Application of an electronic nose to the discrimination of coffees

J. W. Gardner, H. V. Shurmer and T. T. Tan
Department of Engineering, University of Warwick, Coventry CV4 7AL (UK)

Abstract

An investigation has been carried out into the response of an array of twelve tin oxide sensors to the headspace of coffee packs. Discriminant and classification function analyses are performed on the array response to each of three commercial coffees (covering two different blends and two roasts) as well as one coffee which has been subjected to a range of six roasting times. Multivariate functions are calculated from the entire data set (90 samples) or alternatively using half of it, to permit cross-validation. A success rate of 89.9% is achieved with the former procedure in classifying the three commercial coffee odours directly from the response (change in sensor conductances) of the array. This value falls to 81.1% when half of the data set is used for cross-validation. Preprocessing the array data, by normalizing the response of each sensor over the array, is found to increase the success rate (to 95.5%) on the entire data set only. The effect on coffee odour of a set of six roasting times (zero to 11.5 min) is also investigated and found to be considerable, some sensors registering an increase in conductance by a factor of three. A 100% group classification is achieved with zero and long roasting times, the overall success rate being 88.1%. The main conclusion is that tin oxide gas sensors can be used to discriminate between both the blend and roasting level of coffee, confirming their potential application in an electronic instrument for on-line quantitative process control in the food industry.

1. Introduction

A prime requirement of the food industry is a sensitive method of assessing volatiles, for identification, authentication, process control and product blending or formulation. The aim of the present work was to demonstrate the feasibility of using an electronic nose for a range of food applications [1–3]. Coffee quality is assessed by expert coffee tasters largely on the basis of its aroma and flavour, and the highest-quality beans command a considerable premium. Coffee volatiles are numerous and varied in their aroma quality, potency and concentration. Most of the volatiles are derived from initially non-volatile components of the raw bean, which break down and react during roasting, forming a complex mixture. Pyrolysis, other reactions and interactions of components, such as sugars, amino acids, organic acids and phenolic compounds, result in the formation of the characteristic aroma and flavour of coffee. The final composition of volatiles depends on a number of factors, including species/variety of bean, climatic and soil conditions during growth and storage, after both harvesting and roasting, time and temperature of roasting, as well as the roasting equipment used. Green coffee beans are generally regarded as having no agreeable aroma or flavour, but they do possess a large number of volatiles, most of which increase in concentration during coffee roasting, although some tend to decrease through degradation. Green and roasted beans contain an appreciable amount of aliphatic hydrocarbons, probably derived from the oxidation of green bean lipids during storage or transport prior to roasting. Green beans, on average, contain about 13% lipid material, three-quarters of this being triglyceride. There are also appreciable amounts of diterpene, triterpene and sterol esters, free diterpenes, triterpenes, sterols and phosphatides. Roasting has little effect on the volatiles produced in this way and causes negligible change in their overall concentration. In green coffee beans, it has been reported that 21 hydrocarbons, 10 carbonyls, 10 esters, five sulphur compounds, two alcohols, four heterocycles and methylindoline may be detected, including furans, thiophenes, sulphides, aldehydes and ketones [4].

Roasting temperature, time, method of roasting and cooling all affect the volatile composition. The degree of roast affects the formation and degradation of the different volatiles and the final coffee quality. A study of the effect of roasting time on the concentration of volatiles has shown that some
reach the peak of their concentration during a commercial roast, whilst others do not. Some volatiles decrease significantly when roasting is prolonged, whereas others increase, including phenols, pyrroles and furfuryl alcohol. Aldehydes, proponone and total phenols also increase throughout roasting, slowly at first, then at a greater rate, without reaching a peak. A more comprehensive report on coffee volatiles, their reactions and their aroma is outside the scope of this paper, but the details are given by Clarke and Macræ [4].

2. Pattern-recognition techniques

The procedure employed to analyse the data was based upon DFA (discriminant function analysis) and MANOVA (multivariate analysis of variance) techniques. The experimental data were initially screened to check for missing values, outliers and proximity to normal probability density functions. Several transformations were assessed to ensure that any analytical requirements were met [5] as well as to optimize the output from the sensor array [6].

The main aim of discriminant function analysis is to predict group membership from a set of dependent variables or predictors [5–10]. DFA is concerned with the problem of separating out different groups on the basis of available information, whereas MANOVA determines whether group membership produces reliable differences on a combination of predictors. If it does, then the combination of variables can be used in DFA to predict group membership. On the other hand, the main aim of MANOVA is to determine whether group membership is associated with reliable differences in combined dependent variable scores, whereas in DFA it is to determine whether predictors can be combined to predict group membership reliably.

The general arrangement of the pattern-recognition system is shown in Fig. 1. The output from the n-dimensional array of semiconducting sensors is fed into a preprocessor that defines the optimal response parameter, given the nature of the sensors and pattern-recognition technique employed. Features that characterize the odour patterns may be identified from supervised pattern recognition, stored in a data base and then extracted for comparison with new data to predict or classify unknown outputs in terms of a coffee type or flavour attribute from organoleptic tests.

3. Experimental studies

Three commercial types of coffee (blend) were supplied by Lyons Tetley Ltd. for testing. These coffee types were as follows, with the letters (a) to (i) used as labels: (a) a medium roasted coffee of blend type 1; (b) a dark roasted coffee of blend type 1; (c) a dark roasted coffee of blend type 2. Additional coffee types (roasts) were prepared by the Campden Food and Drinks Research Association (CFDRA) from a single coffee blend with a series of bean roasting times. These were unroasted coffee (d); and coffee roasted for 6 min (e); 7.5 min (f); 9.5 min (g); 10.5 min (h); and 11.5 min (i). An array of 12 different commercial tin oxide gas sensors with partially overlapping sensitivities was used to evaluate the coffee headspace [3].

4. Data analysis

4.1. MANOVA

The data were initially screened and found to be reasonably well behaved, i.e., few outliers and approximately normal probability distribution functions. The analysis of the data was carried out using a commercial statistical software package, SPSS [9]. MANOVA was used to analyse the data, confirming that group membership produced reliable differences on a combination of predictors, with the sensor outputs S1 to S12 as the dependent variables and the groups of coffee (a)–(c) the independent variable (one independent variable with three levels).
4.2. Discriminant function analysis

Discriminant function analysis was carried out on the sensor array response obtained for the three commercial coffees (30 samples of coffee (a), 30 samples of coffee (b) and 30 samples of coffee (c)) and the set of roasted coffees (7 samples of coffee at each roasting time, (d)–(i)).

The raw array output (change in conductance) and a preprocessed signal, which consisted of the ratio of change in sensor conductance to the sum of squares over the array, was analysed for the three commercial coffees (a)–(c). The latter normalization procedure has been reported as the optimal response parameter with which to characterize sensor performance [6]. The data were used in the following manner: (i) all the samples in each group were used to derive the classification coefficients and then used to classify themselves (supervised learning); (ii) half the samples in each group were used to derive the classification function and half to cross-validate the classification.

Experimental results are likely to contain some degree of sensor drift and sample variation within the data sets. For minimization of errors and, in particular, guarding against the baseline drift, two steps were implemented:

1) Samples from different coffee types were measured alternately, e.g., eight samples of coffee type (a) and then eight samples of coffee type (b), etc. This reduced the chance of falsely discriminating between groups due to baseline drift.

2) For the cross-validation analysis, the samples were randomly chosen, 15 being used for classification and 15 for cross-validation. The roles of the classifying samples and the cross-validation samples were then reversed and the average of these two analyses was taken as the final result.

The first method of analysis was then performed on the raw data. For the commercial coffees, the grouped data for 30 samples were used to derive the classification coefficients in the analysis, where each coffee had a distinct classification function. Data for each sample were inserted into each classification equation to develop a classification score for each coffee. Each sample was assigned to the group for which it had the highest classification score, a process sometimes called jack-knife classification. The results obtained with all 90 samples included in the analysis gave an overall chance of 89.9% success in classifying the samples, coffee type (c) having the lowest success rate for correct classification. The first two discriminant functions were calculated and the classified scores plotted in Fig. 2. When comparing the group centroids (c), the first discriminant function, $Z_1$, best separates out the coffee types labelled (a) and (b) while the second discriminant function, $Z_2$, best separates out coffee types (a) and (c). Some dispersion in the individual coffee type (c) scores caused an overlap with group (a) and thus a lower degree of discrimination. In this supervised analysis, the success rate for coffee classification improved from a value of 89.9% to 95.5% by preprocessing the data via the suggested normalization procedure. These results agree with those from a previous study on the choice of the preprocessing expression for tin oxide gas sensors [6]. The normalization procedure helps reduce the effect of sample variation.

In the second method of analysis, the classification equation was based on only 15 out of 30 samples per coffee. The 15 samples of each coffee used to derive the classification equation gave a slight decrease in overall success in classifying coffee groups (see Fig. 3 for a plot of the individual scores and group centroids). The success rate is now lower at 88.9% for classifying samples from which the equation was derived. This gives success rates of 93.3%, 93.3% and 80.0% for coffee groups.
(a), (b) and (c), respectively, for correctly classifying the group of samples. When the remaining 15 samples of each group were used for cross-validation, the success rates became 86.7%, 80.0% and 73.7% for groups (a), (b) and (c), respectively. When these remaining samples were used as the calibration samples, they were themselves classified with success rates of 100%, 93.3% and 86.7%, respectively. Cross-validation of these results gave corresponding figures of 86.7%, 100% and 73.3%. The average success rates obtained from all these results were 88.9% for self-classification and 81.1% using cross-validation.

The coffee types labelled (d) to (i), representing Reference 4 discusses the measured changes in different roasting times, were prepared by CD-FRA and are also analysed. However, the data set could not be fully screened due to there being only seven samples in each group. The results obtained are nevertheless worth noting, as the percentage of grouped cases correctly classified was 88.1%. This compares very favourably with classification by chance, which is only about 16.7%. The group centroids, as shown in Fig. 4, tend to move from left to right with increasing roasting time. Coffee type (d) is for zero roast, which is well away from the other coffee types (100% classified correctly). Coffee types (c) and (f) are closer to each other than to any other coffee (85.7% success and 71.4% success for groups (c) and (f), respectively). Misclassification was only found to occur for either group (e) or (f). Similarly, coffee types (g) and (h) are closer to each other than to any other coffee, 100% and 71.4% being correctly classified for coffee types (g) and (h), respectively, with misclassification only between the two mentioned groups. Coffee type (i) is well separated from any other coffees and was classified correctly for all samples.

5. Results and discussion

Reference 4 discusses the measured changes in the concentrations of some key volatile components of coffee over the range of roasting times used in our experiments, indicating, in particular, a rise in the level of furfuryl alcohol and phenol with roasting time. As tin oxide gas sensors are particularly sensitive to combustible materials, this should lead to a change in the array output. Results for six samples of each of the roasting times (d) to (i) showing the responses (fractional changes in conductance) for all twelve sensors are indicated in Fig. 5. This confirms the expected conductivity increase with roasting time for sensors 1 to 11, but sensor 12 appears to be uninfluenced. Scrutiny of these results shows that the sensitivity to blend is weaker than it is to roast.
6. Conclusions

The feasibility of using an electronic nose to classify coffee aromas has been demonstrated in this investigation. The success in classifying three commercial coffees was greater than 80.0% and the success in classifying samples prepared by CD-FRA was greater than 88.1%. Improvement in the experimental procedures, equipment and design can be expected to enhance further the discriminating power of the electronic nose for coffees. A mass-flow system that automatically acquires the data with much higher reproducibility and consistency (instead of a static rig using a syringe injection of headspace) should improve the quality of the data acquired. Reducing the error in the data itself then could further improve the success rate of classifying coffee blends and roasts and, in addition, there is scope for further advancement of sensor technology, particularly in the selectivity and stability of the sensors. The development of integrated tin oxide sensors and conducting polymer devices are priority items at Warwick and are expected to facilitate improvements in resolution [11, 12].

Acknowledgements

This study was carried out in collaboration with the Campden Food and Drinks Association as part of a programme commissioned by the Ministry of Agriculture Fisheries and Food. We also acknowledge the contributions of our colleagues, Philip Bartlett and George Dodd, to the Warwick Electronic Nose project.

References