Unsupervised grouping of industrial textile dyes using K-means algorithm and optical fibre spectroscopy

Ana M. Cubillas[*], Olga M. Conde, Pedro Anuarbe, Antonio Quintela and Jose M. Lopez-Higuera
Photonics Engineering Group, University of Cantabria, Avda. Castros S/N, 39005 Santander, Spain

ABSTRACT

A method for the unsupervised clustering of optically thick textile dyes based on their spectral properties is demonstrated in this paper. The system utilizes optical fibre sensor techniques in the Ultraviolet-Visible-Near Infrared (UV-Vis-NIR) to evaluate the absorption spectrum and thus the colour of textile dyes. A multivariate method is first applied to calculate the optimum dilution factor needed to reduce the high absorbance of the dye samples. Then, the grouping algorithm used combines Principal Component Analysis (PCA), for data compression, and K-means for unsupervised clustering of the different dyes. The feasibility of the proposed method for textile applications is also discussed in the paper.

Keywords: Optical Fibre Sensor, Absorption spectroscopy, Textile dyes, Principal Component Analysis, K-means clustering.

1. INTRODUCTION

Colour matching assessment is a very important requirement in textile industry, especially in the dyeing stage. In this case, if the colour produced is different from the colour specified in the dyeing recipe, the manufactured goods have to be reworked or rejected [1]. This results in significant additional costs and an increased dumping of toxic wastes into the environment. Conventional colour determination employs colorimeters or spectrophotometers for this purpose. Colorimeters provide three tristimulus values, i.e. the CIELAB values [2], whereas spectrophotometers measure the reflectance curve and compute the CIELAB values from it. In both cases, the CIELAB values are calculated after the dyeing stage. Therefore, a method to assess the colour of the dyebath prior the dyeing process could be advantageous in textile industry.

An alternative way to determine the colour could be by analyzing the absorption spectrum using optical fiber sensor technology. Indeed, colour can be completely characterized by the absorption spectrum in the visible region. Additionally, the absorption spectrum contains extended spectral information than the CIELAB values and it is independent of the spectral power distribution and the observer chosen [3]. In this regard, optical fibre spectroscopy is a very well suited technique for monitoring the absorption of liquids. This technique has many advantages such as its simplicity and ease of operation, efficiency, high-performance and relatively low cost [4]. Optical fibre spectroscopy experiments have been previously reported for liquids such as olive oil, lubricant oil, beer and textile dyes [5-8].

In this work, a method for the grouping of textile dyes based on their optical spectral properties is demonstrated. For that purpose, seven different textile dyeing recipes that account for the most common errors in dyeing mixing are considered. Unsupervised clustering methods are used to classify the dyeing samples into groups in order to discover the similar and dissimilar samples [9-11]. For that purpose, K-means clustering algorithm is chosen as it is an efficient and simple method which has proved its validity in many applications [11]. The performance and feasibility of the method in textile applications are also assessed in this paper.

2. EXPERIMENTAL SETUP

The measurement setup shown in Fig. 1 was employed to obtain the transmission and absorption spectra of textile dyes. The different samples were placed inside a cuvette holder (Ocean Optics, CUV-UV) in disposable plastic cuvettes of 0.5 cm path light. Light from a deuterium-tungsten-halogen source (Ocean Optics, DH-2000) was coupled to the cuvette holder. The transmitted light through the system was measured using a high-resolution UV-Vis-NIR spectrometer

*cubillasam@unican.es; phone +34 942200877 (12); fax +34 942200877; website http://gif.teisa.unican.es

(Ocean Optics, HR2000+CG) in the 200-1100 nm range. Optical fibres of 600 μm core diameter were used to couple light between the light source, the cuvette holder and the spectrometer. The system was calibrated using a cuvette containing distilled water as blank reference, and a dark reference, which was taken when the light path to the spectrometer was blocked.
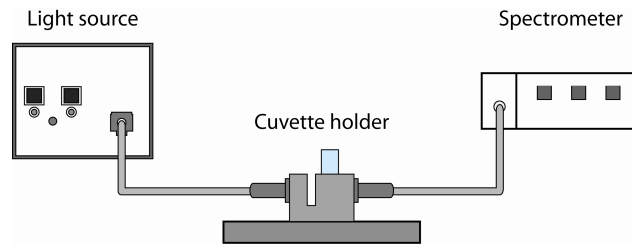

Fig. 1. Experimental setup used for the absorption spectrum measurements of textile dyes.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1 Characterization of textile dyes

Seven different dyeing recipes (F_1-F_7) of three reactive colorants extensively employed in the textile industry (Levafix Amber, Red and Blue [12]) were considered in the experiments (see Table 1). The colorants and dyeing recipes were provided by Textil Santanderina S.A. Specifically, F_1 is the dyeing recipe consisting of a mixture of the three colorants at different concentrations: Amber (C_1), Red (C_2) and Blue (C_3).The other dyeing recipes are specified so that they account for the most common mixing errors in textile dyeing processes: i.e. when one colorant is not present in the mixture (as in formulas F_2, F_3 and F_4) and when one colorant is in a wrong concentration (as in F_5, F_6 and F_7). Furthermore, these dyeing recipes were replicated in our laboratories with different precision scales, to take into account other experimental measurement errors. Therefore, the data set comprised 28 dye samples of 7 different dyeing recipes.

Table 1: Specifications of the different dyeing recipes used in the experiments.

Formula	Amber	Red	Blue
F_1	C_1	C_2	C_3
F_2	0	C_2	C_3
F_3	C_1	0	C_3
F_4	C_1	C_2	0
F_5	$C_1/2$	C_2	C_3
F_6	C_1	$C_2/2$	C_3
F_7	C_1	C_2	$C_3/2$

Textile dyes, as used in dyeing processes, are very thick and optically opaque, so they need to be diluted to allow their correct characterization. In our case, the total absorbance of the mixtures could reach values over 120 AU (AU, Absorbance Units). Due to the sensitivity of the system used, the spectral absorbance of the samples needs to be limited to 2 AU [13]. Therefore, the samples were diluted in distilled water an Optimum Dilution Factor (ODF), calculated as shown in Fig. 2. From the multivariate calibration models of the individual colorants composing the formulas, the overall spectrum absorption of the mixture was obtained following the procedure described in [14]. This step is shown as spectrum synthesis in Fig. 2. Then, the ODF was calculated as the dilution needed so that the total absorbance of the synthesized spectrum was below the absorption limit of the system (i.e. 2 AU).
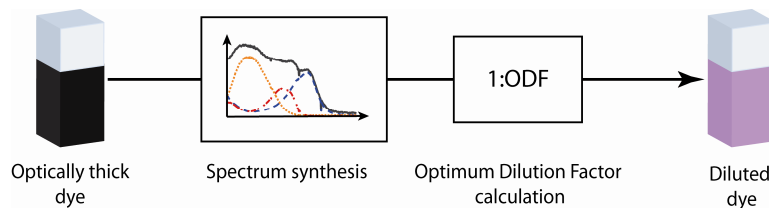

Fig. 2. Procedure followed to dilute the dyes an optimum factor.

Following this process, the different dye samples were diluted in distilled water and their absorption spectra obtained. More than 350 absorption measurements from the 28 dye samples were taken at different days and conditions. An example of the corrected absorption spectra of a set of dyes is shown in Fig. 3. As can be seen, there is a strong spectral overlap between the formulas, which difficulties their differentiation. Furthermore, the most information from the spectra is in the visible region, which is directly related to colour.
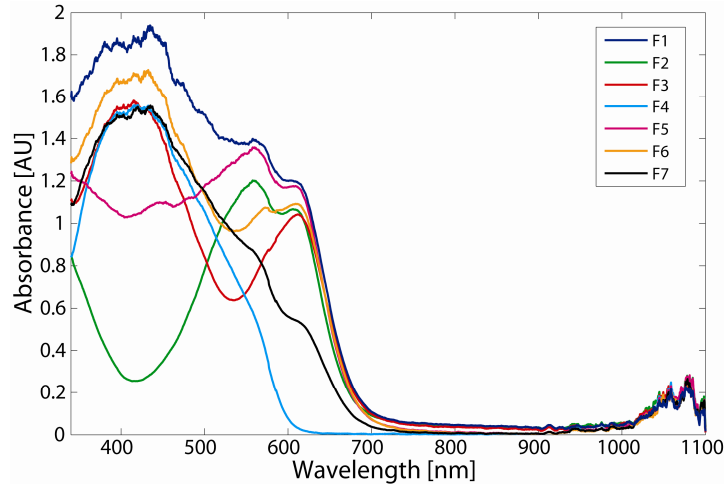

Fig. 3: Corrected absorption spectra of the different dyeing recipes.

3.2 Unsupervised grouping of textile dyes

The spectral grouping algorithm proposed for dye differentiation was a combination of PCA and K-means clustering. PCA was first used to reduce the dimensionality of the data set as each absorption spectrum consisted of 2048 components. As a result, two to three principal components (PCs) were sufficient to represent most of the variability in the data set. We chose to retain 2 PCs (65.3695%, 27.1736% of variance explained, respectively), which represented 92.5431% of the total variance.
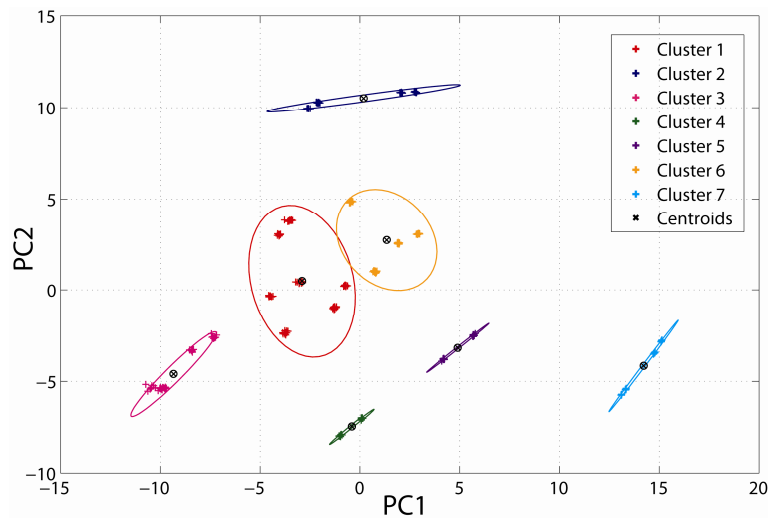

Fig. 4: Result of the PCA compression and K-means clustering.

K-means clustering was applied following PCA to sort and group the different samples based on their spectral attributes. K-means algorithm is an unsupervised centre-based clustering algorithm, in which the number of clusters (k) is assumed to be fixed [11]. It consists of two phases: the initialization phase, in which the algorithm randomly assigns the samples into the k clusters, and an iteration phase, in which the algorithm computes the distance between each sample and each cluster and assigns each sample to the nearest cluster. The iteration process finishes when an error optimization function

is minimized [11]. In our case, 7 clusters and Euclidean distance metric were chosen. The result of PCA and K-means clustering algorithms is shown in Fig. 4. In the figure, the different clusters are represented as a centroid and the boundary as two times the standard deviation of the points within the cluster. As can be seen, there is a clear grouping of samples into 7 clusters, some of them showing very little variability. However, two of them (the red and yellow clusters) are very close, due to the high variability of the samples within the clusters. These correspond to F6 and F3, respectively, which are very similar in composition and, therefore, in colour. It can be concluded that, with this method, the different dyeing recipes are clearly distinguishable by their absorption spectrum, which is also related to colour.

4. CONCLUSIONS

In this paper, a method for the unsupervised clustering of optically thick dyes with application to textile industry is demonstrated. Seven different textile dyeing recipes that account for the most common errors in dyeing mixing were considered. The dye samples were characterized using an extrinsic optical fibre spectroscopy setup in the UV-Vis-NIR in the 200-1100 nm range. Multivariate calibration methods were applied to calculate the optimum dilution factor needed to reduce the high absorbance of the dyes. PCA was first used to reduce the dimensionality of the absorption spectra of dyes and K-means algorithm was applied to sort and group the dye samples into clusters. As a result, the different dye recipes were successfully grouped into 7 clusters. This method could help to indentify dyes with similar absorption spectra, and thus colour.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Senthilkumar, M., "Modelling of CIELAB values in vinyl sulphone dye application using feed-forward neural networks," Dyes and Pigments, 75(2), 356-361 (2007).
[2]   Colorimetry, C., "CIE Publication No. 15.2," Commission Internationale de I'éclairage, Vienna, 19–20 (1986).
[3]   Dupont, D., "Study of the reconstruction of reflectance curves based on tristimulus values: comparison of methods of optimization," Color Research & Application, 27(2), 88-99 (2002).
[4]   Burns, D., and Ciurczak, E., [Handbook of near-infrared analysis] CRC, (2008).
[5]   Mignani, A., Smith, P., Ciaccheri, L., Cimato, A., and Sani, G., "Spectral nephelometry for making extravirgin olive oil fingerprints," Sensors and Actuators B: Chemical, 90(1-3), 157-162 (2003).
[6]   Mignani, A., Ciaccheri, L., Díaz-Herrera, N., Mencaglia, A., Ottevaere, H., Thienpont, H., Francalanci, S., Paccagnini, A., and Pavone, F., "Optical fiber spectroscopy for measuring quality indicators of lubricant oils," Measurement Science and Technology, 20(3), 4011 (2009).
[7]   Sahin, U., Ulgen, A., Kekec, A., and Gokmen, A., "Real-time monitoring of indigo concentrations in the dyebath with a laser diode spectrometer," Textile Research Journal, 74(3), 193 (2004).
[8]   Cubillas, A., Conde, O., Anuarbe, P., Gutierrez, M., Martinez, V., and Lopez-Higuera, J., "Optical fibre spectroscopy sensor for the quantitative determination of industrial textile dyes." Vol. 7503, 750322-750322-4 (2009).
[9]   Massart, D., and Vandeginste, B., [Handbook of chemometrics and qualimetrics] Elsevier Science, (1997).
[10]   Adams, M., [Chemometrics in analytical spectroscopy] Springer Verlag, (2004).
[11]   Gan, G., Ma, C., and Wu, J., [Data clustering: theory, algorithms, and applications], (2007).
[12]   [http://www.dystar.com].
[13]   Viggiano, J. A. S., [Spectrophotometric Analysis of Beer and Wort] Rochester Institute of Technology, (2006).
[14]   Conde, O., Cubillas, A., Anuarbe, P., Gutierrez, M., Martinez, V., and Lopez-Higuera, J., "Fiber-optic spectroscopic sensor for reactive dye mixture spectrum synthesis in textile industry." IEEE Sensors 2009, Institute of Electrical and Electronics Engineers, 136-139 (2009).