The Malaysian Journal of ANALYTICAL SCIENCES

APPLICATION OF CHEMOMETRIC TECHNIQUES TO COLORIMETRIC DATA IN CLASSIFYING AUTOMOBILE PAINT

(Aplikasi Teknik-Teknik Kimometrik untuk Data Kolorimetrik bagi Pengkelasan Cat Kereta)

Nur Awatif Rosli[1], Rozita Osman[1]*, Norashikin Saim[1], Mohd Zuli Jaafar[2]

[1]Faculty of Applied Sciences,
Universiti Teknologi MARA Shah Alam, 40450 Shah Alam, Selangor, Malaysia
[2]Faculty of Plantation and Agrotechnology,
Universiti Teknologi MARA, 77300 Merlimau, Melaka

**Corresponding author: rozit471@salam.uitm.edu.my*

Received: 23 November 2014; Accepted: 27 June 2015

Abstract

The analysis of paint chips is of great interest to forensic investigators, particularly in the examination of hit-and run cases. This study proposes a direct and rapid method in classifying automobile paint samples based on colorimetric data sets; absorption value, reflectance value, luminosity value (L), degree of redness (a) and degree of yellowness (b) obtained from video spectral comparator (VSC) technique. A total of 42 automobile paint samples from 7 manufacturers were analysed. The colorimetric datasets obtained from VSC analysis were subjected to chemometric technique namely cluster analysis (CA) and principal component analysis (PCA). Based on CA, 5 clusters were generated; Cluster 1 consisted of silver color, cluster 2 consisted of white color, cluster 3 consisted of blue and black colors, cluster 4 consisted of red color and cluster 5 consisted of light blue color. PCA resulted in two latent factors explaining 95.58 % of the total variance, enabled to group the 42 automobile paints into five groups. Chemometric application on colorimetric datasets provide meaningful classification of automobile paints based on their tone colour (L, a, b) and light intensity These approaches have the potential to ease the interpretation of complex spectral data involving a large number of comparisons.

Keywords: colorimetric, Cluster Analysis (CA), chemometric techniques, Principle Component Analysis (PCA), Video Spectral Comparator (VSC)

Abstrak

Analisis cip cat adalah penting kepada penyiasat forensik, khususnya dalam kes-kes langgar dan lari. Kajian ini mencadangkan satu kaedah yang cepat dalam mengklasifikasikan sampel cat kereta berdasarkan set data kolorimetri; nilai penyerapan, nilai pantulan, nilai kilauan (L), darjah kemerahan (a) dan darjah kekuningan (b) yang diperoleh daripada teknik pembandingan spektrum video (VSC). Sebanyak 42 sampel cat kereta daripada 7 pengeluar kereta telah dianalisis. Set data kolorimetri yang diperolehi dari VSC dianalisa menggunakan kaedah kimometrik iaitu analisis kelompok (HACA) dan analisis komponen utama (PCA). Berdasarkan HACA, 5 kelompok cat kereta telah dijana; kelompok 1 terdiri daripada warna perak, kelompok 2 terdiri daripada warna putih, kelompok 3 terdiri daripada gabungan warna biru dan hitam, kelompok 4 terdiri daripada warna merah dan kelompok 5 terdiri daripada warna biru muda. PCA menghasilkan dua faktor yang menjelaskan 95.58% daripada keseluruhan varian, membolehkan 43 sampel cat kereta yang dianalisis dikumpulkan ke dalam lima kumpulan. Penggunaan kimometrik terhadap set data kolorimetri memberikan maklumat berguna untuk pengkelasan cat kereta berdasarkan tona warna (L, a, b) dan keamatan cahaya (serapan dan kepantulan). Pendekatan-pendekatan ini mempunyai potensi yang memudahkan penaksiran data spektrum yang kompleks yang melibatkan pelbagai aspek perbandingan.

Kata kunci: kolorimetri, Analisis Pengumpulan Kelompok-Kelompok Berhierarki (HACA), teknik-teknik kimometrik, Analisis Komponen Utama (PCA), Pembanding Spectrum Video (VSC)

Introduction

The examination and identification of automobile paints is crucial in road accident investigations. In many instances, automobile paint is the most significant form of trace contact evidence encountered at the incident scene; thereby adding increased importance to its subsequent analysis [1]. It is often necessary to compare the paint fragment found on the road or on clothing of a car accident victim with the paint sample coming from the suspected car in order to establish whether they have the same origin. On the other hand, if no comparative material is available, like hit-and-run accidents, the identification of the paint components may provide information of the kind of paint and its producer [2]. Automobile paint is a complex multi-component system designed to protect the frame of the vehicle and achieve the desired color and finish. Each layer has a characteristic function and comprised of a distinctive permutation of binders, resins, pigments and additives [3]. The analytical approach towards forensic paint examination relies on the analysis of each layer via a combination of microscopic and spectroscopic techniques are used to characterise paint traces [4,5].

Colorimetry is the science and technology used to quantify and describe physically the human colour perception, which was also established by CIE (Commission Internationale de I'Eclairage) in 1931. Even though limitations are well recognized, the CIE system of colorimetry remains as the only internationally agreed metric for clot measurement [6]. Analysis of automobile paints was done using video spectral comparator (VSC) using quantitative color description with *L*a*b** or CIELAB color system. The concept of color can be divided into two parts: brightness and chromaticity. For example, the color white is a bright color, while the color grey is considered to be a less bright of that same white. In other words, the chromaticity of white and grey are the same while their brightness differs. The 'brightness' (L*, lightness, luminosity or luminance) can range from '0' (black) to '100' (white), and constitutes the vertical axis of a three-dimensional color space. There are two accompanying horizontal axes of 'chromaticity': red-green (+*a** to -*a**), and yellow-blue (+*b** to -*b**). Negative values indicate the 'cold' colors green and blue, and positive integers the 'warm' colors red and yellow [7-9]. The *L*a*b** color space system is used in industry [8] and in meat production [10-12], to define standards for the quality control of products [9].

Chemometrics can be defined as the application of mathematical and statistical techniques used to retrieve more information from complex databases. This technique is an indispensable tool for elucidating patterns in complicated data matrices [13,14] and eliminates the subjectivity associated with the visual comparison of complex data from questioned and known samples, as it ensures that quantitative, impartial measures of the data obtained. Ultimately, this form of statistical analysis may allow the user to obtain information about the vehicle form an unknown paint sample (such as make, model or provenance), based upon its classification into one of the pre specified groupings [4].

In this study, the main objective was to classify automobile paints of various colors using colorimetric data sets (absorption value, reflectance value, luminosity value (L), degree of redness (a) and degree of yellowness (b) obtained from video spectral comparator (VSC) analysis. Since the data obtained had multivariate nature and several variables may be correlated, chemometric methods such as hierarchical agglomerative clusters analysis (HACA) and principal component analysis (PCA) were employed for better interpretation of the data. These techniques have proven to be efficient methods for analysing large data sets in environmental, chemical, biological and ecotoxicological case studies [15].

Materials and Methods

Samples Preparation

Chips of 42 automobile paint samples of various colors and manufacturers were examined as shown in Table 1. The manufacturer, model and color of each vehicle were recorded. A scalpel was used to pry paint chips off the underlying metal, ensuring all layers were present in the sample. The chips of paint were then preserved in bottles to prevent from contaminations.

Table 1. Sample identification for automotive paint samples

Number	Manufacturer	Model	Color	Sample ID
1	Proton	Saga	Black	ABK
2			Red	AR
3			Silver	AS
4			Light Blue	ABL
5		Wira	Blue	ABL
6			White	AW
7			Red	AR
8			Silver	AS
9			Light Blue	ALB
10		Satria Neo	Black	ABK
11		Persona	Red	AR
12		Perdana	White	AW
13		Iswara	Silver	AS
14	Perodua	Myvi	Black	BBK
15			Black	BBK
16			White	BW
17			White	BW
18			Light Blue	BLB
19		Kelisa	Blue	BBL
20			Silver	BS
21		Kancil	White	BW
22			Silver	BS
23		Viva Elite	Blue	BBL
24	Mitsubishi	Lancer	Black	CBK
25			Red	CR
26	Honda	Civic	Red	DR
27			White	DW
28		Jazz	Light Blue	DLB
29		CRV	Light Blue	DLB
30	Toyota	Vios	Black	EBK
31			Silver	ES
32		Estima	Black	EBK
33			White	EW
34		Fortune	Blue	EBL
35		Avanza	Red	ER
36		Corolla SE	Silver	ES
37	Nissan	Serena	Silver	FS
38			Light Blue	FLB
39	Suzuki	Swift	Black	GBK
40	Unknown	1	Black	UBK
41		2	Silver	US
42		3	Silver	US

Video Spectral Comparator (VSC) Analysis

All automobile paint samples were analysed using video spectral comparator (VSC 5000, Foster and Freeman, UK) based on CIE system (International Commission on Illumination, Vienna) expressed as L*, a* and b* which representing lightness (L*), redness (+a*), greenness (-a*), yellowness (+b*) and blueness (-b*). Absorption and reflectance spectra were generated from excitation wavelengths ranging from 400-1000 nm using VSC-5000-HR (Foster and Freeman, UK).

Chemometric Methods: Principal Component Analysis (PCA)

PCA is a multivariate statistical technique commonly used to display and analyse the structure of multivariate data representing the original dataset in a new reference system characterised by new orthogonal variables called principal components (PCs). Principal component analysis aims to explain the maximum amount of variance available in the original dataset using as few PCs as possible. The PCs can be used for an effective representation of the system under investigation, with a lower number of variables than the original case. The coordinates of the samples in the new reference system are called scores, while the coefficients of the linear combination describing each PC, i.e. the weights of the original variables on each PC, are called loadings. The graphical representation of the scores allows the identification of groups of samples showing similar behaviors (samples close to each other in the graph) or different characteristics (samples far from each other).

Hierarchical Agglomerative Clusters Analysis (HACA)

Cluster analysis is an unsupervised pattern recognition technique that uncovers intrinsic structure or underlying behavior of a data set without making a prior assumption about the data, in order to classify the objects of the system into categories or cluster based on their similarity [16]. Hierarchical agglomerative clustering is the most common approach in which clusters are formed sequentially, by starting with the most similar pair of objects and forming higher clusters step by step. The Euclidean distance usually gives the similarity between two samples and a distance can be represented by the difference between analytical values from both the samples [17]. In this paper the agglomerative hierarchical method was applied. In the hierarchical method objects are grouped on the base of inter-object distances. The statistical program used to compute chemometric analysis (PCA, HACA) using XL Stat 2014 and Excel 2007.

Results and Discussion

The purpose of this study was to classify the automobile paints based on colorimetric data sets using chemometric approach. Automobile paints of specific color gave specific colorimetric data, which can be used to identify the color of paint samples.

Video Spectral Comparator (VSC)

Colorimetric data sets obtained by VSC consisted of absorption value (Abs %), reflectance value (Ref %), luminosity value (L*), degree of redness (a*) and degree of yellowness (b*). Table 2 shows the colorimetric data sets for automobile paints samples. High absorption values (89-99%) for black, blue and red paint samples were observed compared to those of white, silver and light blue paint samples. Lightness (*L**) values for all samples decreased with increased in absorption value (Abs %). In the CIELab system, parameter *a** demonstrated the degree of redness. Positive a* values obtained from black, blue and red paint samples showed that redness tone was present in these samples. Red paint samples gave the highest *a** value (27-52). Negative a* values were observed in silver, white and light blue showed a greenish tone in these colors. Parameter b* was the most discriminant variable for the paint samples as shown Table 2. The physical meaning of these positive b* values are related to yellow paint color: the higher b* is, the greater the yellow intensity will be [18].

Table 2. Colorimetric data for automobile paints of different colors

Colour	Abs %	L*	a*	b*
Black	91	36.3	2	-12
	95	27.7	5.5	-21.1
	99	11	11.7	-36.2
Blue	93	34.1	5.8	-20.7
	96	26.7	7.3	-22.9
	96	24.2	7.4	-27.3
Red	89	48.4	27.4	0.4
	95	46.7	52.5	11.7
	97	34.9	39.3	0.2
Silver	10	95.6	-2.9	1.4
	46	79.5	1	-3.9
	61	68.4	-11.1	-5.8
White	8	96.9	-0.8	2
	27	88	-1	3
	33	86.1	1.6	-6.2
Light Blue	70	61.4	-7.2	-22.4
	72	61.5	2.3	-20.2
	77	57.5	1.5	-26.5

L*: luminosity value, a*: degree of redness and b*: degree of yellowness

Chemometric Analysis: Principal component analysis (PCA)

PCA is the most powerful pattern recognition technique that is usually coupled with HACA. It provides information on the most significant parameters, which describes the whole data set by excluding the less significant parameters with minimum loss of original information [19-21]. In this study, PCA was applied to the colorimetric data sets to compare the compositional patterns between the automobile paint samples and to identify the factors that influence their variation. PCA of the automobile paint samples data revealed that 95.58% of the total variance in the data was accounted for in the first two PCs. Figure 2a shows the 2-dimensional scores plot, which is a plot of the projected sample using PC1 and PC2 as a new coordinate system, revealed 5 distinct classes are present based on the colour of the automobile paints (red, black + blue, light blue, silver, silver + white. The first PC (F1) accounting for 58.53% of the total variance was correlated with absorbance value, reflectance value, luminosity value (absorbance) and luminosity value (reflectance). The second PC (PC2) accounting for 37.06% of the total variance was correlated with degree of redness (absorbance), degree of redness (reflectance), degree of yellowness (absorbance) and degree of yellowness (reflectance). The scree plot (Figure 2b) was used to identify the number of PCs to be retained in order to comprehend the underlying data structure [16,20,22]. In the present study, the scree plot showed a pronounced change of slope after the second principal component. (F2).

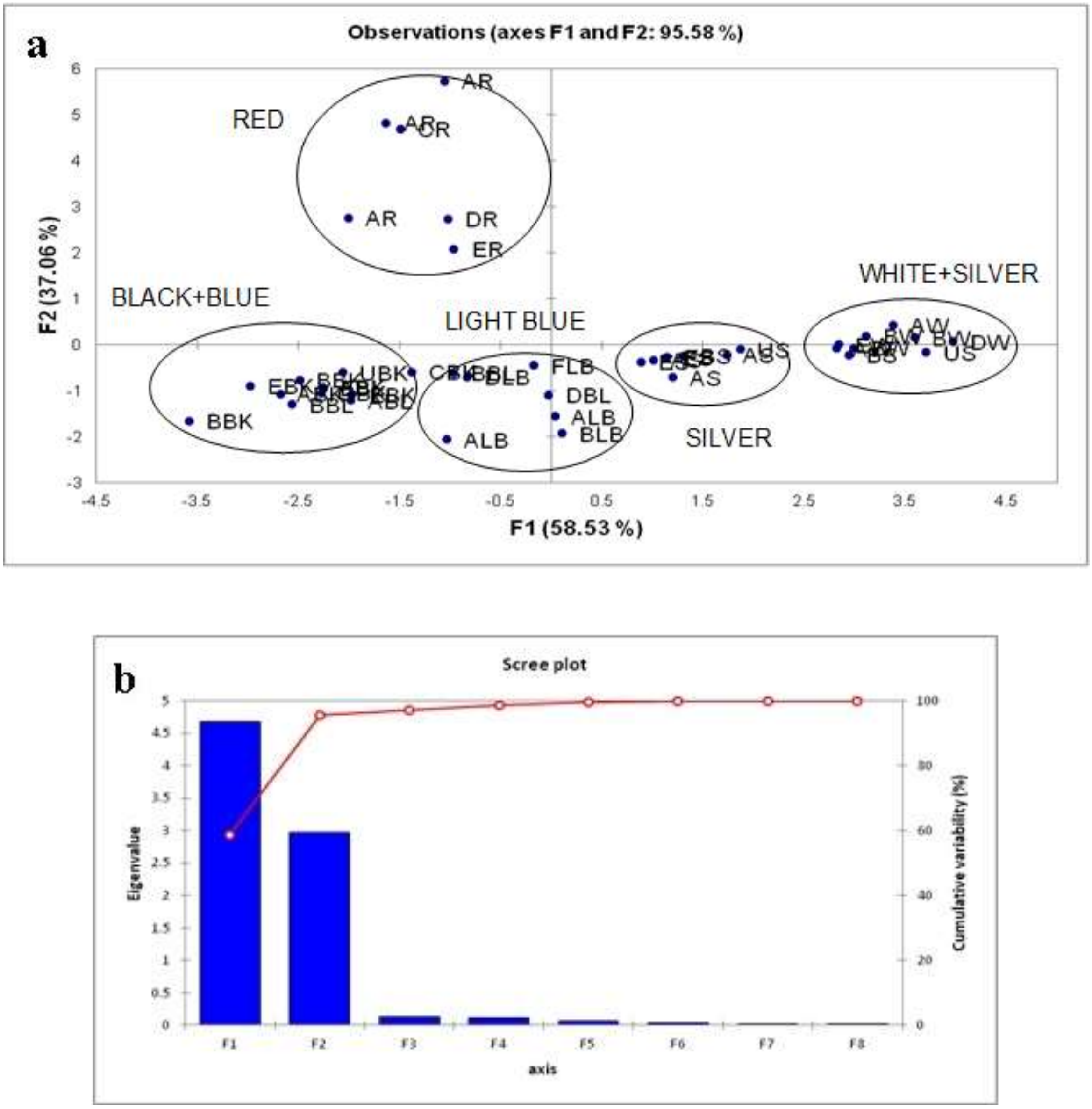

Figure 1. PCA (a) loadings and (b) scree plot for the first two principal components obtained for all automobile paints.

Based on PCA results, only the two principal components with eigenvalue >1 was considered to perform varimax rotations. The parameter loading for the two components from PCA of the data sets after varimax rotation (VFs) is given in Table 3. Rotation of the axis defined by PCA produced a new set of factors, each one involving primarily a subset of the original variables with as little overlap as possible, so that the original variables are divided into groups somewhat independent of each other. Thus, factor analysis of the present data set of automobile paints further reduced the contribution of less significant variables obtained from PCA. Liu et al. (2003) classified the factor loadings as ‘strong’, ‘moderate’ and ‘weak’ corresponding to absolute loading values of > 0.75, 0.75-0.50 and 0.50-0.30, respectively.

VF1, which explains 58.51% of total variance has strong positive loadings on reflectance value (ref %), luminosity value (absorbance), (abs L) and luminosity value (reflectance), (ref L) and strong negative loading on absorbance value (abs %), which can be interpreted as the important factors to classify the color of automobile paint samples. VF2, which explain 37.07% of total variance has strong positive loadings on degree of redness (absorbance) (abs A), degree of redness (reflectance) (red A), degree of yellowness (absorbance) (abs B) and degree of yellowness (reflectance) (ref B).

Table 3. Loadings of VSC data sets on the first two rotated PCs for automobile paint samples

Variables	**VF1**	**VF2**
Abs %	**-0.964**	0.133
Ref %	**0.970**	-0.120
Abs L*	**0.980**	0.024
Ref L*	**0.978**	0.055
Abs a*	-0.405	**0.897**
Ref a*	-0.412	**0.898**
Abs b*	0.530	**0.808**
Ref b*	0.527	**0.816**
Eigenvalue	4.682	2.964
% Total variance	58.514	37.068
Cumulative % variance	58.514	95.582

Bold values represent strong loadings (>0.75), Absorbance value (Abs %), reflectance value (Ref %), Luminosity absorbance (Abs L*), Luminosity reflectance (Ref L*), Degree of redness absorbance (Abs a*), Degree of redness reflectance (Ref a*), Degree of yellowness absorbance (Abs b*), Degree of yellowness reflectance (Ref b*)

Hierarchical agglomerative cluster analysis (HACA)

HACA was applied to the data sets of automobile paint samples of different colours and manufactures, in order to identify similarities in paint components between the samples. The HACA analysis grouped the colorimetric data of 42 automobile paint data sets into five clusters (Figure 2). In general, all samples from different car manufactures can be grouped according to their colors. Cluster 1 consisted of silver automobile paint color. It can be seen that cluster 2 and 3 contain a combination of white+silver and black+blue colours, respectively. The similarity of choromaticity value of some blue, black and white, silver paint colors resulted in overlapping of their group. Cluster 4 and cluster 5 consisted red and light blue automobile paints, respectively. The finding of this study suggested that the automobile paints of various manufacturers could be grouped based on its color. Therefore, VSC technique can be used as a reliable screening approach in the classification of automobile paints. The results could be improved by combining VSC technique and spectroscopy techniques such as fourier transform infrared (FTIR) spectroscopy in order to get a systematic method in classifying and discriminating automobile paints.

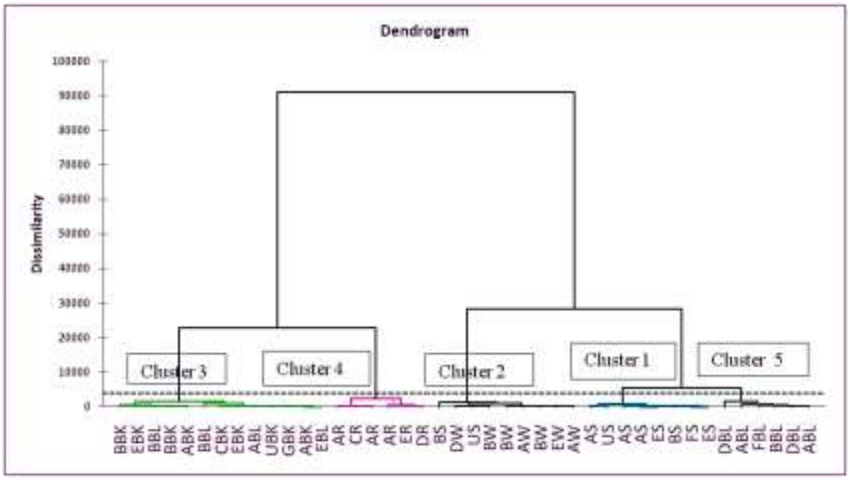

Figure 2. Dendrogram showing clustering of color of automobile paints

Conclusion

Using VSC analysis, automobile paints samples generated a complex multidimensional data sets. Thus, chemometric approach was applied for a systematic classification and discrimination of their colors. This approach is rapid and uses non-destructive technique for the analysis of automobile paints. From the results obtained by the PCA and HACA, automobile paint samples were grouped in 5 clusters according to their color tone (L, a, b) and light intensity (absorbance and reflectance). This study showed that VSC is a reliable technique in verifying similarities and dissimilarities in colors of paint chips in forensic investigation.

Acknowledgement

The author would like to acknowledge the financial support obtained from Ministry of Higher Education (MOHE), Malaysia for this project (Project number: 600-RMI/ERGS 5/3 95/2012) and Universiti Teknologi MARA.

References

1. Maric, M., Bronswijk, W.v., Lewis, S.W. and Pitts, K. (2014). Synchrotron FTIR characteristic of automotive primer surfacer paint coatings for forensic purposes. *Talanta* 118: 156-161.
2. Zięba-Palus, J. (1999). Application of micro-Fourier transform infrared spectroscopy to the examination of paint samples. *Journal of Molecular Structure*, 511-512: 327-335.
3. Bentley, J. (2001). Composition, manufacture and use of paint. In: B. Caddy (Ed.), *Forensic Examination of Glass and Paint: Analysis and Interpretation* (pp. 123-141). Taylor & Francis, London.
4. Maric, M., Bronswijk, W.v., Lewis, S.W., Pitts, K. and Martin D.E (2013). Characterisation of chemical component migration in automotive paint by synchrotron infrared imaging. *Forensic Science International* 228: 165-169.
5. Zięba-Palus, J. and Michalska, A. (2014). Photobleachig as a useful technique in reducing of fluorescence in Raman spectra of blue automobile paint samples. *Vibrational Spectroscopy*74: 6-12.
6. Herendeen, D. L., Allen, J.A. and Lera T. (2011). Philatelic shade discrimination based on measured colour. Institute for Analytical Philately, Inc. *London Philatelist,* 120 (1384): 105-117.
7. Wyszecki, G. and Stiles, W.S. (2000). Colorimetry. In: Wyszecki, G., Stiles, W.S. (Eds.), *Color Science. Concepts and Methods, Quantitative Data and Formulae* (pp. 117-248). New York, USA: Second Ed. John Wiley & Sons.
8. Hirschler, R., (2010). Electronic colour communication in the textile and apparel industry. *Redige,* 1: 43-61.
9. Neuhauser, S. and Handler, J. (2013). Colour analysis of the equine endometrium: Comparison of spectrophotometry and computer-assisted analysis of photographs within the *L*a*b** colour space system. *The Veterinary Journal,* 197: 753-760.

10. Hammell, K.L., Laforest, J.P. and Dufour, J.J. (1994). Evaluation of the lean meat colour of commercial pigs produced in Quebec. *Canadian Journal of Animal Science,* 74: 443-449.
11. Northcutt, J.K., Buhr, R.J. and Rowland, G.N. (2000). Relationship of broiler bruise age to appearance and tissue histological characteristics. *Journal of Applied Poultry Research, 9,* 13-20.
12. Chatelain, Y., Guggisberg, D., Dufey, P.A., Vergères, G. and Hadorn, R. (2007). farbmessung an Fleisch und fleischerzeugnissen. *ALP Science* 507: 1-23.
13. Mok, D.K.W. and Chau, F.T. (2006). Chemical information of Chinese medicines: a challenge to chemist, Chemometric *Intelligent Laboratory System* 82: 210-217.
14. Maree, J., Kamatou, G., Gibbons, S., Viljoen A. and Vuuren, S.V. (2014). The application of GC-MS combined with chemometrics for the identification of antimicrobial compounds from selected commercial essential oils. *Chemometrics and Intelligent Laboratory Systems* 130: 1-10.
15. Singh, K. P., Malik, A., Mohan, D., Sinha and Singh, V. K. (2005). Chemometric data analysis of pollutanys in wastewater-a case study. *Analytica Chimica Acta* 532: 15-25.
16. Vega, M., Pardo, R., Barrado, E. and Deban, L. (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research* 32: 3581-3592.
17. Otto, M. (1998). Multivariate methods. In R. Kellner, J.M. Mermet, M. Otto and H.M. Widmer (Eds.), *Analytical Chemistry* (pp. 916). Wiley-VCH, Weinheim, Germany.
18. Alonso-Slaces, R.M., Guyot, S., Herrero, C., Berrueta, L.A., Drilleau, J-F., Gallo, B. and Vicente, F. (2005). Chemometric classification of Basque and French ciders based on their total polyphenol contents and CIELab parameters. *Food Chemistry* 91: 91-98.
19. Helena, B., Pardo, R., Vega, M., Barrado, E., Fernandez, J. M. and Fernan, L. (2000). Temporal evaluation of groundwater composition in an alluvial aquifer (Pisurga River, Spain) by Principle Component Analysis. *Water Research* 34(3): 807-816.
20. Singh, K. P., Malik, A., Mohan, D. and Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)-a case study. *Water Research* 38: 3980-3992.
21. Kannel, P. R., Lee, S., Kanel, S. R., and Khan, S. P. (2007). Chemometric application in classifiction and assessment of monitoring locations of an urban river system. *Analytica Chimica Acta* 582: 390–399.
22. Jackson, J. E. (1991). A User's Guide to Principal Components. Wiley, New York.
23. Liu, C. W., Lin, K.H. and Kuo, Y. M. (2003). Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Science Total Environment* 313: 77-89.