Plastic Identification and Comparison by Multivariate Techniques with Laser-Induced Breakdown Spectroscopy

Jesús Anzano,[1] Beatriz Bonilla,[1] Beatriz Montull-Ibor,[1] Justiniano Casas-González[2]

[1]*Laser Laboratory & Environment, Department of Analytical Chemistry, Faculty of Sciences, University of Zaragoza Pedro Cerbuna #12, 50009 Zaragoza, (34)976761292, Spain*
[2]*Department of Condensed Matter Physics, Faculty of Sciences, University of Zaragoza, Spain*

ABSTRACT: The classification of plastics is very important in the recycling industry. A quick online classification allows the installation of the equipment in this line of work. Whether qualitative or quantitative analysis, the basic component of any laser-induced breakdown spectroscopy (LIBS) measurement is the emission spectrum recorder from single plasma. Each fire of the laser atomizes a portion of the sample in the pulse focal volume and produces plasma that excites and re-excites the atoms to emit light. The plasma light is collected and recorded in an ensuing measurement. In this sense, the LIBS technique offers all possible advantages: speed, the possibility of online analysis, nondestructive analysis, and so on. In this article, we discuss details related to the analysis of the emission spectrum. The plastics used in this study were low-density polyethylene (PE), high-density PE, polypropylene, polystyrene, and poly(ethylene terephthalate). Hierarchical cluster analysis was proven to be the best method because the four polymers could be divided into two clusters, which allowed their identification and classification in a fast and easy way that could be carried out with commercial software and could be implanted online in a recycling factory, as conventional data analysis techniques are limited to the qualitative identification and calculation of elemental abundances. Principal component analysis on LIBS spectra can be used to better describe the chemical variations in the samples and to extract a greater understanding of the chemical structure. © 2011 Wiley Periodicals, Inc. J Appl Polym Sci 121: 2710–2716, 2011

Key words: laser ablation; plastics; polymer synthesis and characterization

INTRODUCTION

The sorting of plastics into categories is important because it is the first step in the recycling of plastic waste. The recycling of waste plastics has become an important issue because of the enormous use of plastics in every aspect of modern life.

Most current plastic sorting is carried out by hand. Manual sorting is a simple process that needs very little technology, but it leaves much room for improvement. This is a labor-intensive, costly, and inefficient method for sorting plastics. Furthermore, it is difficult to differentiate between the resin types used in packages through the visual means used in manual sorting. For this reason, the Society of the Plastics Industry instituted a voluntary labeling system.

Even with this labeling system, it is still difficult to distinguish polymer types manually because of the condition of the plastics as they reach the separation facility. The bottles may be crushed, cracked, or covered, and the resin label rendered useless. In this way, some automated methods had been developed that propose a series of treatments for the classification of the plastics on the basis of Raman's technique,[1] IR spectroscopy,[2] and of course, the flotation classic's technique[3] and electrostatic techniques.[4]

Raman's technique[1] has some disadvantages. For example, the Raman effect is very weak, so the detection needs sensitive and highly optimized instrumentation. Sample heating through the intense laser radiation can destroy the sample or cover the Raman spectrum and cannot be used for metals or alloys. Furthermore, the mathematical techniques used are more complex than those proposed in this article.

In our laboratory, in previous studies, we developed a new technique known as *laser-induced breakdown spectroscopy* (LIBS). This technique is based on the impact of a laser pulse on a solid sample. The

formed plasma emits light that is carried to a spectrograph, and measurement of the emission spectrum is made. One of the major advantages of the technique is its high speed because the laser pulse duration is just on the order of nanoseconds. Another advantage of the technique is its ease of implementation on online systems; this removes the necessity of carrying out sample pretreatment. This technique has been previously applied in the characterization of plastic-based films from cinematographic films.[5] Furthermore, double-pulse LIBS was used for the detection of plastic materials,[6] and it has been successfully applied in the study of many materials, such as archaeological samples[7] and glasses.[8] Also, previous studies have demonstrated the accuracy of the technique in plastics detection and metal analysis.[9]

LIBS is a technique that can provide all the necessary requirements: It is fast, can be used online, is nondestructive, and so on. The literature has indicated instant identification of postconsumer plastics compared with reference spectral libraries stored in a computer where simple statistical correlation methods, including linear and rank correlations, were used;[10] a method based on the ratios of the characteristic elements of the plastics;[11] and another method based on classification with distance libraries.

The application of multivariate techniques in LIBS was previously described, but not enough studies have been reported. In this way, principal component analysis (PCA) and cluster analysis (CA) were used by Amador-Hernández[12] to classify the spectral information provided by all of the components. Multivariate statistical analysis techniques were coupled with LIBS to identify preservative types and to predict elemental content in preservative-treated samples.[13] Partial least-squares regression was also used[14] to predict the composition of samples. Least-squares regression was used with the LIBS technique in the quantitative analysis of samples from the jewelry industry,[15] for their characterization,[16] and to study the matrix effects in steel samples.[17] PCA has also been applied in the forensic and environmental fields,[18] in soil sample analysis,[19] and for detection of biological contaminants on surfaces.[20]

In this article, a new, fast, and easy method for plastic classification with LIBS with PCA is described. PCA is a useful statistical technique that has found application in fields such as face recognition and image compression and is a common technique for finding patterns in data and expressing the data in such a way as to highlight their similarities and differences. We present a short explanation for the technique, but more details about it can be found elsewhere.[21,22] Furthermore, we present a complete analysis of plastic waste by hierarchical cluster and a comparison among the different cluster algorithms that can be used in this kind of analysis.

The group of commercial plastics, also termed *commodity plastics*, consists of the most used polymers in terms of volume and number of applications. They are mainly polystyrene (PS), polypropylene (PP), high-density polyethylene (PE), low-density PE, and poly(ethylene terephthalate) (PET). These four polymers (PET, PE, PP, and PS) are the only ones whose recycling has an economic interest,[23] so they are the plastics that we analyzed.

EXPERIMENTAL

Instrumental setup

The instrumentation used consisted of a Nd:YAG laser, an xyz stage carrying the sample (Standa 011957), a spectrograph, and an intensified charged coupled device detector. The Nd:YAG laser (Brilliant Quantel, Q-Switched; http://www.quantel.com) had a 115-mJ laser pulse energy at the second harmonic of 532 nm and a 4.4-ns pulse duration. The plasma light was collected and transported to a spectrograph by a lens and optical fiber. The lens position was adjusted by a diode laser (Andor, HE-OPI-0009). An Echelle spectrograph (Andor Mechelle ME5000, focal length = 195 mm, F/7, $\lambda/\Delta\lambda$ = 5000, spectral range = 200–975 nm; http://www.andor.com) was coupled to an intensified charged coupled device detector (Andor iStar DH734, 1024 × 1024 pixels, 13.6 × 13.6 μm^2/pixel, 18-mm intensification diameter). This system was calibrated with an Hg:Ar lamp (Ocean Optics, HG-1, Hg-Ar lines 253–922 nm; http://www.oceanoptics.com). An integration time of 500 ns was optimized.

Samples and procedure

The analyzed samples[24] were PE (samples 1–6), PET (samples 7–9), PP (samples 10–12), and PS (samples 13–15). Sample pretreatment was not necessary. An average of 10 shots in one point for each sample (the samples did not break) was used for the analysis.

Important characteristics of the selected plastics (PE, PET, PP, and PS) are as follows. Polyolefins, such as PE and PP, are simpler polymer structures that do not need plasticizers, although they do use additives such as UV and heat stabilizers, antioxidants, and in some application, flame retardants. The polyolefins pose fewer risks and have the highest potential for mechanical recycling. Both PE and PP are versatile and cheap. PE can be made either hard or very flexible without the use of plasticizers. PP is easy to mold and can be used in a wide range of applications. PS is widely used for foam

insulation and is equally important for hard applications, such as cups and toys. Its production involves the use of known (benzene) and suspected human carcinogenic substances (styrene and 1,3-butadiene). Styrene is also known to be toxic to the reproductive system. PS can be technically recycled, but recycling rates are low. PET is made from ethylene glycol and dimethyl terephthalate. PET is generally used in packaging (e.g., bottles) and often contains additives, such as UV stabilizers and flame retardants. PET recycling rates are high compared to those of other plastics.

Software

Two different commercial software packages were used for PCA. The first one was Minitab v. 15.1.1.0. This software can analyze 500 pixels. The second one was developed by Camo, is called The Unscrambler, version 9.1, and could analyze the 27,000 pixels of our spectra. For the hierarchical CA, the program SPSS 14.0 was used.

PCA is based on the consideration of a data matrix $\mathbf{X}$ with $N \times K$ dimensions, where N is the number of samples and K is the number of variables, expressed as a product of two matrices with their more significant characteristics: $\mathbf{X} = tp' + E$, where t is the score ($N \times q$), p is the loading ($K \times q$), and q is the number of principal components [PCs; $q \leq \min(N,K)$]. In the PCA model, the importance of a variable is indicated by its variance, and E is the residual. The variance explained by every PC is the algebraic concept of the associated eigenvalue. It is important to point out that the PCs are a new set of new variables with several important properties: (1) the information explained by a PC is not present in the other PCs; (2) the first PC explains most of the information, the second explains less, and so forth; and (3) a very reduced number of PCs can explain almost all of the information we had in the original set of atomic peaks. It derived from the latter property, that random noise and irrelevant spectral artifacts are left out for the last PCs; this is very advantageous, and the signal/noise correction is not necessary. The preprocessing of data is a recommended practice before the application of PCA because the results are based on the variance patterns of the original data. There are four preprocessing strategies: mean centering, auto scaling, range over scaling, and variance scaling. In this study, the preprocessing strategy used by the software was mean centering.

On the other hand, hierarchical CA is a class of techniques used to classify cases in groups that are relatively homogeneous and heterogeneous among them, on the basis of a defined group of variables. This analysis begins with the calculation of the dis-

tance matrix (which can be calculated in different ways: Euclidean distance, Euclidean distance squares, cosine, Pearson's correlation, Chebychev, Blocks, or Minkowski) of the cases that contain the existent distances between each element and the remaining ones. From this, the two most closed elements are looked for, and they are grouped in a conglomerate where the agglomeration methods can be the nearest neighbor, furthest neighbor, Ward's method, between-groups linkage, centroid clustering, and within-groups linkage. The next elements are grouped from this point in the same way. The result is expressed in a dendrogram, in which the groupings of the samples can be observed according to their similar characteristics. The versatility of this analysis resides in the possibility of using different types of measures and to select a great variety of methods. For this reason, it has been carried out as a comparison of all these possible cases to determine which of them delivered the best results.

RESULTS AND DISCUSSION

Correlation PCA

Minitab v. 15.1.1.0 allows an intuitive approach to PCA. It allows one to carry out PCA in two versions: by correlation and by covariance. The results obtained with this software were carried out in all of these cases by correlation.

A visual inspection of the 10-shot average was carried out on each of the samples. Thanks to this visual inspection of the spectra, two ranges of wave longitudes were selected because the spectra of the samples in these ranges had differences for the different kinds of samples; the ranges were 497.80–505.21 nm and 655.19–658.52 nm. With PCA applied to these ranges, two intuitive graphics were obtained, which are shown in Figures 1 and 2. In the first range, PC1, two groups of samples could be separated: PS + PP and PET + PE. With PC2, it was
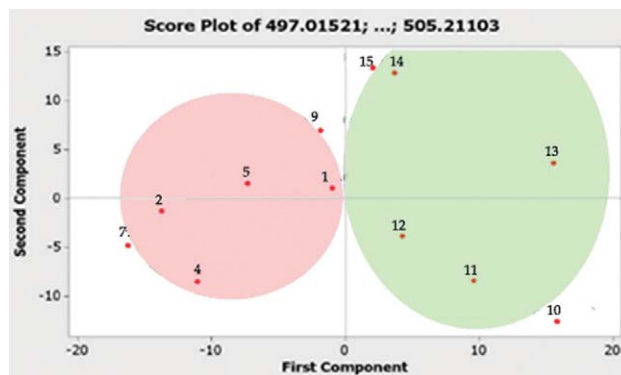

Figure 1 PCA results at 497.80–505.21 nm. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
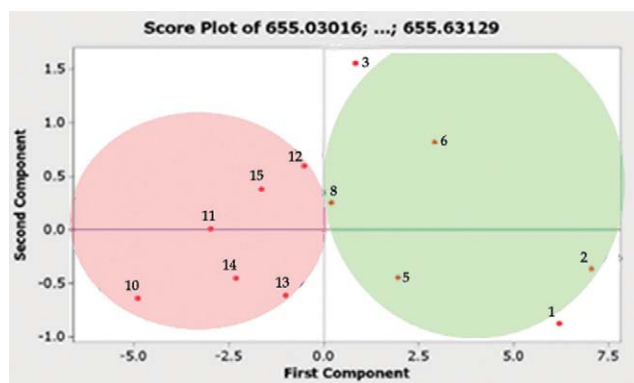
Figure 2 PCA results at 655.19–658.52 nm. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

possible to separate PP from PS. In the second one, with PC1, two groups of samples could be separated: PS + PP and PET + PE.

The independent analysis of 10 shots was investigated for each plastic type, and graphs of poor visualization were obtained, where, because of the quantity of accumulated points, it was impossible to correctly visualize the groups.

Carrying out the analysis with the two ranges of wavelengths mentioned previously, we performed a new analysis of the elementary emission lines of the plastics: H (656 nm), C1 (247 nm), O (777 nm), N (744 nm), C2 (516 nm), and CN (388 nm). With PC1, two groups of samples could be separated: PS + PP and PET + PE. Similar results were obtained for the other lines, except for H (656 nm), which only allowed the differentiation of PE from the rest of the plastics with PC1 and for CN (388 nm), where the separation was impossible.

The analysis of the interference of several wavelengths was carried out. As shown in Figure 3, with PC1, it was possible to differentiate PP + PS from PET + PE, and for PC2, it was possible to separate PET from PE. With this procedure, it was possible to separate the samples into three different groups: PP + PS, PE, and PET. Similar results were obtained for H 656, C 247, and O 777.

We carried out by PCA the study of the ratios (relations between the intensities of the characteristic elements of the plastics). In this case, just PS, PE, and PET were studied. The results obtained are shown in Figure 4. We observed that it was possible to differentiate between PS and PELD from the PET and PE with PC1 (also being hard plastics). On the other hand, with PC2, it was possible to differentiate PS from PELD and PET from PE. The graph allows the classification of the plastics. The study of the loadings was carried out to check the wavelengths, which gave the classification shown in Figure 3. PS,

PET, and PE were classified as C 516 : C 247 and O : C 247, C 247 : N, and H:O, respectively.

In conclusion, a good selection of the range of wavelengths in PCA was of vital importance for the success of the analysis. If the wavelength ranges studied were not present in the spectra of the plastics, no separation was observed by PCA. A high number of samples is necessary to be able to form the data clouds but not in such a big quantity because this does not allow the correct visualization of the graphics. In many of the studied cases, the outliers were identified previously by the test of Grubbs, which is more visual that their identification by PCA. The range of wavelengths 497.80–505.21 nm clearly separated the plastics better than the other studied range, 655.19–658.52 nm. With regard to the analysis of elementary wavelengths, the worst results were obtained for the two elected ranges for visual inspection, which were able to separate the plastics for groups more than one by one. According to the grade of precision required in the separation, this analysis was very useful. In the ratios studied, because the number of wavelengths decreased considerably, it was possible to observe the graph of loadings to determine what wavelengths exactly produced the separation of the plastics.

The numeric results of the pattern of PCA could be introduced in a classifier that used the method for the next neighboring to discover the effectiveness in the separation among the classes.

Covariance PCA

The Unscrambler, version 9.1, allowed a bigger depth to the study by PCA, mainly because it allowed the introduction of the complete spectrum (27,000 pixels) for its direct study by PCA, without the need for previous selection by visual inspection of the wavelengths. The graphs could also be obtained in three dimensions; representing the first 3
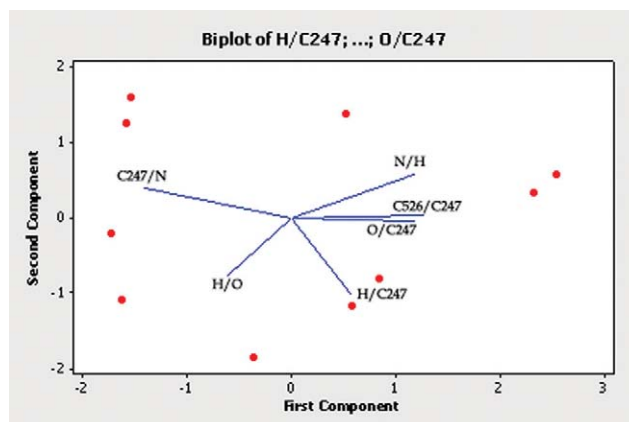

Figure 3 Loadings. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
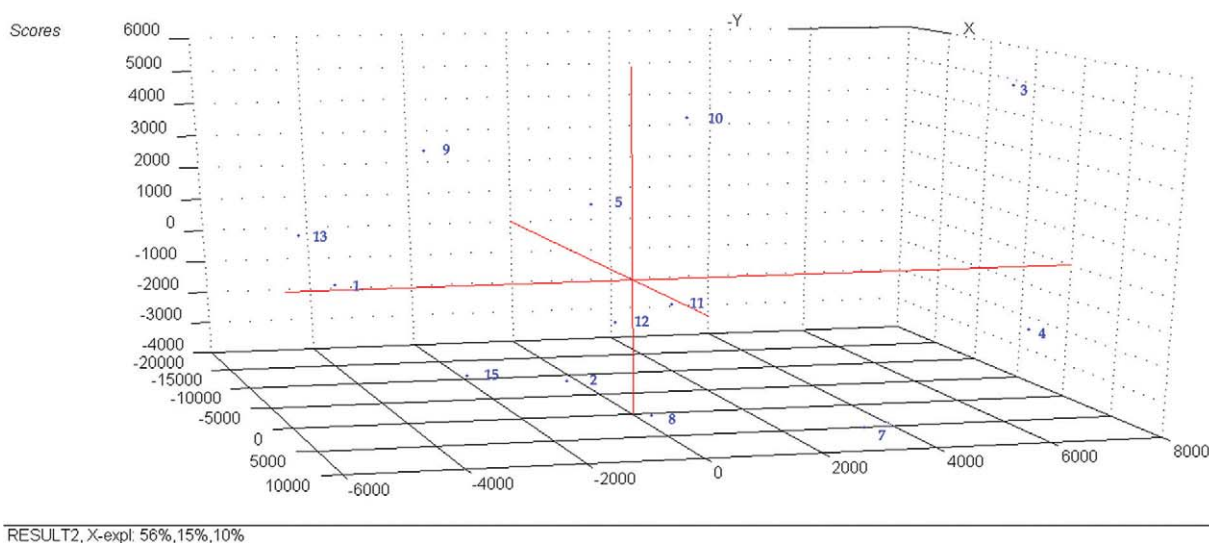
Figure 4 Scores in three dimensions. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

PCs together improved the interpretation and visualization of the data clouds. Another difference with the previous program was that was only possible to carry out PCA by the covariance method. It also allowed graphics between PC1 and PC2, PC2 and PC3, and so on.

The first thing that was carried out with this software was the complete introduction of the spectrum for all of the samples and their study by PCA. In Figure 4, the three-dimensional graphic for the three first components is shown without the outliers. With PC1, it was possible to differentiate PP + PS and PE + PET, and with PC2, it was possible to differentiate PP from PS. In this case, PC1 explained 56% of the variance, PC2 explained 15% of the variance, and PC3 explained 10% of the variance, so with the first three components, 81% of the variance was explained.

With the loadings, it was possible to distinguish the wavelength ranges over which made the graphic separations for PE + PET (389–410, 772–805, 500–520, and 644–680 nm), PS (569–590, 323–347, and 470–480 nm), and PP (415–441 and 442–450 nm).

It was interesting to carry out a new PCA with just this range of wavelengths. Like before, with PC1, it was possible to differentiate PP + PS and PE + PET, and with PC2, we could differentiate PP from PS. In this case, PC1 explained 60% of the variance, PC2 explained 13%, and PC3 explained 10% (total of variance explained: 83%).

In conclusion, this software allowed bigger versatility for the visualization of the data and the obtained results. These results were similar to the PCA with correlation, and the first three PCs explained most of the variance of the data. Furthermore, it was possible to detect the wavelength ranges that affected these PCs and to take into account the full spectrum of 27,000 pixels obtained by LIBS, and it was possible to reduce the ranges of wavelengths of the spectra to be analyzed.

Hierarchical cluster

Hierarchical clusters were carried out for all kinds of distances in the first step, which gave different matrix distances, which were used for the next steps. Once the distance was chosen, every method was checked, and the comparison of all of these combinations is shown in Table I. In this table, the clusters formed by the different plastics are shown. Some blanks are present; this was when, in the dendrogram, clusters could not be appreciated. It is also indicated, for each case, the composition, in types of plastic, for each obtained cluster.

We observed that the best separation was obtained for the Euclidean distance with the first four methods, with the Euclidean distance squares with the last four methods, and with the cosine for the first two methods.

In these cases, it was possible to differentiate two clusters, one formed by the plastics PE and PET and the other one formed by PS and PP. One of these dendrograms is shown in Figure 5. Thanks to this analysis, it was also possible to identify two outliers, samples PE6 and PS14. With the rest of distances and algorithms, a clear separation among the plastics was not possible.

CONCLUSIONS

PCA reduces the dimensionality of a group of data and offers transformation with a reduced number of variables, which, despite that, continue to describe

TABLE I
Comparison Between the Methods and Distances in Hierarchical Cluster Analysis

Method	Euclidean distance	Euclidean distance squares	Distance				
			Cosine	Pearson's correlation	Chebychev	Blocks	Minkowski
Nearest neighbor	PE + PET PS + PP	—	PE + PET PS + PP	PE + PET PS + PP + PE	PE + PET PS + PP PE + PET2[a]	—	PE + PET PS + PP2[a]
Furthest neighbor	PE + PET PS + PP	PE + PET PS + PP	PE + PET PS + PP	PE + PET PS + PP + PE	PE + PET PS + PP PE + PET	PE + PET PET + PE + PS + PP PP + PS	PE + PET PS + PP
Ward's method	PE + PET PS + PP	PE + PET PS + PP	—	—	PE + PET PS + PP PE + PET	PE + PET PS + PP PE + PET	PE + PET PS + PP
Between-groups linkage	PE + PET PS + PP	PE + PET PS + PP	PE + PET + PS PS + PP + PE	PE + PET PS + PP2[a]	PE + PET PS + PP PE + PET3[a]	—	—
Centroid clustering	—	PE + PET PS + PP	PE + PET PS + PP	PE + PET PS + PP + PE	—	PE + PET + PP4[a]	—
Within-groups linkage	—	PE + PET PS + PP	PE + PET PS + PP + PE	PE + PET PS + PP	PE + PET PS + PP PE + PET	PE + PET PS + PP	PE + PET PS + PP

[a] Number of samples without cluster.
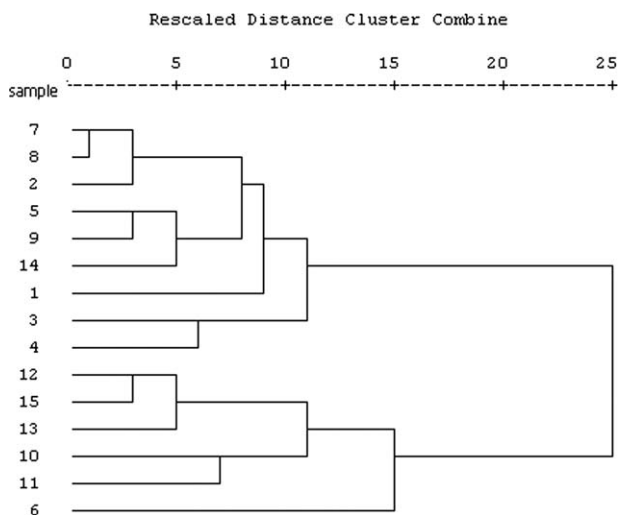
Rescaled Distance Cluster Combine

Figure 5 Dendrogram.

the data in their biggest part. PCA can be used as a classifier and detection of outliers because the representation of the first scores, vectors, or components offer an image with the classes of the separate cases. Its application to a technique that offers a high quantity of information, such as the LIBS spectra, guarantees a correct and reliable simplification that can be applied to satisfactory classification of consumer plastics.

Most of the current commercial technologies have processed capabilities and time ranges that limit their users to large-volume recyclers. The system proposed in this article would be best applied if the plastic intake volume is large enough. Hierarchical CA proved that the four polymers (PET, PE, PP, and PS) are the only ones whose recycling has an economic interest and can be divided in two clusters; this will allow their identification and classification, in a fast and easy way that can be carried out with commercial software and can be implanted online in a recycling factory.

References

1. Allen, V.; Kalivas, J. H.; Rodríguez, R. G. Appl Spectrosc 1999, 53, 672.
2. Murase, A.; Sato, N. Appl Spectrosc 1999, 53, 745.
3. Pongstabodee, S.; Kunachitpimol, N.; Damronglerd, S. Waste Manage 2008, 28, 475.
4. Hearn, G. L.; Ballard, J. R. Resour Conserv Recycling 2005, 44, 91.
5. Gaspard, S.; Oujja, M.; Rebollar, E.; Abrusci, C.; Catalina, F.; Castillejo, M. Spectrochim Acta Part B 2007, 62, 1612.
6. Gottfried, J. L.; DeLucia, F. C., Jr.; Munson, C. A.; Miziolek, A. W. Spectrochim Acta Part B 2007, 62, 1405.
7. López, A. J.; Nicolás, G.; Mateo, M. P.; Piñón, V.; Tobar, M. J.; Ramil, A. Spectrochim Acta Part B 2005, 60, 1149.
8. Carmona, N.; Oujja, M.; Gaspard, S.; García-Heras, M.; Villegas, M. A.; Castillejo, M.; Spectrochim Acta Part B 2007, 62, 94.
9. Moench, I.; Sattmamn, R.; Noll, R. Proc SPIE-Int Soc Opt Eng 1997, 3100, 64.
10. Anzano, J.; Casanova, M. E.; Bermúdez, M. E.; Lasheras, R. J. Polym Test 2006, 25, 623.
11. Anzano, J.; Lasheras, R.-J.; Bonilla, B.; Casas, J. Polym Test 2008, 27, 705.
12. Amador-Hernández, J.; Fernández-Romero, J. M.; Luque de Castro, M. D. Anal Chim Acta 2001, 435, 227.
13. Martin, M. Z.; Labbé N.; Rials, T. G.; Wullschleger, S. D. Spectrochim Acta Part B 2005, 60, 1179.
14. Jurado-López, A.; Luque de Castro, M. D. Spectrochim Acta Part B 2003, 58, 1291.
15. Jurado-López, A.; Luque de Castro, M. D. Talanta 2003, 59, 409.
16. García-Ayuso, L. E.; Amador-Hernández, J.; Fernández-Romero, J. M.; Luque de Castro, M. D. Anal Chim Acta 2002, 457, 247.
17. Laville, S.; Sabsabi, M.; Doucet, F. R. Spectrochim Acta Part B 2007, 62, 1557.
18. Martin, M. Z.; Labbé N.; André N.; Harris, R.; Ebinger, M.; Wullschleger, S. D.; Vass, A. A. Spectrochim Acta Part B 2007, 62, 1426.
19. Bousquet, B.; Sirven, J. B.; Canioni, L. Spectrochim Acta Part B 2007, 62, 1582.
20. Merdes, D. W.; Suhan, J. M.; Keay, J. M.; Hadka, D. M.; Bradley, W. R. Spectrosc Global Ed [Online], 2007, 20. http://spectroscopymag.findpharma.com.
21. Smith, L. I. A Tutorial on Principal Components Analysis; 2002.
22. http://neon.otago.ac.nz/chemlect/chem306/pca/Theory_PCA/index.html (accessed December 2007).
23. Fernández, J. L.; Sánchez, J.; Hernanz, A.; Alberdi, J.; Barcala, J. M.; Navarrete, J. J. J Automática (España) 1999.
24. Anzano, J.; Bonilla, B.; Montull-Ibor, B.; Lasheras, R. J.; Casa-González, J. J. Polym Eng 2010, 30, 177.