Malaysian Journal of Forensic Sciences (2015) 6(1):48-53

Fourier Transform Infrared (FTIR) Spectroscopy and Principal Component Analysis (PCA) of Unbranded Black Ballpoint Pen Inks

Mustafa Kamil[a], Muhammad Naeim Mohamad Asri[b], Wan Nur Syuhaila Mat Desa[b], Dzulkiflee Ismail[b*]

[a]*Faculty of Science and Technology, Universiti Islam Sains Malaysia, 71800 Bandar Baru Nilai, Negeri Sembilan.*
[b]*Forensic Science Programme, School of Health Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian.*

ABSTRACT: Within forensic examinations of questioned documents, characterisation of ballpoint pen ink plays an extremely important role. In this work, inks from eleven (11) unbranded black ballpoint pens collected at an educational fair were subjected to non-destructive, *in-situ* analysis of Fourier Transform Infrared (FTIR) Spectroscopy. The spectroscopic data within the region of 1800-1200 cm^{-1} were analysed using the chemometrics techniques of principal component analysis (PCA) and hierarchical cluster analysis (HCA). The score plot of the PCA reveals seven clusters (five homogenous and two heterogeneous clusters) while the dendrogram of the HCA reveals six clusters (two homogenous and four heterogeneous clusters) which indicate that the unbranded black ballpoint pen inks may share similar ink formulations which can complicate their identifications.

Keywords: FTIR, PCA, unbranded ballpoint pen, ink analysis, questioned document examination

Introduction

Questioned documents (QD) are defined as documents suspected of being fraudulent or whose source, history and authenticity are dubious. Questioned documents are normally associated with criminal activities such as fraud, forgery, blackmailing, illicit drugs trafficking and homicide therefore proper and systematic analyses of such documents may help in solving crimes related to the documents [1].

Ink analysis is one of the important aspects in forensic questioned document examinations to determine the authenticity or validity of a document therefore it is extremely imperative to have accurate and discriminating ink analysis protocol. Ballpoint pens are perhaps the most popular writing implements and are widely used hence it is not surprising to see their association with criminal cases. Ballpoint pen is made up of plastic casing which houses the ink reservoir and a ball which is continuously coated with the ink. In the act of writing the ball rotates and eventually transfers the ink onto the writing substrate. Although the detailed composition of ink is not known, in general it consisted of dyes, pigments, lubricants, surfactants and resins.

Any examination and analysis performed on a questioned document must take into consideration the limited amount of ink available on the document. For this reason, and for reasons of preserving the integrity of the questioned document, an overwhelming consideration in questioned document analysis has been to prioritise non-destructive examinations over destructive ones [2, 3]. One of the non-destructive instrumental techniques commonly employed for forensic ink analysis is the Fourier Transform Infrared (FTIR) spectroscopy. Recent development of the Attenuated Total Reflectance (ATR-FTIR) has revolutionised forensic ink analysis by providing a simple, rapid, non-destructive and cost effective analytical approach [4-7].

On determining the provenance of inks, the difficulty lies in the fact that most ink examination processes remain manual and subjective and suffer from the drawbacks of human interpretation. Therefore, advances in computation and availability of modern statistical software incorporating multivariate chemometrics capabilities make it possible for data to be interrogated with relative ease in order to objectively identify, patterns or trends that have potential to increase the discriminating power of an analytical technique.

In forensic science, chemometrics techniques of Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) are the most commonly use technique [8]. These techniques are described as unsupervised methods of pattern recognition i.e. they are able to objectively identify pattern within large and complex datasets and presenting them graphically for ease of classification.

Principal Component Analysis (PCA)

PCA reduces the dimensionality of the original data set by calculating a new set of variables known as Principal Components. These Principal Components (PC's) are calculated from linear combinations of the original variables and are organised according to the amount of variance in the original data set that they explain which are uncorrelated to one another. Typically, only the first two PC's are required to explain the majority of the variance (80 – 90%) in the original dataset, and are used to generate a graphical output called a score plot. In the score plot, the samples are arranged in space relative to each other according to their first two PC scores. The samples with similar scores will occupy similar position; whilst those with dissimilar scores will be positioned some distance away, thus allowing clusters to be identified [9-12]. When performing PCA clustering, prior knowledge of the provenance of the samples or objects studied is required to avoid misclassification of the samples or objects.

Hierarchical Cluster Analysis (HCA)

HCA orders a large complex data set into clusters which are graphically displayed in a tree diagram called a dendrogram which shows the level of similarity (via a similarity index (%)) between individual sample and groups of samples relative to the entire dataset. The dendrogram is produced typically using agglomerative methods where clustering starts with individual sample and proceeds sequentially until all samples are linked together to form clusters. Several agglomerative methods of linking clusters and distance measures are available but the most common is the single linkage (where the distance between two clusters is the minimum distance between a variable in one cluster and a variable in the other cluster) with the Euclidean distance measure. Different linkage and distance measure may result in different outcomes and the chosen method for a particular dataset is usually based on a trial and error process. The ideal outcome is a dendrogram which exhibits clusters with a relatively large similarity index and comparatively small distances between connected clusters [12].

The aims of this present study are to analyse unbranded black ballpoint pen inks using ATR-FTIR technique and to evaluate whether or not discrimination of such inks can be achieved using chemometrical techniques of Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA).

Experimental

Sample Collection

The unbranded black ballpoint pens used in this study were collected during an educational fair from eleven (11) institutions participating in the event. The pens are listed in Table 1. For ease of identification, each ballpoint pen was given specific reference code. All pens were inscribed with institution name on their plastic casing except for the pen collected from Kementerian Pengajian Tinggi (KPT) (reference code K) which bore no institution name. The plastic casing of the pens were of different shapes, sizes and colours.

Table 1: List of black ballpoint pens collected during the education fair and their references codes

Pen Name	Reference Code
Cyberjaya University	A
University Malaysia Terengganu (UMT)	B
University Science of Islamic Malaysia (USIM)	C
Institute Pendidikan Guru Malaysia (IPGM)	D
Institute Kemajuan Ikhtisas Pahang (IKIP)	E
University Malaya (UM)	F
University Science of Malaysia (USM)	G
University Pertahanan Malaysia (UPNM)	H
University Tun Hussien Onn Malaysia (UTHM)	I
Academic Affairs of University Technology Mara (UiTM)	J
Kementerian Pengajian Tinggi (KPT)	K

Sample Preparation

Ink spot (of each pen) was drawn separately onto a small square of 0.5 x 0.5 cm on white 80 gram A4 paper (Tesco, Malaysia). The ink spots were allowed to stand at room temperature for approximately five (5) minutes prior to spectra acquisition. The spectra were acquired using Bruker FTIR spectrometer (USA) equipped with a diamond ATR sampling interface. The diamond ATR sampling interface was wiped clean with ethanol soaked tissue before and after spectrum acquisition. The reference spectrum of air was also obtained prior to sample measurements to ensure that the ATR-FTIR was working correctly. Spectral measurements were taken over a spectral range of 4000 cm^{-1} to 650 cm^{-1} with resolution of 4 cm^{-1}. Triplicate analysis was performed for each sample.

Chemometrics Analysis

The PCA and HCA were performed using Minitab Version 16.2.3 (Minitab Incorporated, State College, PA, USA). Prior to the chemometrics analyses, the IR spectral data were imported and stored in a data matrix containing *m* rows and *p* column (*m* represented the spectral intensities while *p* represented different wavenumbers) in Microsoft Excel (Microsoft Incorporated, USA) spreadsheet for data pre-processing.

Results and Discussion

Assessing the ink spectra of the eleven ballpoint pens showed that most of the variability occurred predominantly in the region of 1800 – 1200 cm^{-1}. In the region of 4000 – 2000 cm^{-1}, the spectral variations are attributed to the atmospheric water vapour (H_2O) and carbon dioxide (CO_2) as described by Kher *et. al* [13]. Since this region contains only few bands, it was excluded from data analysis so for chemometrics analysis, only region 1800 -1200 cm^{-1} was considered. The ink spectrum of one of the ballpoint pens (reference code B) is shown in Fig. 1 with window showing the region considered for chemometrics analysis.

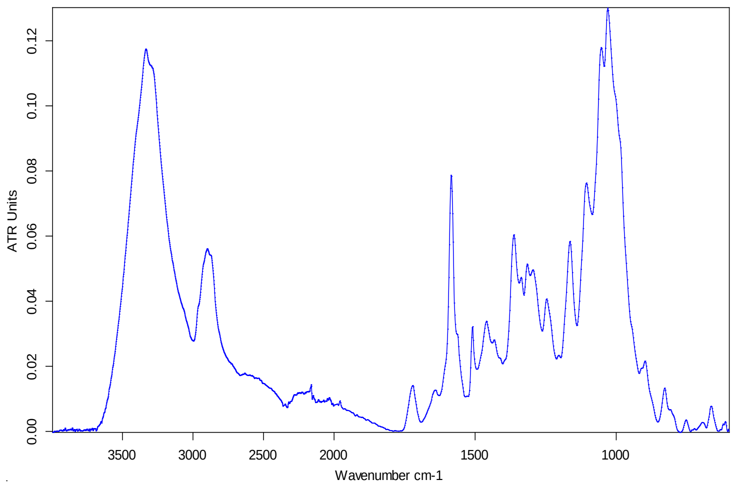

Fig. 1: Ink spectrum collected from ballpoint pen ink with reference code B. The window represents the region considered for chemometrics analysis

PCA

PCA performed to the IR spectral data in the region of 1800 – 1200 cm^{-1} produced score plot as shown in Fig. 2. Seven clusters (designated as cluster A, B, C, D, E, F and G) are evident in the score plot when the first two principal components *i.e.* PC1 and PC2, which accounted for 85.6% of the total variance in the data set were considered. Except for cluster B and D, the rest of the clusters comprise of inks from the same ballpoint pen. Cluster B comprises of inks from ballpoint pens G and I while cluster D comprises of inks from ballpoint pens K, B, E and H.

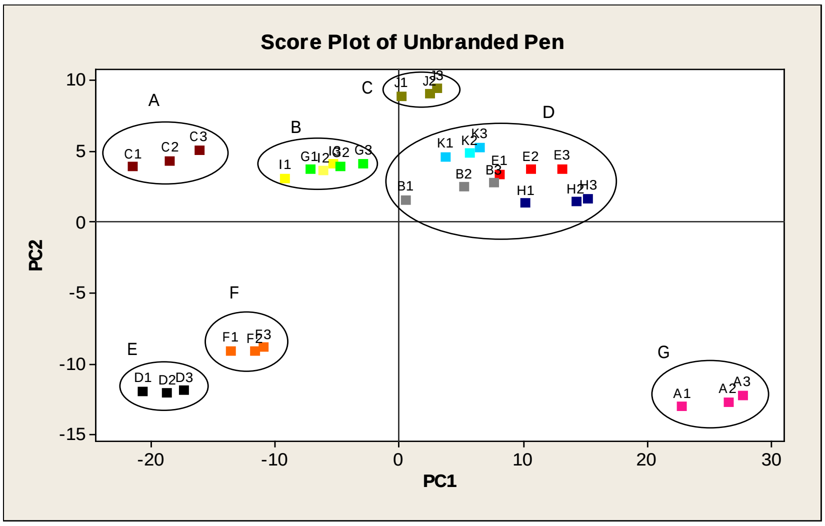

Fig. 2 Score plot of the unbranded ballpoint pen inks

HCA

HCA was performed using the Euclidean distance as the distance measure and single linkage strategy to link clusters within the data set. Figure 3 shows the dendrogram produced from the data set. At similarity index of approximately 86%, six (6) clusters (designated as cluster A, B, C, D, E and F) are evident in the dendrogram. Except for cluster A and E which comprise of inks from the same pens, the rest of the clusters comprise of inks from different pens. Cluster B comprises of inks from Pen B, E and H, cluster C comprises of inks from Pen J and Pen K, cluster D comprises of Pen G and I while the final cluster i.e. cluster F comprises of inks from Pen D and Pen F.

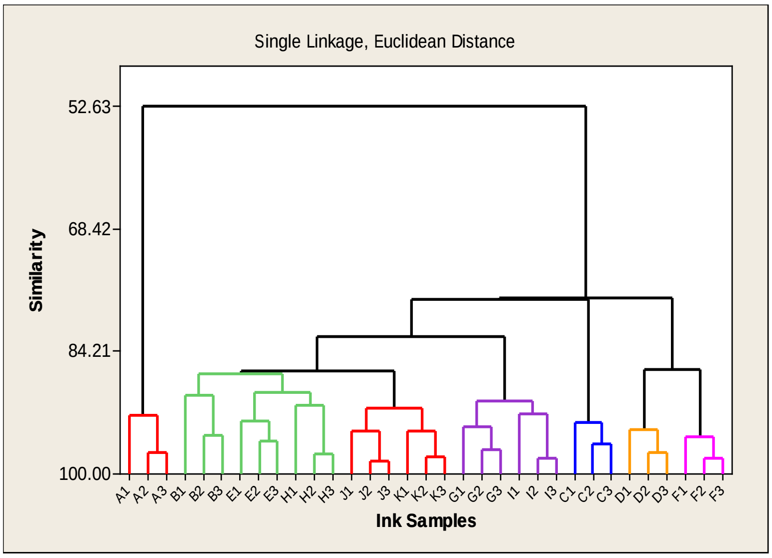

Fig. 3: Dendrogram of the unbranded ballpoint pen inks

Both PCA and HCA have failed to establish homogenous or neat clusters corresponding to the number of pens considered in this study. The inability of the chemometrics techniques to establish homogenous or neat cluster for some of the unbranded black ballpoint pens is probably due to the fact that the pens used inks with the same chemical formulations. This is not surprising since such pens may be manufactured by the same manufacturer (although with different plastic casing of various shapes, sizes and colours) that uses ink with the same chemical formulations or by different manufacturer (hence the different in term of shapes, sizes and colour of the plastic casing) that uses ink with the same chemical formulations. From forensic questioned document perspective, this situation may complicate the identification of the ink deposited on a document and eventually the pen used to write the document.

Conclusion

Careful attention should be given when examining and analysing inks of this nature since similarities in term of IR spectra do not necessarily mean that the inks originate from the same pen and vice-versa. Although non-destructive technique such as ATR-FTIR preserves the integrity of document, situation like this warranted combination of examination techniques including destructive techniques for examples high performance liquid chromatography (HPLC) and gas chromatography (GC) to positively confirm on the identity of the inks.

Acknowledgements

The authors are grateful for the financial support from the Universiti Sains Malaysia (USM) under the Short Term Grant Scheme (304/PPSK/61313030) which enabled this work to be conducted.

References

1. Claybourn, M. and Ansell, M., Using Raman Spectroscopy to Solve Crime: Inks, Questioned Documents and Fraud. Science and Justice, 2000. 40: p. 261-271.
2. Ellen, D., Chapter 7: The Materials of Handwritten Documents: Substances and Techniques, in Scientific Examination of Documents: Methods and Techniques. 2006, Taylor and Francis. p. 115-148.
3. Brunelle, R.L. and Crawford, K.R., Chapter 5: Forensic Comparison and Identification of Writing Inks by Thin Layer Chromatography and Densitometry, in Advances in the Forensic Analysis and Dating of Writing Ink. 2003, Charles C. Thomas. p. 72-107.
4. Causin, V., Casamassima, R., Marega, C., Maida, P., Sciavone, S., Marigo, A. and Villari, A., The Discrimination Potential of Ultra Violet - Visible Spectrophotometry, Thin-Layer Chromatography, and Fourier Transform Infrared Spectroscopy for the Forensic Analysis of Black and Blue Ballpoint Inks. Journal of Forensic Sciences, 2008. 53(6): p. 1468-1473.
5. Wang, J., Luo, G., Sun, S., Wang, Z. and Wang, Y., Systematic Analysis of Bulk Blue Ballpoint Pen Ink by FTIR Spectrometry. Journal of Forensic Sciences, 2001. 46(5): p. 1093-1097.
6. Zieba-Palus, J. and Kunicki, M., Application of the Micro-FTIR Spectroscopy, Raman Spectroscopy and XRF Method Examination of Inks. Forensic Science International, 2006. 158: p. 164-172.
7. Dirwono, W., Park, J.S., Agustin-Camacho, M.R., Kim, J., Park, H., Lee, Y., and Lee, K., Application of Micro-Attenuated Total Reflectance FTIR Spectroscopy in the Forensic Study of Questioned Documents Involving Red Seal Inks. Forensic Science International, 2010. 199: p. 6-8.
8. Morgan, S.L. and Bartick, E.G., Chapter 13: Discrimination of Forensic Analytical Chemical Data Using Multivariate Statistics, in Forensic Analysis on the Cutting Edge: New Methods for Trace Evidence Analysis, R.D. Blackledge, 2007, John Wiley and Sons Inc. p. 333-374
9. Mat-Desa, W.N., Nic Daeid, N., Dzulkiflee, I. and Savage, K., Application of Unsupervised Chemometric Analysis and Self-Organising Feature Map (SOFM) for the Classification of Lighter Fuels. Analytical Chemistry, 2010. 82: p. 6395-6400.
10. Adam, C.D., In Situ Luminescence Spectroscopy with Multivariate Analysis for the Discrimination of Black Ballpoint Pen Ink Lines on Paper. Forensic Science International, 2008. 182: p. 27-34.
11. Thanasoulias, N.C., Parisis, N.A. and Evmiridis, N.P., Multivariate Chemometrics for the Forensic Discimination of Blue Ballpoint Pen Inks Based on their UV-Vis Spectra. Forensic Science International, 2003. 138: p. 75-84.
12. Everitt, B.S., and Dunn, G., Applied Multivariate Data Analysis. 2001, London: Arnold.
13. Kher, A., Mulholland, M., Green, E. and Reedy, B., Forensic Classification of Ballpoint Pen Inks Using High Performance Liquid Chromatography and Infrared Spectroscopy with Principal Component Analysis and Linear Discriminant Analysis. Vibrational Spectroscopy, 2006. 40: p. 270-277.

Additional information and reprint request:
Dzulkiflee Ismail
Email: dzulkiflee@usm.my
Forensic Science Programme
School of Health Sciences
Universiti Sains Malaysia
16150 Kubang Kerian
Kelantan, Malaysia
Tel: +609-7677820