UNIVERSITI SAINS MALAYSIA

SIMULATED SOURCE DETERMINATION OF BLUE BALLPOINT PEN INKS USING FTIR SPECTROSCOPY, PEARSON PRODUCT MOMENT CORRELATION AND PRINCIPLE COMPONENT ANALYSIS

Dissertation submitted in partial fulfillment for the

Degree of Bachelor of Science (Hons.) in Forensic Science

NUR DIYANA BT ABD RAZAK

School Of Health Sciences

Universiti Sains Malaysia

Health Campus

16150, Kubang Kerian

Kelantan, Malaysia

2015

ACKNOWLEDGEMENTS

In the Name of Allah the Most Gracious and Merciful

Alhamdulillah, all praises to Allah for the blessing and strength Firstly, I would like to express my gratitude and love to Allah for giving me the strength and guidance whilst completing this piece of work. I would like to express my deepest gratitude and thanks to my supervisor, Dr. Dzulkiflee Ismail, Lecturer in Forensic Science, School of Health Sciences, Universiti Sains Malaysia for the support, encouragement, suggestion and contribution of time throughout this research process and completion of this dissertation.

Next, I would like to express my gratitude to the all staffs at Questioned Document Department of Jabatan Kimia Malaysia (JKM) for their full cooperation and guidance given to me throughout my visit to their laboratory. I would also like to express my thanks and love to my parents and siblings for their full support.

Last but not least, I would like to thanks Muhammad Naeim bin Mohamad Asri and also all my classmates for their courage and never ending spirit during the writing period of this thesis. Their love and care always strengthen and inspire me. Thank you all.

NUR DIYANA BT ABD RAZAK 112121

Table of Contents

LIST OF ABBREVIATION

ATR	: Attenuated Total Reflectance
CE	: Capillary Electrophoresis
CM	: Centimeter
DRIFTS	: Diffuse Reflectance Infrared Fourier Transform Spectroscopy
FTIR	: Fourier Transform Infrared Spectroscopy
GC	: Gas Chromatography
HPLC	: High Performanceced Liquid Chromatography
JKM	: Jabatan Kimia Malaysia
LDI-MS	: Laser Desorption Ionization Mass Spectrometry
PCA	: Principle Component Analysis
PDRM	: Polis Diraja Malaysia
PPMC	: Pearson Product Moment Correlation
SEM	: Scanning electron microscopy
TLC	: Thin Layer Chromatography
VSC	: Video Spectral Comparative

LIST OF FIGURE

LIST OF TABLE

ABSTRACT

Ballpoint pens are the most common writing implement used to write a document therefore it is not surprising to see their association with criminal cases such as insurance fraud, homicide, drug dealing and kidnapping. The ink deposited by a ballpoint pen is the course of writing a dubious document can reveal useful information about the ballpoint pen. This study uses pair-wise manual comparison, statistical technique of Pearson Product Moment Correlation (PPMC) or simply Pearson's r and chemometric technique of principle component analysis to determine the ballpoint pen responsible for writing a simulated threatening note. Eleven (11) blue ballpoint pen inks in which one of them was used to write a threatening note were subjected to Fourier Transform Infra Red (FTIR) spectroscopy. The FTIR spectra of the blue ballpoint pen inks were first interrogate using pair- wise i.e. one to one manual comparison followed by PPMC and PCA. Of three techniques employed, only PCA through its principle component score plot had successfully determined the blue ballpoint pen inks responsible for writing the simulated threatening note . This study revealed that combination of FTIR and PCA can be used to determine the original source of ink provided that the reference ink is available.

ABSTRAK

Pen mata bulat adalah tulisan yang paling biasa digunakan untuk menulis dokumen , oleh itu tidak hairanlah untuk melihat penghubungan mereka dengan kes-kes jenayah seperti penipuan insurans, pembunuhan, urusan dadah dan penculikan. Dakwat yang dikeluarkan oleh pen mata bulat semasa menulis dokumen yang meragukan boleh mendedahkan maklumat yang berguna mengenai pen mata bulat itu. Kajian ini menggunakan "pair-wise" manual perbandingan, teknik statistik Produk Korelasi Momen Pearson (PPMC) atau hanya r Pearson dan teknik chemometric analisis komponen prinsipal untuk menentukan pen mata bulat yang bertanggungjawab untuk menulis nota yang mengancam simulasi. Sebelas (11) pen mata bulat berdakwat biru di mana salah satu daripada nya telah digunakan untuk menulis nota yang mengancam tertakluk kepada Fourier Transform Infra Red (FTIR) spektroskopi. Pertamanya, spektrum FTIR pen mata bulat berdakwat biru dianalisis menggunakan "pair-wise" iaitu satu per satu perbandingan manual diikuti oleh PPMC dan PCA. Tiga teknik yang digunakan, hanya PCA melalui komponen prinsip skor plot telah berjaya menentukan pen mata bulat biru berdakat biru yang bertanggungjawab untuk menulis nota yang mengancam simulasi. Kajian ini mendedahkan bahawa gabungan FTIR dan PCA boleh digunakan untuk menentukan sumber asal dakwat dengan syarat bahawa dakwat rujukan disediakan.

CHAPTER 1 : INTRODUCTION

1.1 Document in Forensic Science

A document can be defined as anything upon which a mark is made for the purpose of conveying a message. Documents may be involved not only in forgery cases but also in robberies, suicides and burglaries. Examples, in suicides case sometimes the victim left a message or notes for their family. This evidence can be used for the police. A questioned document is any signature, handwriting, typewriting whose source or authenticity is in dispute or dubious (USLEGAL, 2001).

In forensic, a document is any medium that contains printed or handwritten markings whose source or authenticity is in doubt. A number of these examinations involve the ink that the document was written with. The identity of the writing implement can be known using several technique (Siegel *et al.*, 2005). In Malaysia, Department Of Chemistry and also Polis Diraja Malaysia (PDRM) are the institution that carried out question document analyses (USLEGAL, 2001)

Almost century forensic science has addressed the difficult problem of the scientific examination of documents. A lot of the techniques developed over that period are still in use today. The technology used to produce documents continues to grow and with the propagation of the computer and image processing tools the methods used to produce forgeries are more easier (Barbosa *et al.*, 2014).

.

Government and also private sector were used ink analysis method to ensure the authencity or fake of question document. The examination of inks is often made to distinguish between inks in order to evaluate the authenticity of a document. This can make the detection and confirmation of alteration to documents with significant financial value like checks, insurance claims, wills, contract and tax return(Halim *et al.*, 2012a).

1.2 Forensic Ink Analysis

In forensic, Ink analysis is an important tools that can expose useful information about questioned document(Senior *et al.*, 2012). Forensic scientist can evaluate authencity of unknown document based on the different types of ink. It is possible to identify the brand of the pen even though it is difficult to determine whether an individual pen was used to write a document. This need some kind of classification and discrimination among the different brands of inks (Feraru and Meghea, 2014).

Ballpoint pens are usually famous in the most writing and also widely used in fraud and criminal cases. Ink analysis is the perfect technique which that can be use to know the authencity of the suspicious document (Feraru and Meghea, 2014).

1.3 History of Ball Point Pen

Based on the chronology, it is difficult to determine the history of the ballpoint pen. Many inventors try to improvise ballpoint pen to become better in several decades. In 1879 in Providence, Rhode Island, Alonzo T. Cross invented the stylographic fountain pen, a precursor of the ball-point pen. He entered competition with Duncan Mackinnon, the other stylographic pen inventor. In 1880 A. T. Cross divided his business from his father's and renamed his company the A.T. Cross - Pen and Pencil Manufacturer.(Cosmopolis, 2002).

In 1884 L.E. Waterman, a New York City insurance salesman, designed the first workable fountain pen, the fountain pen becomes the major writing instrument for the next sixty years. There are four fountain pen manufactures take over the market which are Parker, Sheaffer, Waterman and Wahl-Eversharp (Cosmopolis, 2002)

In 1938, A Hungarian journalist named Laszlo Biro invented the first ballpoint pen. Biro had noticed that in the newspaper printing the type of ink used dried quickly, leaving the paper dry and smudge-free. He wanted to create a pen using the same type of ink. The thicker ink would not flow from a regular pen and Biro had to devise a new type of point. He did so by fitting his pen with a tiny ball bearing in its tip. As the pen moved along the paper, the ball rotated picking up ink from the ink cartridge and leaving it on the paper.

Then the ballpoint pen made a foray into the British market. Laszlo Biro, the inventor continued his display of entrepreneurial flair and managed to safe significant financial scores to his name, however, he soon lost the Biro Company to Henry Martin who took over the manufacture of ball pen from Laszlo Biro (Bellis, 2015).

Martin made the ballpoint pen a recognized name (Biro Swan) in the market, but the credentials to bring ball pen to the position that it enjoys today, go to Marcel Bich. He founded a company that manufactured his own patented ball pen designs. BIC Cristal, that directly succeeded Biro, enjoys the largest share in the market today (Verma, 2012).

1.4 Mechanism of Ballpoint Pen

A ballpoint pen can be described as a writing instrument which features a tip that is automatically refilled with ink. It consists of an exactly formed metal ball seated in a socket below a reservoir of ink. Ink is delivered when the pen is moving along a writing surface example paper (How Product Are Made, 2015).

As the operator moves the pen across the surface the rotation of the ball transfer ink from the sealed chamber to the writing surface. The ball has two key functions which are to seal the ink chamber and to transfer the ink to the writing surface. The ink is fed by gravity to the upper surface of the ball (University of Reading, 2014). Figure 1.1 shows the mechanism of ballpoint pen.

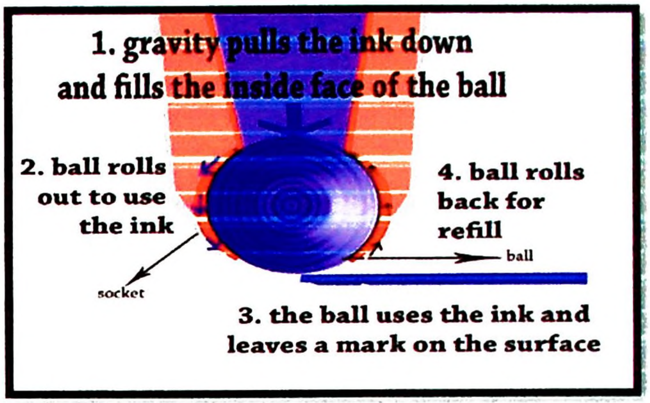

Figure 1.1: The Mechanism of Ballpoint Pen (Harlem, 2011)

1.5 Problem Statements

The problem statement in this study is:

I. Is the eleven blue ballpoint pen inks can be characterize using Fourier Transform Infrared Attenuated Total Reflectance (FTIR-ATR)

II. Is the FTIR spectra profile of the blue ball point pen inks can be analyses using Pearson Moment Correlation and chemometrics technique of Principle Component Analysis

1.6 Significant of the Study

In this part, the significant of the study were discussed. The findings of this study are expected to helps forensic question document examiner to successfully determine which pen had been used to write a document. Other than that, this method can be used to help characterize ballpoint pen inks to support more objective conclusion rather than subjective conclusion.

1.7 Expected Outcomes

The expected outcome in this study are the eleven blue ballpoint pen inks will produce different FTIR spectra profiles .Next, the Pearson Moment Correlation and Principal Component Analysis (PCA) will be able to “link” the ink on the simulated threatening note to one of the blue ballpoint pen inks.

1.8 Objectives of the Study

This section discuss the three main objective in the study. Firstly, to characterize eleven blue ballpoint pen inks using Fourier Transform Infrared Attenuated Total Reflectance (FTIR-ATR).Secondly, to analyze the FTIR spectra of the blue ball point pen inks using Pearson Moment Correlation and chemometrics technique of Principle Component Analysis and finally to assess the feasibilities of statistical and chemometrics techniques for source determination of blue ballpoint pen inks.

CHAPTER 2 : LITERATURE REVIEW

2.1 Introduction

Ink analysis is an important forensic procedure that can tell useful information about questioned documents. Most of its applications regard the detection and confirmation of alterations to documents with important financial value such as wills, contracts, insurans claim and tax returns (Thanasoulias *et al.*, 2003). The authenticity of document can be known by performed the examination of ink to differentiate between ink (Halim *et al.*, 2012a).

Forensic ink analysis involved the examination of documents using the naked eye, oblique lighting conditions and using special optical filters. It can be performed using optical, spectroscopic and chromatographic methods (Ahmad and Yacob, 2005).

The study emphasizes on ballpoint pen inks available in Malaysian market as these pens are often used in the occasion of signing important documents . In this study, the ballpoint inks were analysed using infrared spectroscopy. The data obtained from these techniques may be further analysed using chemometric methods such as Principle Component Analysis (PCA) and discriminant (DA) because the employment of multivariate analysis allows the extraction of more information based on the similarities and differences among samples in a dataset (Halim *et al.*, 2012b).

Forensic experts often use optical methods and thin layer chromatography (TLC) for detecting fraud in such documents. The known methods make use of electronic microscopes and chemical reagents that may destruct the material tested. The objective of most analyses is to determine whether two pieces of written text originated from the same ink, therefore, comparison of different writing inks on a document is the main point of the most investigations (Djozan *et.al*, 2008).

According to Barbosa , he using a large number of samples of, a reasonably large number of ballpoint pens of different manufacturers. The ink from the different ballpoint manufacturers available in the Brazilian market has different color distributions and aging patterns. The documents were digitized using a flatbed scanner in true color 300 dpi resolution. The Mahalanobis distance was used to cluster the different samples forming clear different distributions for each ballpoint pen manufacturer. The method planned is a simple yet effective technique to analyze alterations performed at a later stage of the ballpoint pen writing of a document which used a pen from a different manufacturer(Barbosa *et al.*, 2014).

2.2 Destructive Techniques

The inks can be categorized by non-destructive or destructive methods depending on whether a sample needs to be taken from the document, a process that would alter it. In destructive methods, a portion of the ink has to be removed from the document prior to the analysis (Saferstein, 2001). Example of the destructive methods such as High Performance Liquid Chromatography (HPLC), Thin Layer Chromatography (TLC), Capillary Electrophoresis (CE) and Gas Chromatography/Mass spectroscopy (GC/MS) (Dirwono *et al.*, 2010).

Transmission FTIR Spectroscopy was chosen as a technique for inks analysis although it is destructive method due to cost effectiveness, good signal-to-noise ratio (absorb 100% energy) and it can produce the best quality of the spectrum. However, it is difficult to differentiate the pen inks and errors become occur when analyze the IR data manually . As alternative, the data obtained from this technique may be further analysed using chemometric methods(Halim *et al.*, 2012a).

Studied on multivariate chemometrics for the forensic discrimination of blue ballpoint pen inks based on their Vis spectra showed that the results of UV-Vis spectra of ink were difficult to compare. The ink sample size used was very small to overcome this problem. Each of the spectrums representsthe average of the absorption from the same batch. The chemometric application such as cluster analysis (CA),principal component analysis (PCA) and discriminant analysis(DA) was successively used to calculate the discriminant model. UV-VIS examination may provide indications that the document has been stained by chemicals or other materialsthat may affect the ink analysis (Nicholas *et al.*, 2003).

Observation of thin layer chromatograms (TLC) under alternative light sources, the use of infrared luminescence and microspectrophotometry have also been employed in an attempt to achieve better characterization of the TLC bands . Although TLC remains the preferred method for ink analysis due to its low cost and relative simplicity, high performance liquid chromatography (HPLC) has been used as an alternative destructive technique that has the advantage of higher resolution and is also capable of detecting colorless components in the ink matrix (Roux *et al.*, 1999).

2.3 Non-Destructive Techniques

Example of non-destructive methods are Scanning electron microscopy (SEM), Video Spectral Comparative (VSC) and Raman Spectroscopy . However, these techniques are complex and needs a long time consuming analytical steps . Besides that, it requiring very expensive instrument (Senvaitiene *et al.*, 2005)

Harris and MacDougall concluded that diffuse reflectance (DR) FTIR might be a viable method for non-destructive analysis of ink on paper. However, the spectra of ink on paper that Harris obtained did not compare with their reference spectra of the inks themselves (Harris and MacDougall, 1989).

Diffuse Reflectance Infrared Fourier Transform Spectroscopy (DRIFTS) was utilized for the analysis of blue ballpoint pen ink samples on paper using KBr as a background. This analysis was found to give a poor discrimination between the ink spectra. Micro-ATR was found to be a simpler technique for obtaining spectra of the ballpoint pen ink samples (Kher *et al.*, 2006)

One hundred and eight samples of blue ink, made in China, were collected and analyzed by FTIR spectroscopy. These spectra were associated with the chemical components of the inks and were separated into two groups. The spectra data of these inks, such as peak location and peak intensity, were described and summed up by an artificial intelligence of pattern recognition, and 35 subgroups of the two groups were obtained on the basis of their correlation coefficient (Wang *et al.*, 2001).

Denman et al. 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 (Denman *et al.*, 2010)

2.4 Components in Ballpoint Pen Ink

There are approximately thousand different ballpoint pen inks that are available in the market. Pen inks are viscous liquids made of a mixture of several dyes and pigments that provide colour, organic solvents that serve as a vehicle. Besides, it also contain resins that sustain the right viscosity and adhering properties, and other agents like antioxidants, preservatives and trace elements (Braz *et al.*, 2013). In general, ink of pen consists of three elements which are colouring materials, vehicle and resin

2.4.1 Colouring Materials

The most important component of the ballpoint pen ink is the colouring material which comes in the form of dyes, pigments or combination of both. Dyes are soluble in the liquid body of the ink, which is also known as the vehicle where as pigments are finely ground multi-molecular granules that are insoluble in the vehicle (Feraru and Meghea, 2014).

An ink's colour comes from either a pigment which is water insoluble or from a dye that is water-soluble. The dye eosin lends red ink its colour and is made by adding the element bromine to a fluorescent compound. Inks that use pigments include metallic gold ink which surprisingly uses a copper-zinc alloy and white ink which contains titanium oxide. Carbon black, a pigment derived from coal and oil, is an important part of black ballpoint pen ink (Johnson, 2001). For blue ink, several dyes are used but the more common ones include substitute triphenylmethane dyes, copper phthalocyanine blue and crystal violet.

2.4.1.1 The Vehicle

The vehicle is a faint bluish-black solution that is difficult to examine. The vehicle, whose composition affects the flowing and drying characteristics of the ink can consist of oils, solvents and resins (Feraru and Meghea, 2014). The ink vehicle can be either plant-based (linseed, rosin, or wood oils), which dries by penetration and oxidation, or solvent-based (such as kerosene), which dries through evaporation. The vehicle is a faint bluish-black solution that is difficult to examine (Ausley, 1998).

2.4.1.2 The Resin

Resins can either naturally occurring or synthetic polymeric material that begin in a highly viscous state and hardens with treatment. Typically, it is soluble in alcohol, but not in water (Wise GEEK, 2003). Natural resins are typically fusible and flammable organic substances that are transparent and are yellowish to brown in colour.

Most natural resins come from trees, especially pines and firs. Resin formation occurs as a result of injury to the bark from wind, fire, lightning, or other cause (The Editors of Encyclopædia Britannica, 2015). According to Johnson (2014) resins are also used for stabilizing the polymers to prevent clotting of the ballpoint pen ink.

2.5 History of FTIR Spectroscopy

FTIR stands for Fourier Transform Infrared . In the 1880s, Albert Abraham Michelson, inverted the interferometer to further his studies of the speed of light. The development of FTIR is imposible without the use of Michelson Interferometer (Smith, 1996)

In 1949 Peter Fellgett an astrophysicist used an interferometer to measure light from celestial bodies and produced the first Fourier transform infrared (FTIR) spectrum. However for many years, only a few advanced research groups with access to large, expensive computers and with personnel capable to wait up to 12 hours to transform an interferogram into a spectrum used Fourier transform infrared spectroscopy (Hsieh, 2008).

The first FTIR spectrometers were large and expensive, and were found primarily in a few well-to-do research labs. In the late 1960s when microcomputers able to do the Fourier transform became available, commercial FTIR spectrometers appeared. The 1966 development of the Cooley-Tukey algorithm, which quickly does a Fourier transform (the

Fast Fourier Transform or FFT), was also instrumental in the commercialization of FTIR spectrometers (Hsieh, 2008)

2.6 Theory of FTIR Spectroscopy

Infrared radiation is passed through a sample. Some of the infrared radiation is absorbed by the sample and some of it will passed through or transmitted . The resulting spectrum represents a fingerprint of a sample with absorption peaks which correspond to the frequencies of vibrations between the bonds of the atoms making up the material. No two compounds produce the same infrared spectrum that is principle of individual (Thermo Nicolet Corporation, 2001)

This spectroscopy can analysis and observed sample area as small as 20-200µm in diameter in both transmission or normal- reflection mode. Based on chemical information content, the wavelength range between 2.5 and 20 µm is of special interest. This region is referred as medium infrared (Buffeteau *et al.*, 1996).

According to quantum mechanics , the molecule can take up an amount of energy to reach the first vibrationally excited state. A molecule which is irradiated with a continuous spectrum of infrared energy may absorb light quanta which have this energy. The spectrum of the remaining radiation shows an absorption band at a frequency of complex vibrational and rotational movements of the molecules. Vibrations which adapt the molecular dipole moment are visible in the infrared. Different forms of vibrations are discriminated, resulting in different, binding forces and binding angles of atoms in a molecule .Thus, complex molecules show numerous options of internal vibrations (Buffeteau *et al.*, 1996).

In the identification of unknown compounds. It has been worldwide accepted that the position of absorption bands is expressed in wavenumbers [cm^{-1}], which is the reciprocal of the wavelength. Double and triple bonds will give higher wavenumbers than single bonds (Schmitt and Flemming, 1998).

2.6.1 FTIR Spectrometer

The attenuated total reflectance (ATR) accessory shines an Infrared laser at the angle of total reflectance into a crystal. Some of the light escapes the crystal and reflects off a sample sitting on top of the crystal. The computer attached to the FTIR performs a Fourier transform on the energy reflected back through the crystal, with and without a sample, displaying the characteristic spectrum for the sample (Gremlich and Gunzler, 2002).

A general FTIR spectrometer consists of five components which are light source, interferometer, sample compartment, detector and a computer. Figure 2 shows the schematic of FTIR spectrometer.

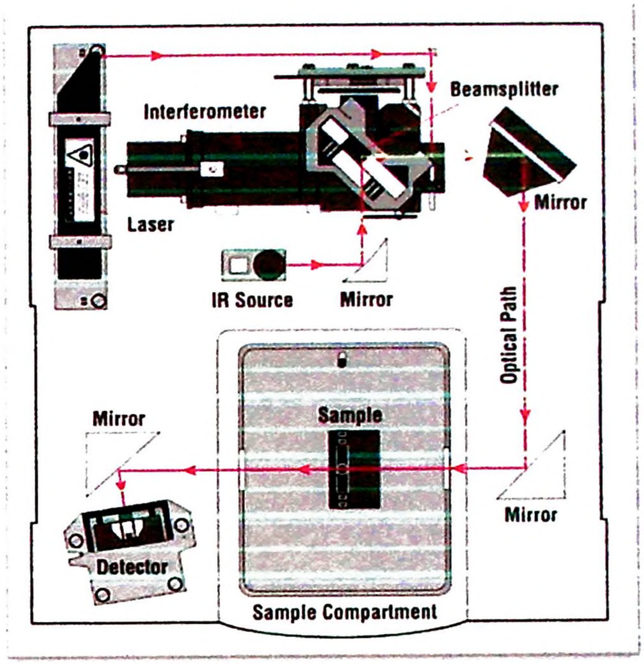

Figure 2.1: Schematic of FTIR spectrometer

(http://chemwiki.ucdavis.edu/Physical_Chemistry/Spectroscopy/Fundamentals/The_Power_of_the_Fourier_Transform_for_Spectroscopists accessed on 22 May 2015)

2.6.1.1 Light Source

Infrared energy is released from a glowing black-body source. This beam passes through an aperture which controls the amount of energy presented to the sample.

2.6.1.2 Interferometer

The beam enters the interferometer where the spectral encoding takes place. The interferogram signal are produce then exits the interferometer.

2.6.1.3 Sample Compartment

Depending on the type of analysis being ,the beam enters the sample compartment where it is transmitted through or reflected off of the surface of the sample. This is where specific frequencies of energy, which are uniquely characteristic of the sample, are absorbed.

2.6.1.4 Detector

The beam finally passes to the detector for final measurement. The detectors used are specially designed to measure the special interferogram signal.

2.6.1.5 Computer

The measured signal is digitized and sent to the computer where the Fourier transformation takes place. The final infrared spectrum is then presented to the user for interpretation and any further manipulation (Thermo Nicolet Corporation, 2001)

2.7 History of Pearson Product Moment Correlation (PPMC)

Mostly people belief that Karl Pearson developed this statistical measure himself. Although Pearson did develop a rigorous treatment of the mathematics of the Pearson Product Moment Correlation (PPMC), it was the imagination of Sir Francis Galton that originally conceived modern notions of correlation and regression. Galton, a cousin of Charles Darwin and an accomplished 19th century scientist in his own right, has fascination with genetics and heredity provided the initial inspiration that led to regression and the PPMC (Stanton, 2001)

Galton's first insights about regression sprang from a two-dimensional diagram plotting the sizes of daughter peas against the sizes of mother peas. As described below, Galton used this representation of his data to illustrate basic foundations of what statisticians still call regression. Many instructors may also feel more comfortable starting with correlation and building up to regression .According to Rodgers and Nicewander (1988), the empirical and theoretical developments that defined regression and correlation as statistical topics were presented by Sir Francis Galton in 1885. Then, In 1895, Karl Pearson published Pearson's *r*.

In the late 19th century, Sir Francis Galton was measuring many different biological and sociological variables and describing their distributions using the existing methods for single variables. However, there was no way to quantify the degree of relationship between two variables. In 1896, Karl Pearson create Pearson Product Moment Correlation to address this need. The coefficient is one of the most frequently employed statistical methods in the social and behavioral sciences and is frequently used in theory testing, instrument validation, reliability estimation, and many other descriptive and inferential procedures (Walk and Rupp, 2010).

2.7.1 Pearson Product Moment Correlation (PPMC)

Pearson product-moment correlation (PPMC) is a measure of the strength of the linear relationship between two variables. Sometimes it is reffered as Pearson Moment Correlation or Pearson Coefficient. The relationship between the variable must be linear whereas the correlation coefficient does not adequately represent the strength of the relationship between the variables. The symbol for Pearson correlation is "*p*" when measure in population and "*r*" when measured in a sample (Lane, 2014)

Correlation coefficient is defined as number or function that indicates the degree of correlation between two sets of data or between two random variables and that is equal to their covariance divided by the product of their standard deviations (Merriam-Webster, 2015)

Pearson's *r* can range from -1 to 1. An *r* of -1 indicates a perfect negative linear relationship between variables means as the value of one variable increase the other variable is decrease. When the value of *r* is 0 it indicates no linear relationship between variables meaning there is no association between the two variables.Lastly, when the value of *r* is +1 it indicates a perfect positive linear relationship between variables. That means when the value of one variable increase the other variable also increase (Laerd Statistics, 2013)

In recent years, statistical treatments of the data have been additionally planned in order to improve the analysis and the interpretation of the result. Pearson correlation coefficients were applied for the objective comparison of 30 black gel pen inks analyzed by laser desorption ionization mass spectrometry (LDI-MS) to ensure that no false differentiation occur. The mass spectra were obtained for ink lines directly on paper using positive and negative ion modes at several laser intensities. This methodology has the advantage of taking into account the reproducibility of the results as well as the variability between spectra of different inks. The influence of brands was also found to be minimal (Weyermann *et al.*, 2012).

2.8 Chemometric Techniques

Principle Component Analysis (PCA) is used to identify patterns in data and to highlight their similarities and differences. Original data from instrumental analysis such as UV-Vis spectra, Fourier Transformed Infrared spectra, gas and liquid chromatograms are being explained using new variables which are known as principal components (PCs). They are derived from linear combinations of the original variables of the original data with its specific loading (Ismail *et al.*, 2014)

Principal Component Analysis (PCA) is a technique which uses sophisticated basic mathematical principles to transforms a number of possibly correlated variables into a smaller number of variables called principal components. PCA has been called one of the most important outcomes from applied linear algebra and perhaps its most common use is as the first step in trying to analyze large data sets. (Richardson, 2009).

In PCA, the Principle Component (PCs) or the new variables are arranged in decreasing order of importance where the first PC accounts for the largest variation in the dataset followed by the second, third and so on. Values in the range of 70 – 90% are suggested to describe the original dataset of PCs. PCA generates two smaller matrices known as loading and score matrices. The latter is of interest as a plot known as score plot can be constructed out of it where groupings or clustering of the objects are produced (Ismail *et al.*, 2014)

2.8.1 Principle Component Analysis (PCA)

Thanasoulias et al. have working 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. 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 move a small portion of the sample into the solvent (ethanol) . This study showed a 100% correct arrangement of the training dataset between inks of different brands (Thanasoulias *et al.*, 2003)

Pattern recognition techniques are methods that are used to successfully determine embedded pattern or useful information in multidimensional datasets. Chemometric technique is one of the pattern recognition techniques which is an approach to analyzing chemical data using both mathematical and statistical methods to determine the core properties of objects or substances that are hard to be measured (Webb, 2002).

When Chemometrics technique of PCA has been coupled with UV-Vis spectroscopy (Thanasoulias *et al.*, 2003; Adam *et al.*, 2008; Senior *et al.*, 2012), HPLC and IR spectroscopy (Kher *et al.*, 2006) as well as ToF-SIMS (Denman *et al.*, 2009; Alamilla *et al.*, 2013) and have shown good discrimination of not only the "simulated inks" but also inks obtained from real cases. As conclusion whatever the combinations of analytical techniques when coupled with chemometrics can become a formidable weapon that can be used by of forensic scientists to combat crime related to documents.