Journal of Thermal Analysis and Calorimetry, Vol. 71 (2003) 147–154

RADIAL BASIS NEURAL NETWORK FOR THE CLASSIFICATION OF FRESH EDIBLE OILS USING AN ELECTRONIC NOSE

Z. Ali[1*], *D. James*[1], *W. T. O'Hare*[1], *F. J. Rowell*[2] *and S. M. Scott*[1]

[1]School of Science and Technology, University of Teesside, Middlesbrough, TS1 3BA, UK
[2]School of Health, Natural and Social Sciences , University of Sunderland, Sunderland, Tyne and Wear, SR2 7EE, UK

Abstract

An electronic nose utilising an array of six-bulk acoustic wave polymer coated Piezoelectric Quartz (PZQ) sensors has been developed. The nose was presented with 346 samples of fresh edible oil headspace volatiles, generated at 45°C. Extra virgin olive (EVO), Non-virgin olive oil (OI) and Sunflower oil (SFO), were used over a period of 30 days. The sensor responses were then analysed producing an architecture for the Radial Basis Function Artificial Neural Network (RBF). It was found that the RBF results were excellent, giving classifications of above 99% for the vegetable oil test samples.

Keywords: edible oils, electronic nose, neural network, piezoelectric quartz, radial basis function

Introduction

The quality control of odours and volatile compounds is important in many sectors ranging from the chemical to the environmental [1]. With respect to fresh edible oils, the investigation of food volatiles is increasingly of interest. The most widely applied and established technique for their evaluation being Gas-Chromatography coupled to Mass-Spectrometry (GC-MS) [2]. However, this equipment is expensive and requires technical skill for satisfactory operation.

With this in mind, significant work has been carried out on inexpensive portable instruments for online factory floor measurements. Much activity has taken place in using arrays of chemical or bio-sensors with limited selectivity to the analyte being measured. These not only allow good reversibility for array re-use, but through analysis of the pattern of sensor responses, selectivity can also be gained. The sensing array, known as an 'electronic nose' if the volatile being sensed has an odour [3], can be based upon various different sensing elements. Metal Oxide Sensors (MOS) [4], Metal Oxide Semi-Conductor Field Effect Transistor (MOSFET) [5], Conducting Polymer Sensors [6], Piezoelectric Quartz Crystal (PZQ) (also known as Bulk Acous-

* Author for correspondence: E-mail: z.ali@tees.ac.uk

1418–2874/2003/ $ 20.00
Akadémiai Kiadó, Budapest

tic Wave Sensors (BAW)) [7] and Surface Acoustic Wave Sensors (SAW) [8] have all been used as electronic nose sensors.

Electronic nose

The system consists of an array of partially selective piezoelectric quartz crystals (PZQ) sensors; the signals from these sensors may be processed using pattern recognition techniques to ascertain the analyte under investigation. When excited by an alternating current these devices resonate at a designed frequency. In PZQ sensors a chemical or biochemical layer on the surface of the crystal allows extraction of an analyte from a sample stream. The sensing is based on the change in device frequency (frequency shift), which is proportional to the mass of material sorbed by the coating on the crystal surface. The relationship between frequency shift and mass change was derived by Sauerbrey [9] and may be calculated using Eq. (1).

$$\Delta f = -2.3 \cdot 10^6 f_o^2 \frac{\Delta M_s}{A} \quad (1)$$

where Δf is the change in frequency of the quartz crystal in Hz; f_o is the fundamental frequency of the quartz crystal in MHz; ΔM_s is the mass of material deposited or sorbed onto the crystal in g and A is the area coated in cm^2.

Apparatus

The PZQ based array has six sensors with fundamental frequencies of 10 MHz. Each PZQ was coated with a commonly utilised gas chromatography stationary phase, each containing a different functional group to allow limited selectivity. A reference PZQ, allows for drift compensation.

PZ quartz crystal coating

The PZQ coatings used were OV-1, Carbowax 20M, OV-17, Diethylene glycol succinate, Silar 10C and OV-210. These were chosen to give a wide range of functional groups and polarities. Dilute solutions (0.1% *w*/*w*) of each coating were prepared in a volatile solvent, either chloroform ($CHCl_3$) or an 80:20 *v*/*v* mixture toluene:methanol. The solutions were applied to both sides of the crystal as the frequencies of the PZQ were monitored so that frequency shifts were similar for each coating. The sensors were conditioned prior to use by passing nitrogen over their surface for six h.

Sampling

10 mL of oil sample in a 125 cm^3 Dreschel bottle was stored at 45°C for 30 min to allow for headspace generation. During sampling the Dreschel bottle was kept at 45°C to allow dynamic headspace analysis. A valve switches between the reference and sample gas, flow rate for both was set at 17 mL min^{-1}. Sampling was performed over a 3-min cycle, 1 min base line reading (reference) and 2 min response (sample). After each reading the

sample chamber was purged with reference nitrogen for 5 min prior to the introduction of the next sample. A total of 346 samples were taken, consisting of 112 Extra Virgin Olive oil, 126 non-virgin Olive oil and 108 Sunflower oil samples.

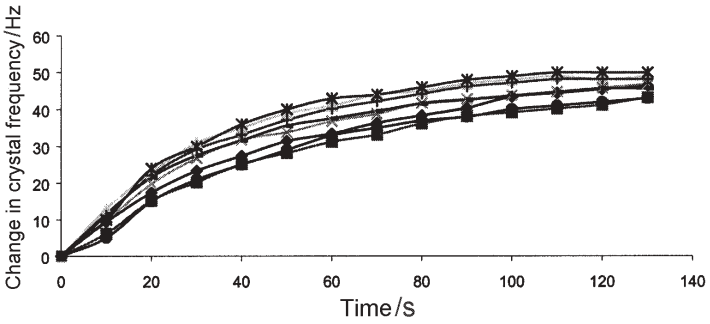

Fig. 1 Typical dynamic sensor response

Figure 1 shows a typical sensor response to an analyte, eight response sets are shown for an OV-210 coated sensor reacting to sunflower oil. The response curves approximate to an open loop first order response. To enable a reasonably stable reading to be made for classification, the response at 120 s was chosen. At this time the sensor has almost reached equilibrium with the analyte, small experimental errors will have little effect on the frequency change observed. The actual equation for the sensor response is a sum of two exponentials as given in Eq. (2) as described by Freeman [10].

$$\Delta f(t)=a_1(1-\exp(-a_2t))+a_3(1-\exp(-a_4t)) \quad (2)$$

where a_1 to a_4 are constants and t is time (s).

Feature extraction

The attributes of the gas (frequency change for each sensor) were calculated at 120 s response, the classification methods were applied using the same feature sets. The histogram of Fig. 2 shows the frequency shift for the 108 sunflower oil samples for the OV-210 coated sensor, the mean frequency change is 38 Hz and the standard deviation is 6.2 Hz. The overall shape of the histogram may be likened to a gaussian function as defined in Eq. (3). Similar analysis of the other sensors with all of the analytes gives similar results.

$$\mu(x_i)=\exp\left[-\frac{(x_i-\bar{x}_i)^2}{2\sigma_i^2}\right] \quad \text{where } i=1, 2,..., n \quad (3)$$

where $\bar{x}$ is the mean of x, the analyte data set for that sensor and σ is the corresponding standard deviation.

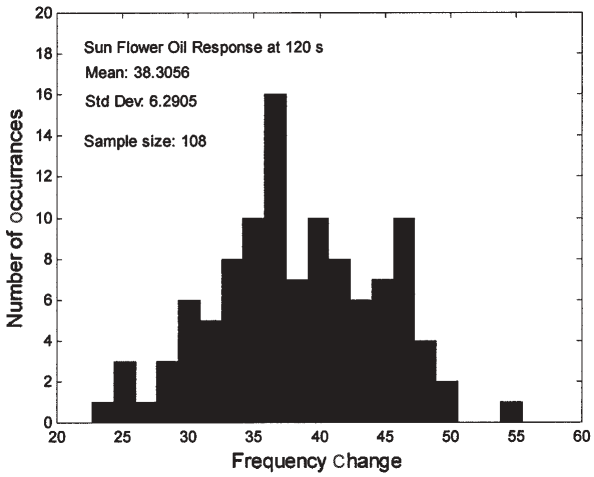

Fig. 2 Histogram of OV-210 sensor with sunflower oil

Principal component analysis

Principal Component Analysis (PCA) is initially used here for visualisation of the data. It is a commonly used multivariate technique [11, 12], which acts unsupervised. PCA finds an alternative set of axes about which a data set may be represented. It indicates along which axis there is the most variation; axes are orthogonal to one another. PCA is designed to provide the best possible view of variability in the independent variables of a multivariate data set. If the principal component scores are plotted they may reveal natural clustering in the data and outlier samples.

Figure 3 shows the first two principal components in a PCA scores plot for 346 oil samples. The data set consists of 112 Extra Virgin Olive (EVO) oil, 126 Olive oil (OI) and 108 Sunflower oil (SFO) samples. The data classes are clearly visible as clusters. The non-virgin olive oil forms a tight cluster to the right of the plot with a centre of (0.12, 0), the sunflower forms a less tight cluster to its left, centre (0.03, 0),

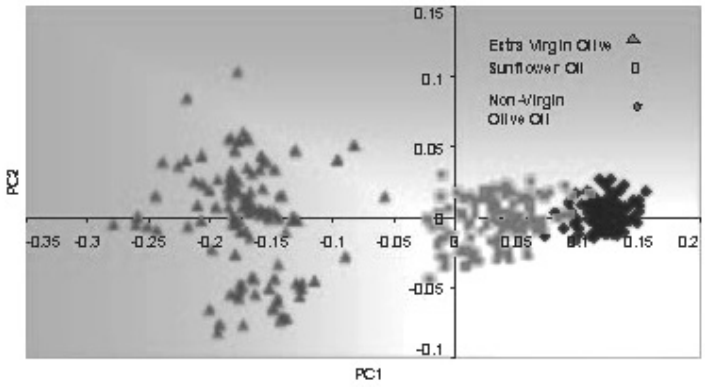

Fig. 3 PCA plot for edible oil data

and there is a small degree of overlap between these two classes. The extra virgin olive oil data forms a loose cluster to the left of the plot, centre (–0.18, 0); there are several outliers to this cluster some of which are closer to the sunflower cluster centre than the extra virgin olive oil cluster centre. The Extra virgin olive oil should be easily separated for classification, however the overlapping region between the non-virgin olive oil and the sunflower oil may cause some problems.

Radial basis neural network

Background

Radial basis neural networks were popularised by Broomhead and Lowe in the late 1980's [13], they are quick to train and conceptually elegant. Evans and co-workers [14] used a RBF network with data from an electronic nose to determine the quality of wheat. The nose used an array of 32 conducting polymer sensors; the output was then pre-processed before classification being made by the RBF network. The reported success rate was 92% correct on a simple judgement between good and bad quality.

Overview

In an RBF network the feature space is normalised $[0.1]^n$ and is filled with M overlapping radial based functions. The functions are continuous and reach a maximum value at the centre of the specific region covered, but assume a near zero value outside of it. There are several types of radial functions, the most popular being the gaussian. One way of describing an RBF network is that each radial function is a fuzzy set membership function in the feature space. Any feature vector x, belongs to one or more of the response regions, it is fuzzified by each radial basis function to the appropriate membership of that region. The output layer of neurones maps the combination of fuzzy memberships into an output representing the overall membership for a particular class. The weights from the hidden layer to the output layer form the fuzzy rules to perform the joins between the sub-clusters of the hidden layer. The masses must be trained to the appropriate activation to perform the correct mappings.

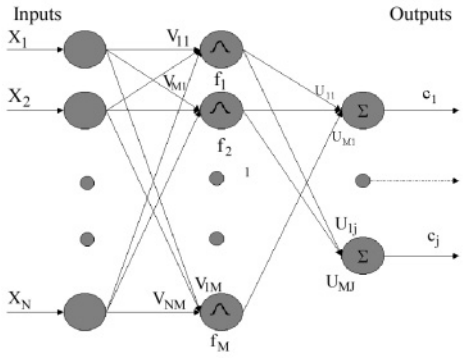

Fig. 4 Architecture of RBF classification network

Architecture

The gaussian functions equation is as given by Eq. (3). The centre of each RBF is placed on a small cluster that represents a subclass; therefore *M* functions cover the feature space. The σ or spread parameter may be adjusted so that covers a larger area; adjacent RBF's usually overlap to some degree. The neurones represented by the *M* centres make the single hidden layer of an *N-M-C* feed forward artificial neural network as shown in Fig. 4. The output layer *C* contains summing neurones with weighted connections to the hidden layer *M* that must be trained in a similar way to a multi layer perceptron network.

Operation

The operation of a trained network consists of presenting an input vector *x*, the input layer normalises the vector to [0.1]. The hidden layer to produce a scaled response then processes the normalised vector. Any input vector close to one of the *M* neurone centres will produce an output *y* that is greater than any other. The vector $y=(y_1,...,y_M)$ that is output from the hidden layer is processed by each neurone of the output layer. It is usual to use a summing function (Eq. (4)) or an averaging squashing function (Eq. (5)) rather than the multi layer perceptron sigmoid function.

$$c_i=\sum_{m=1}^{M} u_{mj}\, y_m \tag{4}$$

$$c_i=(1/S)\sum_{m=1}^{m} u_{mj}\, y_m,\ S=y_1+...+y_m \tag{5}$$

The output vector *c* is then tested against each of the target vectors that identify the classes. The greatest output represents the highest activation and thus the input vector *x* is recognised.

Training

Finding the centres, spread and weights of the hidden nodes constitutes training of an RBF network. For optimal performance of an RBF network the position of the centres and spread of the hidden nodes is critical. The full training algorithm for radial basis function networks of Looney [15] allows adjustment of the hidden neurone centres *v*, the spread parameter σ^2 and the output weights *u*.

Typically the steepest descent algorithm is used to train the output weights *u*, the total sum-squared error, *E* over all the *Q* input vectors is minimised. *t* is the target output vector that identify the classes. σ_m is initialised to 0.05, u_{mj} are set randomly to between (–0.5 to 0.5).

$$E=\sum_{q=1}^{Q}\sum_{j=1}^{J}(t_j^{(q)}-z_j^{(q)})^2 \tag{6}$$

If η is the network-learning rate, the steepest descent formula to optimise the output weights *u* is

$$u_{\mathrm{mj_{new}}} = u_{\mathrm{mj_{old}}} + (2\eta/M)\sum_{q=1}^{Q}(t_j^{(k(q))} - z_j^{(q)})\, y_m^{(q)} \tag{7}$$

Function centres

$$v_{n_{new}}^{(m)} = v_{n_{old}}^{(m)} + [2\eta/(M\sigma^2)]\sum_{q=1}^{Q}\left\{\sum_{j=1}^{J}(t_j^{(k(q))} - z_j^{(q)})u_{mj}\right\} y_m^{(q)}(x_n^q - v_n^q) \tag{8}$$

Spread parameter

$$\sigma^2_{m_{new}} = \sigma^2_{m_{old}} + (2\eta/M)\sum_{q=1}^{Q}\{\sum_{j=1}^{J}(t_j^{(k(q))} - z_j^{(q)})[u_{mj}\, y_m^{(q)} \left\| x^{(q)} - v^{(m)} \right\|^2/(2\sigma_m^4)]\} \tag{9}$$

Choice of architecture

The choice of the number of hidden nodes in an RBF may affect classification performance [12, 14]. Typically a large number of hidden nodes are required to adequately describe the feature space with enough resolution to classify to any degree of accuracy. If however sufficient knowledge of the system is known then a suitable architecture of the network may be obtained from that. The variance for each sensor in the oil data has been shown to correspond to a gaussian distribution, Fig. 2, and then if each hidden node in the network corresponds to a sensor, the network is simply mapping fuzzy sensor responses to classifications. The weight from any given input node to any given hidden node represents the mean frequency of that sensor for a given analyte. The spread parameter σ^2 is a representation of the standard deviation of that sensor to all the oils and the weights from the hidden layer to the output layer represent the fuzzy rules that map the analyte to the classification. There are six sensors in this nose system, therefore six input nodes to be used in the RBF network, each input normalises the crystal response. Six hidden nodes represent the six sensors, producing a fuzzy membership for a given analyte. Three output nodes represent the three classes of oil. The network architecture to be used therefore is 6:6:3.

Table 1 RBF 6-6-3 network results

Class	Non-virgin olive oil	Sunflower oil	Extra virgin olive oil	Total	Correct/%
Non-virgin olive oil	42	0	0	42	100
Sunflower oil	1	33	0	34	97.06
Extra virgin olive oil	0	0	37	37	100
			Total correct	112	99.13

Results using RBF networks

Table 1 shows the results of using a 6-6-3 RBF network trained using 233 data sets and tested on 113 sets selected randomly from the full data set. Only one sunflower oil sample has been misclassified as non-virgin olive oil.

Discussion

The results demonstrate that the fuzzy analogy of the RBF network to the piezoelectric quartz crystal response pattern is a valid one. The PCA scores plot of Fig. 3 shows that the overlapping region of the non-virgin olive oil and sunflower oil could cause difficulty correctly classifying data that lies in this area. The trained fuzzy mapping of the RBF network misclassifies few points in the region of uncertainty.

Conclusions

The RBF network provides a simple and effective method of distinguishing between the vapour signatures of the edible oils. The fuzzy analogy between the sensor response patterns and the radial basis functions of the hidden nodes in the network provided an architecture that accurately modelled the piezoelectric crystal based electronic nose.

* * *

The authors acknowledge ERDF funding under contract 70/41/001 for this project. We also wish to acknowledge support of Kevin Middleton and Jill Legg in the construction of the electronic nose.

References

1 Z. Ali, W. T. O'Hare, T. Sarkodie-Gyan, B. J. Theaker and E. Watson, Conference Proceedings SPIE, Boston, September 1999.
2 R. D. Hiserodt, C. T. Ho and R. T. Rosen, ACS SYM SER, 660 (1997) 80.
3 K. Persaud and G. H. Dodd, Nature, 299 (1982) 352.
4 Y. G. Martin, J. L. P. Pavon, B. M. Cordero and C. G. Pinto, Anal. Chim. Acta, 384 (1999) 83.
5 J. P. Paulsson and F. Winquist, Foren. Sci. Int., 105 (1999) 95.
6 R. Stella, J. N. Barisci, G. Serra, G. G. Wallace and D. DeRossi, Sens. Acts B, 63 (2000) 1.
7 Z. Ali, W. T. O'Hare, T. Sarkodie-Gyan and B. J. Theaker, J. Therm. Anal. Cal., 55 (1999) 371.
8 M. Fang, K. Vetelino, M. Rothery, J. Hines and G. C. Frye, Sens. Acts B, 56 (1999) 155.
9 G. Z. Saubrey, Z. Phys., 155 (1959) 206.
10 N. J. Freeman, I. P. May and D. J. Weir, J. Chem. Soc. Faraday Trans., 90 (1994) 751.
11 H. G. Byun, K. C. Persaud, S. M. Khaffaf, P. J. Hobbs and T. H. Misselbrook, Computers and Electronics in Agriculture, 17 (1997) 233.
12 R. Callan, The Essence of Neural Networks, Prentice Hall., 1999
13 D. S. Broomhead and D. Lowe, Multivariate functional interpolation and adaptive networks, Complex Systems. Vol. 2, 321–355., 1988
14 P. Evans, K. C. Persaud, A. S. Mceish, R. W. Sneath, N. Hobson and N. Magan, Sens. Acts B, 69 (2000) 348.
15 C. G. Looney, Pattern Recognition using Neural Networks, Oxford University Press, 1997