

SYSTEMATIC ASSESSMENT OF ATTENUATED TOTAL REFLECTANCE-FOURIER TRANSFORMS INFRARED SPECTROSCOPY COUPLED WITH MULTIVARIATE ANALYSIS FOR FORENSIC ANALYSIS OF BLACK BALLPOINT PEN INKS

(Penilaian Bersistem Spektroskopi Mikro-ATR-FTIR yang Digabungkan dengan Analisis Multivariat untuk Analisis Forensik Pen Mata Bulat Dakwat Hitam)

L. C. Lee^{1,2}, Mohamed Rozali Othman^{1*}, and H. Pua²

¹*School of Chemistry Sciences and Food Technology,
Faculty of Science & Technology,*

Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor D.E. Malaysia;

²*Program of Forensic Science, Faculty of Health Science,
Universiti Kebangsaan Malaysia, Jalan Raja Muda Abdul Aziz, 50300 Kuala Lumpur, Malaysia*

*Corresponding author: rozali@ukm.my

Abstract

This manuscript aims to provide a new and non-destructive method for systematic analysis of inks on a questioned document. Ink samples were analyzed *in situ* on the paper substrate by micro-ATR-FTIR spectroscopy and the data obtained was processed and evaluated by a series of multivariate chemometrics. Absorbance value from wavenumbers of 2000-675 cm⁻¹ were first processed by cluster analysis (CA), followed by principal component analysis (PCA) to form a set of new variables. Subsequently, the variables set was used for classification, differentiation and identification of 155 sample pens that comprise nine different brands. Results show that nine black ballpoint pen brands could be classified into three main groups via discriminant analysis (DA). Differentiation analyses of nine different pen brands performed using one-way ANOVA indicated only two pairs of brands cannot be differentiated at 95% confidence interval. Finally an identification flow chart was proposed to determine the brand of unknown pen inks. As a conclusion, the proposed method for extracting and creating a new variable set from infrared spectrum was evaluated to be satisfactory for systematic analysis of inks based on their infrared spectrum.

Keywords: Forensic science; ATR-FTIR spectroscopy; classification; differentiation; ink analysis; multivariate analysis

Abstrak

Manuskrip ini bertujuan membekalkan satu kaedah yang baru dan tidak membinasa untuk analisis sistematik dakwat pada dokumen yang dipersoalkan. Sampel dakwat telah dianalisis secara *in situ* pada substrat kertas dengan spektroskopi mikro-ATR-FTIR dan data yang diperolehi telah diproses serta dinilai melalui satu siri kemometrik multivarians. Nilai serapan dari nombor gelombang 2000-675 cm⁻¹ telah diproses dahulu dengan analisis pengelompokan (CA), diikuti dengan analisis komponen utama (PCA) untuk membentuk satu set pembolehubah baru. Kemudian, set pembolehubah berkenaan telah digunakan untuk klasifikasi, pembezaan dan pengenalan 155 pen sample yang terdiri daripada sembilan jenama yang berlainan. Keputusan menunjukkan bahawa sembilan jenama pen mata bulat dakwat hitam boleh diklasifikasi ke dalam tiga kumpulan utama melalui analisis diskriminasi (DA). Analisis pembezaan antara sembilan jenama pen yang berlainan dilaksanakan menggunakan ANOVA satu hala menunjukkan hanya dua pasangan jenama tidak dapat dibezakan pada selang keyakinan 95%. Akhirnya, satu carta alir pengenalan telah dicadangkan untuk menentukan jenama dakwat pen yang tidak diketahui. Sebagai kesimpulan, kaedah yang dicadangkan untuk mengekstrak dan mencipta satu set pembolehubah baru dari spektrum inframerah telah dinilai sebagai memuaskan untuk analisis sistematik dakwat berdasarkan spektrum inframerah mereka.

Kata kunci: sains forensik; spektroskopi ATR-FTIR; klasifikasi; pembezaan; analisis dakwat; analisis multivariat

Introduction

Since the advent of the 20th century, ink analysis has been one of the main focuses of questioned document examination. In brief, ink analysis involves analytical techniques to provide information that assist in identification of the source of questioned ink. Identification refers to the process of determination of the brand, model, manufacturer and year of production of the instrument(s) that could have produced the ink specimen [1].

Recently, destructive techniques such as high performance thin layer chromatography (HPTLC) [1], high performance liquid chromatography (HPLC) [2], time-of-flight secondary ion mass spectrometry (TOF-SIMS) [3], and laser desorption ionization-mass spectrometer (LDI-MS) [4] have been studied extensively with the aim to establish an efficient and effective analytical method to identify ballpoint pen inks. While non-destructive techniques included Fourier transform infrared (FT-IR) spectroscopy using an Attenuated Total Reflectance objective (micro-ATR-FTIR) has not been investigated and exploited thoroughly. Micro-ATR-FTIR utilizes an attenuated total reflection (ATR) accessory which operates by measuring the changes that occur in a totally internally reflected infrared beam when the beam comes into contact with a sample. An infrared beam is directed onto an optically dense crystal into the sample held in contact with the crystal. This evanescent wave protrudes only a few microns beyond the crystal surface and into the sample. FTIR-ATR has advantage over FTIR-KBr in that sample preparation is quicker and less messy, user-to-user spectra variation is minimized and the impact of sample preparation on results is reduced [5]. However, infrared (IR) spectroscopy was known to not be able to produce as much information as those destructive techniques such as HPLC and inductively coupled plasma mass spectrometry (ICP-MS). As such, multivariate analysis can be used to increase the value of non-destructive technique as a good analytical method for ink analysis. Multivariate analysis is a powerful tool when dealing with multi-component systems and offers an interpretative methodology for discerning significant differences from complicated datasets, e.g. IR spectrum of ink [6].

Multivariate analysis such as cluster analysis (CA), principal component analysis (PCA) and discriminant analysis (DA) have been used for classification and discrimination of soil [6], document paper [7], ivory [8] and ballpoint pen inks [2, 9] by using spectra from IR, Raman and UV-visible spectroscopy. As those spectra usually consists over hundreds of data points, there is a need to reduce the number of data points so as to improve the performance of DA while reducing training and utilization times [10]. All of the aforementioned studies [2, 6, 7, 8, 9] employed different feature selection/extraction methods. Brody et al. [8] used eleven integrated areas calculated from Raman spectrum as input variables for DA. Thanasoulis et al. [6, 9] used *K*-means cluster analysis to assist in selection of few raw variables that will then be further reduced by PCA. On the other hand, Kher et al. [2, 7] decomposed the complete spectrum by PCA without any prior selection. The resulting PCA scores were used as DA variables. In addition, Kher et al. [2, 7] also suggested selecting discrete variables based on the ratio of the between-to-within-variability.

The main aim of this study is to explore the potential of micro-ATR-FTIR spectroscopy coupled with multivariate analysis as a non-destructive, fast and easy approach for systematic analysis of black ballpoint pen inks. On one hand, micro-ATR-FTIR has shown its ability to differentiate red seal inks [11] and blue ballpoint pen inks [2]. On the other hand, although the use of multivariate analysis with IR spectrum should be avoided due to multicollinearity problems, literature has demonstrated that such problems could be overcome by applying certain data extraction/reduction methods. In this study, three different data-reduction techniques will be proposed to create a new variable set from IR spectrum of inks. Subsequently, the new set of variables will be evaluated for classification, differentiation and identification of pen inks to determine its feasibility for forensic ink analysis.

Materials and Methods

Samples

A total of 155 black ballpoint pens of nine different brands were sampled from stationery shops in Malaysia. All of these were obtained in multiple packs of the same product to ensure all of them were from the same production batches as the variation between batches will not be considered in this study. All pen brands were allocated a reference number for the purpose of this study. Detailed descriptions of collected samples used in this study are presented in Table 1. Double ATM, from Thailand was the white paper used throughout the study. Each of the 155 pens was used to draw three different small circles to ensure complete coverage of paper by the pens. Experimental

work was carried out on the ink deposited on the white paper not later than one day after drawing to minimize changes due to ink aging.

Table 1. Reference number for all nine brand of list of studied black ballpoint pens

Reference number	Brand	Model	Quantity
1	Bic	BU3 0.8	4
		BU2 0.7	3
		RS 2	4
		NS Med	3
		Org NS fine	3
		Round Stic	4
2	Faster	CX813	1
		CX818	4
		CX1006	4
		CX838	4
		CX608	3
		828	5
3	Faber-Castell	Super click	4
		Ballpen 1.0	4
		Ballpen 0.7	4
		Click ball	4
		Needle ball	4
		Ballpen 0.5	4
4	G-soft	gs 66	3
		gs 77	3
		r 100	4
5	MGM	BP 713	6
		e-Rite	4
6	PILOT	Super GP	4
7	Paper Mate	KM 100 Med	4
		KM 100 Fine	4
		KM 100 S/F	4
		KV 2 Med	4
		KV 2 Fine	4
		KV 2 S/F	4
		KM Med PT	4
		KM Fine PT	5
		KM F	4
8	Stabilo	Liner 808 M	4
		Liner 808 F	4
		Liner 308	3
		Marathon 318	2
		Galaxy 818 M	5
9	UNicorn	BP 318 GS	5
		BP 333 DT	5
		Total	155

Micro-ATR-FTIR Analysis

All experimental spectroscopy was carried out on a Thermo Scientific Nicolet iN10 MX FT-IR microscope with mercury cadmium telluride (MCT) detector. A Ge crystal tip ATR objective was used as micro-sampling accessory. The background spectrum was reacquired after every analysis to reduce variation in the spectra due to instrumental drift. Each spectrum was the result of an average 16 scans at 4 cm^{-1} resolution over a spectral range of 4000 to 675 wavenumbers (cm^{-1}). Three spectrums were collected from each of 155 pens. The IR spectral data was stored in a data matrix in Microsoft Excel spreadsheet.

Software

Data collected was processed and analyzed using statistical package SPSS (Statistical Package for the Social Sciences, Windows version 15.0, SPSS Inc., Chicago, USA).

Results and Discussion

Sampling

In general, three sources of variations present in pen ink samples are: (a) variation between groups; (b) variation within a group (differences between different individual pens of a single group) and (c) variation between replicates within a single individual pen (errors inherent to the method of analysis). Two groups can be considered different if variation (a) is significantly bigger than variation (b) and (c). On the other hand, aforementioned sources of variations could be present at four different levels including batches of production, brands, models and ballpoint sizes. Most studies found across literatures draw only one sample to represent a particular group and seldom take into account variation (b) and (c). This study will consider all three sources of variations simultaneously. As such, we can evaluate the method more thoroughly and explicitly distinguish the differences between each group in more detail. Finally the resulting conclusion can be applied to the population rather than to the sub-population. In this study, the evaluation of new set of variables for classification, differentiation and identification of pen inks at brand-level will be considered first. In other words, the variation between groups refers to differences between pens of different brands (inter-brand variability). To encounter all possible variations within a brand (intra-brand variability), various models from that particular brand were drawn. For determination of variation (b) and (c), each brand was represented by more than one individual pen and results of triplicate analysis were treated as individual reading.

Infrared Spectrum

Although all the spectra were scanned from $4000\text{--}675\text{ cm}^{-1}$, only the regions from $2000\text{--}675\text{ cm}^{-1}$ were analyzed. The IR absorption peaks at the region between 4000 and 2000 cm^{-1} and this was mainly due to water vapor and carbon dioxide from the atmosphere and only a few weak and broad bands attributable to ink [2]. The respective spectrums of nine different brands of black ballpoint pen inks analyzed by micro-ATR-FTIR spectroscopy in the region of $2000\text{--}675\text{ cm}^{-1}$ are shown in Figure 1. In general, all spectra from nine different brands were showing similar spectrum patterns. All exhibit two prominent peaks at approximately 1584 and 1360 cm^{-1} but with differences in the form of relative peak height and its shape (Figure. 1). They are characteristic peaks of ballpoint pen inks, and have previously been assigned to a skeletal vibration of triarylmethane dye and the C=C stretch vibration of epoxy resin (about 1581 cm^{-1}) [12]. The region between wavenumbers $1100\text{--}1000\text{ cm}^{-1}$ is of less concern as it contains peaks mostly from paper substrate [13]. The vibrational assignment for other peaks has been explained elsewhere [12]. As such, the quantity of crystal violet may play an important role in the differentiation of those nine pen brands.

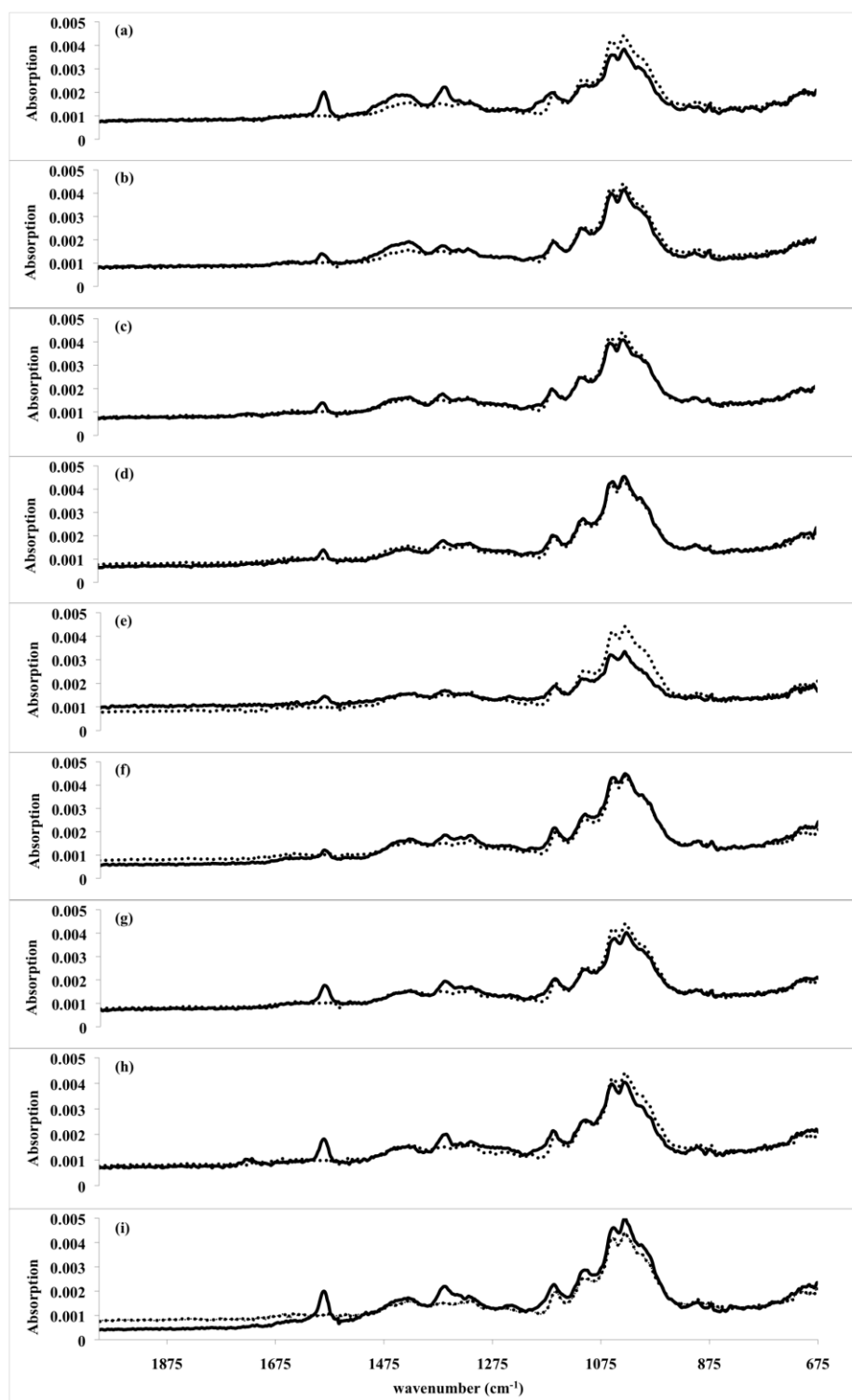


Figure 1. The respective spectral of nine different brands of black ballpoint pen inks (—) and blank paper (....) analyzed by micro-ATR-FTIR spectroscopy in the region of 2000-675 cm^{-1} : (a) Bic; (b) Faster; (c) Faber-castell; (d) G-soft; (e) MGM; (f) Pilot; (g) Paper-mate; (h) Stabilo; (i) UNICorN

Statistical Pretreatment

Normalization of selected IR region

All three spectra of each sample were included to assess the material reproducibility and homogeneity while ensuring representativeness of each sample due to the minimal size of inks contacted by ATR Ge tip and the potential heterogeneity of the inks [8]. Triplicate analysis of 155 pens gave 464 spectra as one replica spectrum of sample 11 was lost. As such, full data set comprised 464 spectra, consisting of absorbance values as a function of wavenumber, were normalized prior to carrying out the multivariate analysis to reduce variation in the data due to different thickness of the pen inks deposited on paper samples. The absorbance values for each spectrum were divided by the total absorbance resulting from each spectrum across wavenumbers 2000-675 cm^{-1} .

Transformation of raw data to new set of variables

In this study, the selected IR region (2000-675 cm^{-1}) composed of 688 raw data. Three different data extraction techniques were proposed to transform the normalized absorbance values into a new set of variables. This is an important step as discriminant analysis (DA) is restrictive with regard to the number of predictor variables versus the number of samples. In addition, a smaller number of input variables also means results can be obtained with reduced computational and economical expense while improving the accuracy and efficiency of the classification tasks [14].

The first approach is simply choosing few discrete representative variables without involving PCA. Absorbance values of wavenumbers that make up the three characteristic IR peaks of crystal violet: 1171, 1173, 1174, 1176, 1178, 1360, 1362, 1363, 1367, 1581, 1583, 1587 and 1589 cm^{-1} were selected as a new set of variables. This approach labeled as “discrete data-DA”. The second approach named “PCA-DA” involved the decomposition of the whole 688 data by correlation PCA. The number of components to be extracted were decided based on the Kaiser criterion [15]. The resulting principal components act as another new set of variables. The third feature reduction technique which labeled as “CA-PCA-DA” was achieved by conducting cluster analysis and principal component analysis sequentially. Initially, cluster analysis was performed on the raw variables (688 data points) over the objects (ink samples) by using Ward’s method and squared Euclidean distance to form eight different clusters. Each cluster is expected to contain variables that carry similar information about the objects. Subsequently, correlation PCA was performed on each of the eight clusters, separately. The number of components to be extracted was decided based on the Kaiser criterion [15]. The new set of variables comprised all selected principal components that were labeled accordingly, e.g. PCA 1_2 refers to the first component produced by PCA conducted on the second cluster of data points.

Systematic Analysis

Classification via discriminant analysis for the full dataset

The selected IR region (2000-675 cm^{-1}) was subjected to three different types of data-reduction techniques. After the feature reduction step, stepwise DA was conducted on each of the three aforementioned-processed dataset to see whether all 464 IR spectral could be discriminated according to their pen brands (category variables). DA was used with leave-one-out classification in the training set. Leave-one-out is a type of cross-validation techniques that enables all of the available data to be utilized for training while still giving an unbiased estimate of the generalization capabilities of the resulting classifier [14]. Stepwise DA was performed by using the Mahalanobis distance method and probability- F as the criterion. Prior probabilities were equal for all groups (brands) as the representativeness of population was unknown [16].

The use of selected discrete variables (the first approach) resulted in the lowest correct classification (24.80 % correct classification) while the decomposition of the whole spectrum by once PCA (the second approach) resulted in 45.30 % correct classification, higher than the first approach. A total of 58.60 % of ink samples were correctly classified when the same dataset was analyzed by several PCA according to eight different clusters formed from it. Therefore, the CA-PCA-DA approach (third approach) was selected as the best approach. There is a few advantages behind the third approach. Firstly, the newly formed components will be more informative than a single wavenumber as principal component is a linear combination of several wavenumbers. PCA conducted on each of the clusters of data points of IR spectrum grouped according to similarity is expected to be more effective in extracting important information from IR spectrum than PCA conducted on whole data points of spectrum. This

also helps in reducing the risk of losing important information. In addition, PCA running on a smaller number of data points would result in a more stable matrix than those running on whole 688 data points all at once. On the other hand, PCA scores are proven to be more effective than raw variables for DA [6, 9]. In addition, by using extracted components as new latent variables, the problem of redundancy is eliminated as the components are orthogonal and uncorrelated.

The CA-PCA-DA approach reduced 688 data points to 36 principal components. Stepwise DA selected 26 components from 36 components as predictor variables. Table 2 presents the first four variables that have the highest canonical discriminant function coefficients with the first two discriminant functions. PCA 1_3 and PCA 3_3 contributed most to the first discriminant function while the second discriminant function was strongly influenced by PCA 2_2 and PCA 2_7. All of these four components contain bands assigned to the aromatic ring vibrations of crystal violet, the primary dye of most black ballpoint pen inks [12].

Table 2. List of first four variables (components) that have highest coefficient with first two discriminant functions and wavenumber(s) that loaded >0.400 onto the respective component/ the first four variables that have highest canonical discriminant function coefficients with first two discriminant functions.

Component	Function		Wavenumber with component loading > 0.400 to component
	1	2	
PCA 2_2	-0.08	-1.09	1593-1578; 1383-1375; 1348-1329; 1304-1300; 1188-1174; 1144-1138
PCA 1_3	1.00	0.02	928-775
PCA 3_3	1.48	0.37	1254-1190; 831-823
PCA 2_7	0.31	1.04	1616-1599; 1574-1522

This indicates the dyestuff components in the ink were responsible for producing maximum variability in the data. Figure 2 shows the projections of the 464 spectra into the space of the first two discriminant functions, displaying 76.19% of the between-to-within group variation in the data. Based on Figure 2, discriminant function 1 was found to be responsible for separation of group 1, 2, 3, 4, 6, 7, and 8 from group 5 as well as 9 while group 8 was separated from the others by discriminant function 2.

As such, all the pen brands may be categorised into three groups according to the LDF 1-2 scores. The first, second and third clusters were composed of brand 5 and 9; brand 3 and 8; brand 1, 2, 4, 6 and 7, respectively. In leave-one-out cross-validation, the overall classification success is relatively poor (58.6%). However, five out of nine groups were able to give correct classification over 70% which is an acceptable level in many studies in pattern recognition [17]. This indicated a reasonable level of differentiation does exist among studied pen brands.

Differentiation analyses

Pair-wise comparison was conducted to determine whether all nine pen brands may be discriminated through the 36 predictor variables. With nine different brands, 36 $((1/2)K(K-1))$ possible pairs of brands have been created. One-way ANOVA was applied to find out whether the means of each predictor variable differed between pen brands, to a 5 % significance level. Subsequently, appropriate post-hoc test was also conducted to determine the pairs of pen brands that differed significantly at particular variables.

Discrimination between nine different brands was also determined at univariants level via one-way ANOVA at 95% confidence interval rather than 99% confidence interval. With the utilization of the 95% confidence interval, there is a 5% probability of committing a type I error (false exclusion). When the confidence interval is increased to 99% that leaves only a 1% chance (or less) of committing a type I error, however at the same time the probability of committing a type II error (false inclusion) increases. With respect to forensic casework, type II errors (saying that two samples originated from the same source when they in actual fact do not) should be reduced or eliminated, if possible. In other words, although both types of errors pose problems for the forensic examiner, arguably from a forensic point of view data that presents a higher percentage of type II errors is preferred over the contrary [18]. Figure 3 shows the summary of results for 36 pen pairs. Out of 36 pairs, only two pairs were statistically indistinguishable meaning that they share very similar (statistically the same) organic profile and were likely from the same source of origin. A discrimination power of 94.4% ($34/36 \times 100\% = 94.4\%$) was achieved.

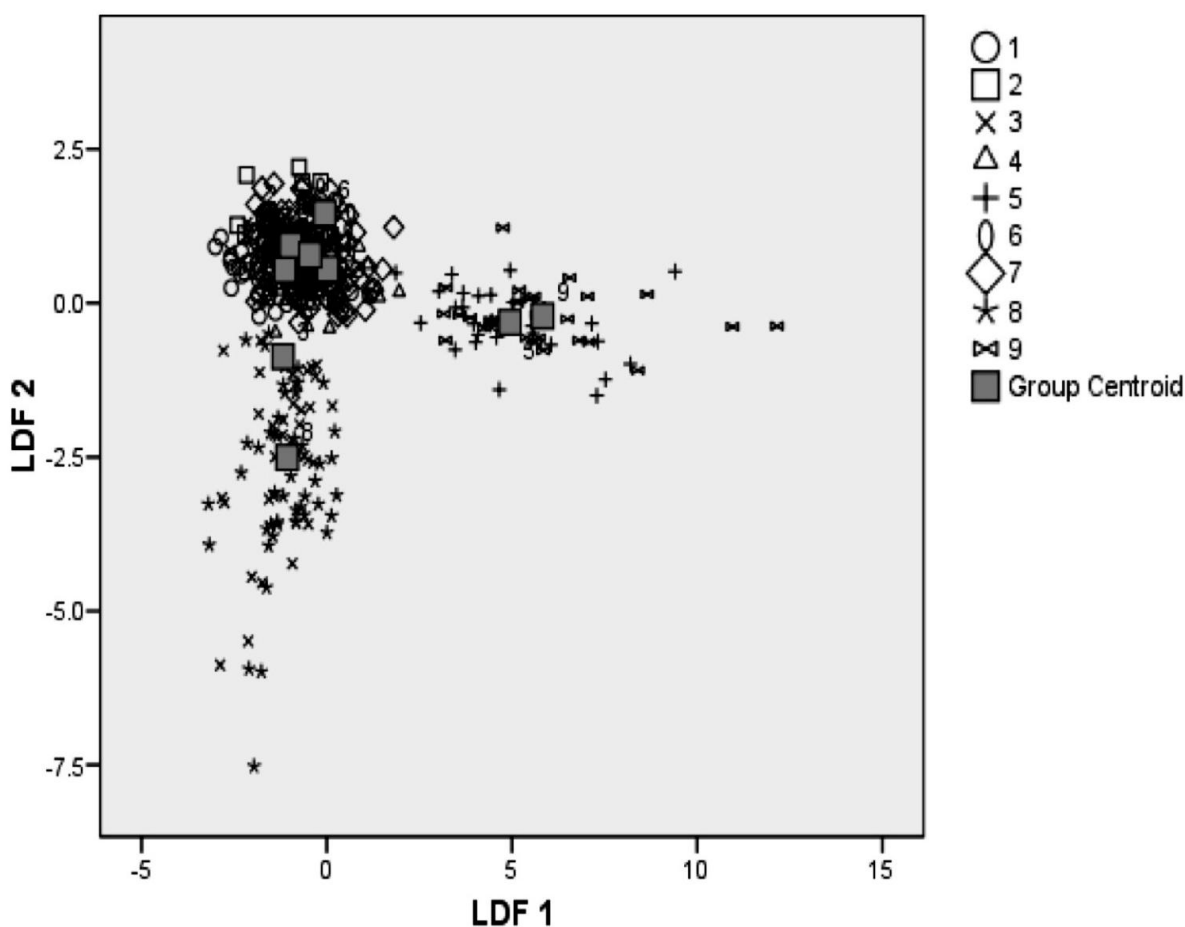


Figure 2. Scores plot of first two canonical discriminant functions showing the separation of group centroids for nine different brands of pen inks/ the projections of the 464 spectra into the space of the first two discriminant functions, displaying 76.19% of the between-to-within group variation in the data.

Identification of pen brands

Having demonstrated that better separation could be achieved through pairwise group comparison, a flow chart was designed for the IR data based on the absence or presence of significant statistical difference of certain principal components in comparing with a selected reference spectrum to identify the pen brand of studied inks objectively. One-way ANOVA showed that 23 out of 36 principal components differed significantly at the 0.05 level. After examining the output of post-hoc tests, it was decided that the IR spectrum of pen brand 5 acts as a reference spectrum. The whole process of identifying a pen brand in comparing their IR spectra with that of the reference spectrum are shown in Figure 4. Firstly, questioned ink X is compared against the reference sample (the IR spectrum of pen brand 5) at PCA 1_1. If the score value of both at PCA 1_1 is significantly different at 95% confidence interval, then sample X could be of brand 1, 2, 3, 4, 6, 7 or 8. On the other hand, insignificant result indicated sample X could likely be brand 9 or brand 5. The identification process may continue until the brand of sample X is eventually identified.

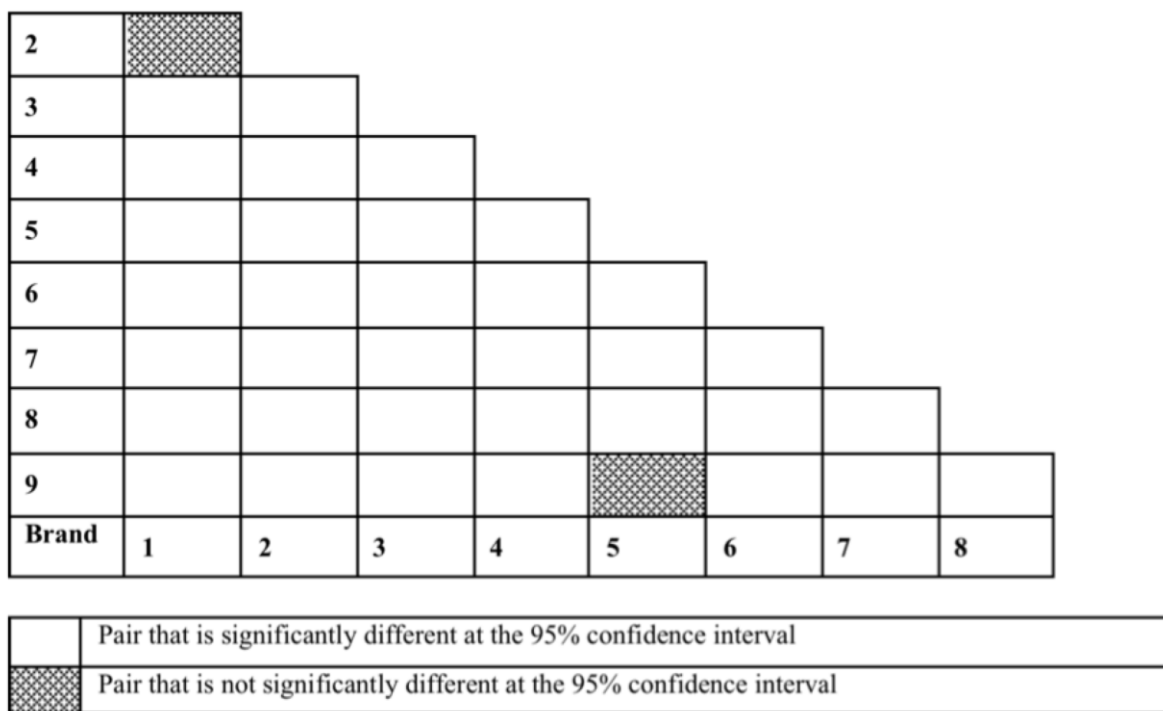


Figure 3. Univariate pair-wise group comparison of nine different brands of black ballpoint pen inks

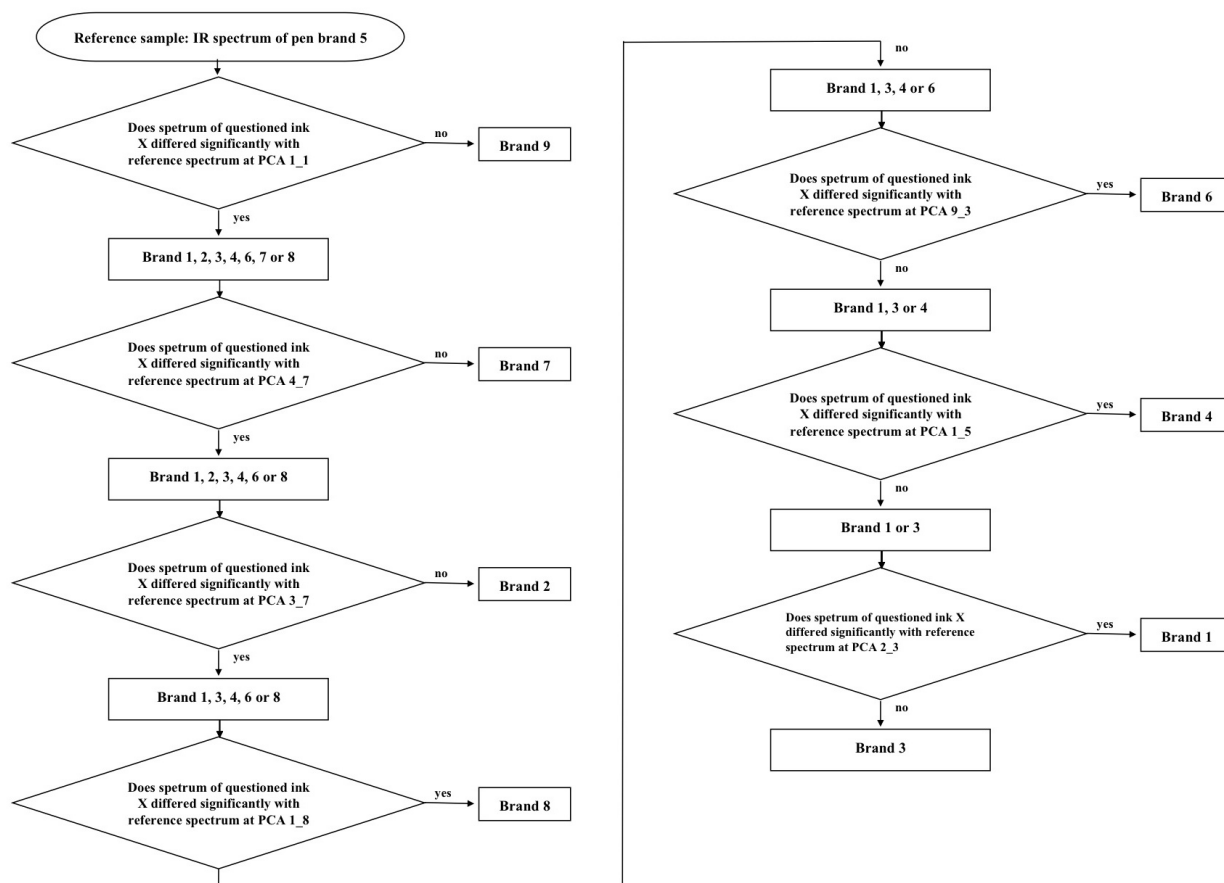


Figure 4. Flow chart for identification of nine pen brands based on their IR spectral

Conclusion

In conclusion, the results of the present study have demonstrated the ability of micro-ATR-FTIR spectroscopy to produced an informative spectrum for classification, differentiation and identification of inks via multivariate analysis techniques. The chief advantages of micro-ATR FTIR spectroscopy are that it is non-destructive, does not require any prior preparation of the sample, requires only a small amount of sample for examination, not time-consuming due to its rapid analysis time, eco-friendly and cost-effective as it does not involve any reagent. On the other hand, multivariate analysis is of great utility in handling high dimensional spectra data and in assessing quantitatively the minor differences of a particular ink relative to other inks that otherwise could not be noticed easily by visual inspection. Since encouraging results were obtained, the same method could be applied to further compare ink from pens of the same brand. If satisfactory results are obtained, an automatic algorithm could be developed.

Acknowledgement

We are grateful for the assistance in the acquisition of infrared data from Mr.Fizol and Mrs. Rusikah bt. Minhad from Jabatan Kimia Malaysia. We wish to thank Prof. Dr. Aziz Jermain for useful discussion on multivariate analysis. We would also like to extend our gratitude to all the staff of the forensic science program from UKM as well as Mrs. Hartinah Annuar. Also thanks to UKM for research grant UKM-DLP-2012-024.

References

1. C. Neumann, R. Ramotowski, T. Genessay, (2011). Forensic examination of ink by high-performance thin layer chromatography--the United States Secret Service Digital Ink Library, *J. Chromatogr. A*. 1218: 2793-811.
2. A. Kher, M. Mulholland, E. Green, B. Reedy, (2006). Forensic classification of ballpoint pen inks using high performance liquid chromatography and infrared spectroscopy with principal components analysis and linear discriminant analysis, *Vib. Spectrosc.* 40: 270-77.
3. J.A. Denman, W.M. Skinner, K.P. Kirkbride, I.M. Kempson, (2010). Organic and inorganic discrimination of ballpoint pen inks by TOF-SIMS and multivariate statistics, *Appl. Surf. Sci.* 256: 2155-63.
4. M. Gallidabino, C. Weyermann, R. Marquis, (2011). Differentiation of blue ballpoint pen inks by positive and negative mode LDI-MS, *Forensic Sci. Int.* 204: 169-177.
5. T.J.U. Thompson, M. Gauthier, M. Islam, (2009). The application of a new method of Fourier Transform Infrared Spectroscopy to the analysis if burned bone, *J. of Archaeol. Sci.* 36: 910-914.
6. N.C. Thanasoulas, E.T. Piliouris, M-S.E. Kotti, N.P. Evmiridis, (2002). Application of multivariate chemometrics in forensic soil discrimination based on the UV-Vis spectrum of the acid fraction of humus, *Forensic Sci. Int.* 130: 73-82.
7. A. Kher, M. Mulholland, B. Reedy, P. Maynard, (2001). Classification of Document Papers by Infrared Spectroscopy and Multivariate Statistical Techniques, *Appl. Spectrosc.* 55: 1192-1198.
8. R.H. Brody, H.G.M. Edwards, A.M. Pollard. (2001). Chemometric method applied to the differentiation of Fourier-transform Raman spectra of ivories, *Anal. Chim. Acta* 427: 223-32.
9. N.C. Thanasoulas, N.A. Parisi, N.P. Evmiridis, (2003). Multivariate chemometrics for the forensic discrimination of blue ball-point pen inks based on their Vis spectra, *Forensic Sci. Int.* 138: 75- 84.
10. Y. Mallet, D. Coomans, O. de Vel, (1996). Recent developments in discriminant analysis on high dimensional spectral data, *Chemom. Intell. Lab. Syst.* 35: 157-173.
11. W. Dirwono, J.S. Park, M.R. Agustin-Camacho, J. Kim , H-M. Park, Y. Lee, K-B. Lee, (2010). Application of micro-attenuated total reflectance FTIR spectroscopy in the forensic study of questioned documents involving red seal inks, *Forensic Sci. Int.* 199: 6-8.
12. J. Wang, G. Lou, S. Sun, Z. Wang, Y. Wang, (2001). Systematic analysis of bulk blue ballpoint pen ink by FTIR spectrometry, *J. Forensic Sci.* 46: 1093-1097
13. V. Causin, C. Marega, A. Marigo, R. Casamassima, G. Peluso, L. Ripani, (2010). Forensic differentiation of paper by X-ray diffraction and infrared spectroscopy, *Forensic Sci. Int.* 197: 70-74.
14. Y. Mallet, D. Coomans, O. de Vel, (1996). Recent developments in discriminant analysis on high dimensional spectral data, *Chemom. Intell. Lab. Syst.* 35: 157-173
15. H.F. Kaiser, (1960). The application of electronic computers to factor analysis, *Edu. Psychol. Meas.* 20: 141-151.
16. Y.H. Chan, (2005). Biostatistics 303: Discriminant analysis, *Singapore Med. J.* 46: 54-61.
17. M.A. Hall, (1999) Correlation based Feature Selection for Machine Learning [PhD thesis]. Hamilton: University of Waikato.
18. B. E. Naes, (2009). Elemental analysis of glass and ink by laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) and laser induced breakdown spectroscopy (LIBS), [PhD thesis], Miami: Florida International University.