Non-parametric analysis of infrared spectra for recognition of glass and glass ceramic fragments in recycling plants

Alessio Farcomeni a, Silvia Serranti b,*, Giuseppe Bonifazi b

a Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università di Roma “La Sapienza”, P.le Aldo Moro, 5 00185 Roma, Italy
b Dipartimento di Ingegneria Chimica, dei Materiali, delle Materie Prime e Metallurgia, Università di Roma “La Sapienza”, Via Eudossiana, 18 00184 Roma, Italy

Accepted 24 January 2007
Available online 11 April 2007

Abstract

Glass ceramic detection in glass recycling plants represents a still unsolved problem, as glass ceramic material looks like normal glass and is usually detected only by specialized personnel. The presence of glass-like contaminants inside waste glass products, resulting from both industrial and differentiated urban waste collection, increases process production costs and reduces final product quality. In this paper an innovative approach for glass ceramic recognition, based on the non-parametric analysis of infrared spectra, is proposed and investigated. The work was specifically addressed to the spectral classification of glass and glass ceramic fragments collected in an actual recycling plant from three different production lines: flat glass, colored container-glass and white container-glass. The analyses, carried out in the near and mid-infrared (NIR–MIR) spectral field (1280–4480 nm), show that glass ceramic and glass fragments can be recognized by applying a wavelet transform, with a small classification error. Moreover, a method for selecting only a small subset of relevant wavelength ratios is suggested, allowing the conduct of a fast recognition of the two classes of materials. The results show how the proposed approach can be utilized to develop a classification engine to be integrated inside a hardware and software sorting architecture for fast "on-line" ceramic glass recognition and separation.

© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

In the glass recycling sector, the presence of glass ceramic fragments mixed with glass resulting from waste collection, can negatively affect the production process and the overall quality of the final recycled product. Glass ceramics are characterized by a melting point higher than that of glass and cannot melt during the furnace cycle (Pannhorst, 1997). As a consequence, the furnace and other glass producing machines can be damaged and the final products (bottles, jars, etc.) can break during the manufacturing or handling process or can show some defects (Fig. 1). Such a problem dramatically increased in these last years due to the introduction on the market of large quantities of glass ceramic manufactured goods, characterized by high thermal-shock resistant properties, such as dishwasher, cookware, stove tops, and cooking surfaces for electric and gas stoves (Holland and Beall, 2002). According to market demand, glass ceramics are nowadays not only opaque white colored, but also transparent, being practically undistinguishable from common glass, both for human senses and for the sorting devices usually utilized in recycling plants.

It is well known that glass collected from both domestic and commercial waste is accompanied by several polluting materials, such as plastic, metal, paper and wood, as well as the previously mentioned glass ceramic or glass-like contaminants. While good strategies have been developed in the last years for removal of several contaminants using automatic on-line sorting systems, no really effective or low cost solution has been found until now for the recognition of glass ceramic fragments inside a glass waste stream for recycling. Glass recyclers are strongly interested in solving such a problem, being one of the major causes of economic loss.

* Corresponding author. Tel.: +39 06 44585925; fax: +39 06 44585618.
E-mail address: silvia.serranti@uniroma1.it (S. Serranti).

0956-053X/S - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wasman.2007.01.019
The only two currently pursued strategies, at the industrial scale, to reduce the presence of glass ceramic contaminants are “reduction at source” and “manual sorting”, the approaches both gave scarce results. The “reduction at source” is strongly conditioned by the fact that citizens, in spite of public education programmes, confuse glass ceramic with glass. “Manual sorting” is usually based on the utilization of trained operators that, looking carefully at the waste stream, try to identify glass ceramic fragments. Such an approach is clearly expensive and not reliable. One of the parameters usually adopted in human-based recognition is related to the reflectance characteristics of fragments; such a property can be estimated in different ways according to human expert knowledge, his level of attention and environmental conditions.

The possibility of recognizing glass and glass ceramic fragments by spectroscopic techniques, both in the visible and infrared field, has been preliminary investigated previously (Bonifazi and Serranti, 2006; Serranti et al., 2006). In those studies the attention was addressed to the evaluation of the spectral response of glass ceramic fragments in respect to their surface characteristics, color and polluting elements and to the comparison of spectral signature of glass fragments (cullet) characterized by a similar set of attributes. Results showed that a hyperspectral approach in the visible and near infrared (VIS–NIR) field could allow recognition of the different products especially in the near infrared field (Bonifazi and Serranti, 2006). According to these results, systematic investigations have been thus carried out in the mid-infrared field (MIR) also. The results demonstrated a higher sensitivity of the approach in this spectral range, thus allowing achievement of a better distinction between glass and glass ceramic pieces (Serranti et al., 2006). In both cases the classification was realized by selecting specific wavelengths and adopting a “band ratio approach” to perform the requested identification.

2. Materials and methods

2.1. Glass and glass ceramic samples

Glass and glass ceramic samples have been collected inside the Bernhard Reiling Glas GmbH recycling plant, located in Marienfeld (Germany).

In this study, the attention has been focused on the application of statistical classifiers for recognizing the near mid-infrared spectral signatures of the two materials, in order to achieve the best percentage of recognition. The use of non-parametric analysis of infrared spectra can be in fact adopted in the development of an integrated hardware and software architecture for “on-line” recognition and sorting of glass ceramic contaminants.

2. Materials and methods

2.1. Glass and glass ceramic samples

Glass and glass ceramic samples have been collected inside the Bernhard Reiling Glas GmbH recycling plant, located in Marienfeld (Germany).
The cullet product object of the investigations is constituted by three different typologies, according to the source glass fragment streams usually processed in the plant, that is:

- flat-glass (FG),
- colored container-glass (CG), and
- white container-glass (WG).

Two color classes of glass ceramic have been collected:

- white glass ceramic (GC-A and RC-A), and
- amber glass ceramic (GC-B).

Concerning glass samples, five different sampling points (indicated by letters A–E) have been identified in the production cycles of both colored and white container-glass, whereas flat-glass have been collected at just one sampling point. Fig. 2 illustrates the flow-sheet of the plant and the location of the sampling points.

Ten representative glass fragments for each typology (flat glass, colored container-glass and white container-glass) and related sampling point in the plant have been selected for a total of 110 samples. Sample selection was performed taking care to collect cullet characterized by different physical attributes, i.e., thickness, color, size, shape and manufacturing.

Glass ceramic samples have been collected inside the glass fragments flow stream fed to the plant. Forty representative glass ceramic white samples and 11 representative amber samples showing different thicknesses, sizes, shapes and manufacturing have been selected.

The reference sample set utilized for the analysis was thus constituted by 161 fragments. A summary of selected cullet with the description of their characteristics (cullet type, color, thickness and size) is reported in Table 1. In Fig. 3 some pictures of selected cullets for different sample types are shown.

### 2.2. FT-IR equipment and measurement

Spectra of glass ceramic and glass samples have been recorded in the NIR–MIR infrared field by a Perkin SPEC-TRUM-ONE™ FT-IR spectrometer. The optical component of the instrument consisted of a dynamically aligned Michelson interferometer, a potassium bromide beam splitter and a deuterated triglycine sulfate (DTGS) detector. The instrument is interfaced to Perkin Elmer SPEC-TRUM-ONE™ software.

Fourier transform infrared spectra were collected from 7800 to 380 cm\(^{-1}\) at intervals of 4 cm\(^{-1}\). All of the spectral analyses, performed on the different series of glass and glass ceramic samples, were carried out in transmittance conditions. The spectrum of each sample was acquired three times in order to validate the results. The smallest wavenumber range, 7800–2230 cm\(^{-1}\), corresponding to the wavelength range 1280–4480 nm, was selected based on preliminary tests of samples that showed the best signal-to-noise ratio in such interval of the near and mid-infrared field.

### 2.3. Statistical approach

In order to develop a recognition procedure for glass and glass ceramic fragments, based on their spectral signature, a statistical classifier was built.

The statistical procedure can be summarized in the following steps:

- spectral data pre-processing,
- feature selection from spectra, and
- classification of the spectra.

Spectral data were first checked for coherence and compliance to the assumptions needed for the statistical methods. A Levene test for the homogeneity of the variances of the observed band intensities across the different wavelengths was done. Such homogeneity is needed for the
The application of statistical methods, as the Student’s $T$ test and the wavelet transform.

The feature selection from glass and glass ceramic spectra has been carried out performing a Student’s $T$ test at each wavelength and then applying the GAUGE multiple testing correction method (Farcomeni, 2004). A wavelet transform (Ogden, 1997) of the data has been also applied, because wavelets are well known to be appropriate for the representation of functions that use to be spiky and/or spatially inhomogeneous, which spectra usually are.

$T$ test. It is used to compare two data groups with respect to a quantitative variable. The null hypothesis is that the two expected values are equal, in this case against a two-sided hypothesis. Under the hypothesis of equal variances, the test statistic is obtained by multiplying the difference of the means by the square root of the number of observations, and dividing the result by the (pooled) standard deviation. The test statistic asymptotically follows a Student’s $T$ distribution with degrees of freedom equal to the number of observations minus 1.

Wavelet transform. A (discrete) wavelet transform is a decomposition of a function with respect to a wavelet basis. Among the special characteristics of a wavelet basis, there is localization in both space and time domain, which allows for a compressed, whitened and often sparse representation of the signal. The decomposition is based on an iterative dichotomy of the frequency band.

Concerning the classification stage of glass and glass ceramic spectra, the $k$-nearest-neighbor ($k$-nn) method (Cover and Hart, 1967) has been adopted as it is considered one of the fastest known classification methods, and because it also gave better performance (i.e., lower classification error) when compared to a few other methods (like linear and quadratic discriminant analysis). Such an approach is a simple classification method that directly exploits the training set to build the classifier: a new object is predicted to belong to the class of the majority of the closest $k$ objects, in terms of a suitably defined distance function.

The classification was performed adopting the procedure described in the following. The available sample set was randomly split in two groups. The first group (training set) was used to do the statistical learning, that is, to estimate the parameters of the model if any or to do the actual classification. In this study the training group to classify new fragments was utilized: a distance was computed between the new fragment and all of the fragments in the training group, and then the closest fragments have been used to classify the new (unlabeled) one. The unlabeled fragment was then assigned to the group to which the majority of the ($k$) closest fragments belong to (that is, glass or glass ceramic). The second group (test set) was utilized to validate the procedure: each fragment in the test set was classified and the proportion of erroneously classified fragments was recorded. Such proportion is a good estimate of the probability of wrong classification for future, truly unknown, fragments. Such an operation was repeated several times in order to better estimate the quality of the classification. Finally, all of the known fragments were used to classify truly unlabeled fragments.

The classification error has been estimated by means of cross-validation. A test set of objects was randomly selected amongst the entire population. The remaining objects were then used to build a classifier and to predict
the class of the test set, obtaining an estimation of the classification error (Hastie et al., 2001). This operation was iterated and the estimates averaged out.

3. Results and discussion

3.1. Spectral features of glass and glass ceramic fragments

The recorded near and mid-infrared absorbance spectra of glass and glass ceramic samples are reported in Figs. 4 and 5, respectively. In Fig. 4a the complete spectra of glass fragments are shown, while in Fig. 4b the average spectral plot is given. A small spike is detected around 1500 nm, while a strong jump is visible around 2700 nm. After that jump, the absorbance remains almost constant until 4000 nm and then starts to increase. Fig. 5a shows the complete spectra of glass ceramic samples, whereas Fig. 5b gives the corresponding average spectral plot. A small spike is detected around 2000 nm, while a stronger one is visible around 3000 nm. Moreover, differently from glass, absorbance of glass ceramic samples rapidly decreases after the second spike. Another peak is then detectable moving towards 4500 nm.

3.2. Cullet spectral classification

3.2.1. Spectral data pre-processing

The Levene test for the homogeneity of the variances of the observed band intensities across the different wavelengths gave a $p$-value of 0.99. Hence the hypothesis that
the variance of the absorbances is homogeneous at different wavelengths is retained.

3.2.2. Feature selection from glass and glass ceramic spectra

A suitable number of wavelengths allowing performance of a better distinction between glass and glass ceramic was primarily selected. The purpose of wavelength selection was twofold: first of all, on-line classification is a time critical task, hence as little information as possible is to be processed. Secondly, classification methods are heavily dependent on the number of predictors, which should be minimal to avoid overfitting, curse of dimensionality, lack of robustness and instability.

Predictors have been selected as follows: (i) a Student’s T test at each wavelength was carried out to detect significant differences between absorbance of glass and glass ceramic at that specific wavelength, and (ii) the GAUGE multiple testing correction method was used to choose the significant wavelengths. As a result, only the absorbances at wavelengths corresponding to the 50 smallest p-values were considered significantly different between glass and glass ceramic.

The significance of the difference between the absorbance at a given wavelength implies that the two groups are well separated at that point. Nevertheless, the distance between the groups can be not satisfactory for classification purposes. For this reason, the previously identified 50 significant wavelengths were sorted by squared difference between average absorbance for glass and average absorbance for glass ceramic. The 10 significant wavelengths (in nm) performing the best separation between glass and glass ceramic fragments have been thus identified; they are: 2882, 2932, 2982, 3032, 3082, 3132, 3432, 3482, 3532, 3582, 3632, 3682, 3732 and 3782. Such a result is in good agreement with what is visually detectable by looking at the spectral plots. In fact, the strongest differences between glass and glass ceramic spectra are detectable in the wavelength ranges around 3000 nm and higher than 3500 nm.

In order to define the best classification strategy to be used by a sorting system working on-line in an actual plant, both in terms of quality of classification and speed of data acquisition and processing, two different approaches have been investigated, identified in the following as “fast” and “optimal” classification methods.

The “fast” classification method is an approach addressed to minimize the number of features to perform the recognition. It is based on the selection of a very small number of wavelengths to be passed directly to the classifier. Only the first 10 previously identified wavelengths have been used.

The “optimal” classification method is based on the utilization of features selection criteria addressed to produce the best in terms of generalization error of the classifier among a number of different subsets of transformed and not transformed wavelengths. It is well known, in fact, as the generalization error can be defined the expected proportion of mis-classifications due to the fact that the training has been done only on a subset of population (Hastie et al., 2001). The “optimal” classification method was given by feeding the classifier with a wavelet transform of the spectra sampled at the first 32 significant wavelengths, of the previously mentioned 50, according to the values assumed by the squared difference of the average of the absorbances for each class of objects. Such wavelengths are (in nm): 1282, 1482, 1532, 1582, 1632, 2582, 2632, 2682, 2832, 2882, 2932, 2982, 3032, 3082, 3032, 3132, 3432, 3482, 3532, 3582, 3632, 3682, 3732, 3782, 3832, 3882, 3932, 3982, 4032, 4082, 4382, 4432 and 4482.

The analysis has been restricted to 32 of the 50 significant variables, both to get a faster on-line classifier and to avoid problems associated with wavelet transforms of functions sampled in a number of points that is not a power of 2.

3.2.3. k-Nearest neighbor classification of glass and glass ceramic spectra

For spectra classification, the sum of the squared differences between the absorbances for the “fast” classification method and the sum of the squared differences between the wavelet coefficients for the “optimal” classification method have been used.

Many values of k have been tried, showing that the minimal error rate was given by k = 3 in both cases. This is not surprising, since k = 3 is a good choice in many applications.

In order to perform the classification based on the “fast” classification method, a specific procedure was developed. It was built according to three different steps:

- sampling of the absorbance spectrum at the 10 significant wavelengths listed above;
- computing of the distances between the new spectrum at those wavelengths and the training spectra; and
- assignment of the new spectrum to the class of the majority of the closest three spectra.

As for the previous approach, an algorithmic procedure was set-up also for the “optimal” classification method. In this case four steps have been identified:

- selection of the 32 significant wavelengths of the absorbance spectrum;
- data expansion with respect to a wavelet basis (wavelet transform). The adopted approach was developed on the base of Daubechies orthonormal compactly supported wavelet of length 8, least asymmetric family (Daubechies, 1992);
- distances computing between the wavelet coefficients of the new spectrum and the coefficients of the training spectra; and
- assignment of the new spectrum to the class of the majority of the closest three spectra.
The error on the training set for the “fast” classification method was 9.3%, giving mis-classification of 15 over 161 objects, 12 of which are glass ceramic fragments.

The error on the training set for the “optimal” classification method was 6.8%, giving mis-classification of 11 over 161 objects, all of which are glass ceramic fragments.

The classification error has been further estimated by means of cross-validation for both of the proposed classification procedures.

A test set of 25 objects was randomly selected from 161. The 136 remaining objects were utilized to build a classifier and predict the class of the other 25, obtaining an estimation of the classification error. This operation has been iterated 1000 times and the estimates have been averaged. Following this approach, the classification error of the “fast” classification method was estimated as 9.7% and the classification error of the “optimal” classification method as 7.6%. The estimated mis-classification rates for glass and glass ceramic are reported in Table 2. From the analysis of Table 2 it appears that, by adopting the “fast” classification strategy, just 1.7% of glass fragments are not correctly recognized, being attributed to glass ceramic. From a practical point of view, the implementation of such a procedure, inside a separator working on-line in a recycling plant, will produce, as a result, that the sorting system design would blow out less than 2 glass fragments out of 100. Concerning the glass ceramic, just 8% of fragments will be mis-classified as glass and would not be blown out by an on-line sorting system but would be processed with glass cullet. Adopting the “optimal” classification strategy, results are obviously better for both materials, in fact just 0.3% of glass will be mis-classified as glass ceramic, whereas just 7.3% of ceramic glass will be mis-classified as glass. In both cases, Table 2 shows that it is easier to correctly classify glass than glass ceramic. This is partially due to the fact that only about 30% (52 out of 161) of the objects in the training set are glass ceramic pieces. Such a choice was adopted because the presence of glass ceramic inside a cullet flow stream is usually low. The adoption of a training set constituted by few elements of glass ceramic contaminants allows the researchers to better explore the quality of the recognition/classification algorithms and to stress the related numerical procedures. Results of both classification strategies can be considered in good agreement with industrial practice and end user needs, both in terms of accuracy and processing speed. The described algorithmic procedures are, in fact, applied on a reduced set of data. In terms of costs, the adoption of such a strategy can result in a reduction of about two-thirds of the costs for the sensors and in an increase in the production (number of fragments investigated per unit of time) of about one-third. These values are, in any case, highly dependent on the plant layout and the quantity and quality, in terms of possible variations, of the feed product to recycle.

The adoption of the “optimal” classification method, on the other hand, will allow better classification results. Such a goal is reached through the acquisition and handling of a larger set of data (32 wavelengths), practically covering the entire investigated wavelength field. This approach results in a slower processing speed of about one-third if compared with the “fast” one. Estimated costs can be determined accordingly.

4. Conclusions

Different recognition/classification procedures have been developed and applied with reference to glass ceramic fragments detection inside a “cullet” (glass fragments to recycle) flow stream. In particular, the spectral signatures of glass and glass ceramic fragments have been investigated in the NIR–MIR field (1280–4480 nm). The detected spectra have been analyzed, adopting a statistical approach. More specifically, two different strategies have been developed and evaluated, named the “fast” classification method and the “optimal” classification method. The “fast” classification method led to a cross-validated estimated classification error of 9.7%. Applying the “optimal” classification method, the cross-validated estimated classification error, on processed data, decreased to 7.6%.

It is important to outline that both of the proposed procedures result in good classification results, for all the typologies of investigated glass and glass ceramic samples. This means that the results are independent from color, shape, manufacturing, size, dirtiness (even paper labels), usually manually collected glass ceramic fragments and loss of good glass related to the adoption of an automatic sorting system.

Probably the best method to be adopted in a real system could be the “fast” one. In fact, even if it resulted in a relatively larger error in classification, compared to the “optimal” one, it is important to note that the 10 selected wavelengths of the “fast” classification method can be grouped into two separate spectral windows of 200 nm each (2800–3000 nm and 3500–3700 nm). This could allow the adoption, in a real system, of a detection-sorting system design based on two separate IR-sensors, working in the two different identified wavelength ranges. Such a solution can allow the operator to minimize costs, increasing at the same time the speed of spectral data acquisition and processing. The possibility to utilize two specialized sensors, operating in the previously defined spectral ranges, allows simplification of the detection equipment design, increasing the data processing. The described algorithmic procedures are, in fact, applied on a reduced set of data. In terms of costs, the adoption of such a strategy can result in a reduction of about two-thirds of the costs for the sensors and in an increase in the production (number of fragments investigated per unit of time) of about one-third. These values are, in any case, highly dependent on the plant layout and the quantity and quality, in terms of possible variations, of the feed product to recycle.

The adoption of the “optimal” classification method, on the other hand, will allow better classification results. Such a goal is reached through the acquisition and handling of a larger set of data (32 wavelengths), practically covering the entire investigated wavelength field. This approach results in a slower processing speed of about one-third if compared with the “fast” one. Estimated costs can be determined accordingly.

Different recognition/classification procedures have been developed and applied with reference to glass ceramic fragments detection inside a “cullet” (glass fragments to recycle) flow stream. In particular, the spectral signatures of glass and glass ceramic fragments have been investigated in the NIR–MIR field (1280–4480 nm). The detected spectra have been analyzed, adopting a statistical approach. More specifically, two different strategies have been developed and evaluated, named the “fast” classification method and the “optimal” classification method. The “fast” classification method led to a cross-validated estimated classification error of 9.7%. Applying the “optimal” classification method, the cross-validated estimated classification error, on processed data, decreased to 7.6%.

It is important to outline that both of the proposed procedures result in good classification results, for all the typologies of investigated glass and glass ceramic samples. This means that the results are independent from color, shape, manufacturing, size, dirtiness (even paper labels),
and that the proposed strategies could be successfully applied inside glass recycling plants for the different glass production lines. In order to define the best sorting strategy, both the quality of classification results and the speed of spectral data acquisition and processing should be taken into account.

The selection of the best classification procedure should be evaluated during an on-line testing phase carried out directly in a glass recycling plant. In fact, the speed of spectral data acquisition and processing of both of the proposed classification methods, “fast” and “optimal”, should be tested considering the speed of the glass stream. In other words, the proper compromise between correctness of classification and processing time should be found, according to user needs and system requirements.

Acknowledgements

The work was financially supported within the framework of the Project CRAFT-1999-71817 (Contract No.: G1ST-CT-2002-50298) “Development of a Novel and High-Speed Spectral Imaging System to Detect Glass-Like Contaminants in the Recyclable, Cost Effectively, Increasing Glass Recycling and Avoiding Landfilling” of the European Union. A special thank you is given to Mr. Rohland Pohl (Reiling Glass Recycling GmbH, Marienfeld, Germany) for samples collection and for his valuable information concerning problems related with glass ceramic recognition and handling inside the glass recycling plant.

References


