal. Chim. Acta 642 (2009) 12–18.
[19] E.D. Moreira, M.J.C. Pontes, R.K.H. Galvão, M.C.U. Araújo, Near infrared reflectance spectrometry classification of cigarettes using the successive projections algorithm for variable selection, Talanta 79 (2009) 1260–1264.
[20] U.T.C.P. Souto, M.J.C. Pontes, E.C. Silva, R.K.H. Galvão, M.C.U. Araújo, F.A.C. Sanches, F.A.S. Cunha, M.S.R. Oliveira, UV–Vis spectrometric classification of coffees by SPA–LDA, Food Chem. 119 (2010) 368–371.
[21] M.J.C. Pontes, C.F. Pereira, M.F. Pimentel, F.V.C. Vasconcelos, A.G.B. Silva, Screening analysis to detect adulteration in diesel/biodiesel blends using near infrared spectrometry and multivariate classification, Talanta 85 (2011) 2159–2165.
[22] R.A. Fisher, The use of multiple measurements in taxinomic problems, Ann. Eugenics 7 (1936) 179–188.
[23] M.C.U. Araújo, T.C.B. Saldanha, R.K.H. Galvão, T. Yoneyama, H.C. Chame, V. Visani, The successive projections algorithm for variable selection in spectroscopic multicomponent analysis, Chemom. Intell. Lab. Syst. 57 (2001) 65–73.
[24] R.K.H. Galvão, M.C.U. Araújo, in: S.D. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive Chemometrics: Chemical and Biochemical Data Analysis, Elsevier, Oxford, 2009, p. 233.
[25] M.J.C. Pontes, A.M.J. Rocha, M.F. Pimentel, C.F. Pereira, Determining the quality of insulating oils using near infrared spectroscopy and wavelength selection, Microchem. J. 98 (2011) 254–259.
[26] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley Longman Publishing Co., Inc, Boston, 1989.
[27] R. Leardi, Genetic algorithms in chemometrics and chemistry: a review, J. Chem. 15 (2001) 559–569.
[28] R.W. Kennard, L.A. Stone, Computer-aided design of experiments, Technometrics 11 (1969) 137–148.
[29] Barbara Stuart, Infrared Spectroscopy: Fundamentals and Aplications, Wiley, Chichester, 2004, p. 179.