ELSEVIER

Analytica Chimica Acta 348 (1997) 503–509

ANALYTICA
CHIMICA
ACTA

Neural network based electronic nose for the classification of aromatic species

J. Brezmes*, B. Ferreras, E. Llobet, X. Vilanova, X. Correig

Departament d'Enginyeria Electrònica, Universitat Rovira i Virgili, Autovia de Salou s/n, 43006 Tarragona, Spain

Received in revised form 21 February 1997; accepted 24 February 1997

Abstract

In this work, an aroma identification system has been developed. Based on an array of semiconductor tin dioxide gas sensors and neural network processing algorithms, the system has proven a 100% success rate in the discrimination of five different aromatic species. An initial nine sensor array was simplified to seven after a PCA analysis detected redundancy between three of the sensors. Data processing and classification performed by a feedforward artificial neural network with a hidden layer and trained with a backpropagation algorithm showed no significant performance differences between the complete and reduced sensor array which confirms the redundancy detected by the PCA analysis. Our results show that a reliable Electronic Nose system can be designed using inexpensive and poorly selective chemical semiconductor gas sensors.

Keywords: Tin oxide gas sensor; Odour recognition; Electronic nose; Principal component analysis; Artificial neural networks

1. Introduction

Despite considerable and sustained attempts to develop new electronic instrumentation capable of mimicking its remarkable ability, the human nose is still the primary 'instrument' used to assess the smell or flavour of various industrial products. The design of a system capable of discriminating between different odours could have many applications in the cosmetic, food and beverage industries. The term 'Electronic nose' has been widely accepted to refer to this kind of system and it could be defined as "an instrument which comprises an array of electronic chemical sensors with partially overlapped sensibilities and an appropriate pattern-recognition system capable of recognising simple or complex odours" [1–4].

Important information applicable to an 'Electronic nose' can be obtained looking into the human olfactory system. For example, although there are believed to be a relatively small number of receptor proteins (about 100) with partially overlapped sensibilities and a low sensibility (in the order of ppm's), subsequent neural processing enhances receptor sensitivity by about three orders of magnitude, removes drift and provides discrimination between several thousand odours.

Multisensor systems have employed conducting polymers [5], Langmuir-Blodgett films [6] and metal oxides [7,8] as the sensing materials. The combination of sensor arrays and pattern recognition techniques has been used with some success in the analysis of the

*Corresponding author. Tel.: +34-77-559619; fax: +34-77-559710.

0003-2670/97/$17.00
PII S0003-2670(97)00202-X

response of tin oxide sensor arrays to tobaccos, wines and other alcoholic beverages [9,10].

Tin oxide semiconductor chemical sensors use their electrical resistance change to detect reducing vapours. The porous tin oxide contains oxygen vacancies in its lattice. Electrons that can be thermally activated are trapped in these vacancies. The chemical process involved in their sensing mechanism is the adsorption and reaction of reducing vapours with oxygen adsorbates. As a result of the adsorption and reaction, the sensor resistance decreases due to the release of electrons trapped in oxygen adsorbates [11,12]. These sensors are inexpensive, easy to control and remarkably sensitive to a wide spectrum of reducing vapours. On the other hand, they have some drawbacks such as being poorly selective and showing a temperature and moisture-dependent behaviour. Nevertheless, these problems can be overcome with proper electronic processing [13–15].

In this paper, we introduce an 'Electronic nose' capable of discriminating among five different aromatic species (cinnamon, red pepper, thyme, pepper, and nutmeg). The system is composed of an array of nine commercially available tin oxide chemical sensors and an Artificial Neural Network (ANN) as the pattern recognition processing algorithm.

In Section 2, the experimental set-up and measuring procedures used in the development of this work are described. In Section 3 the main results are presented and discussed: In the first stage, a Principal Component Analysis (PCA) is performed to evaluate the relevance of each sensor of the array in the identification of the aromatic species. The same procedure is used to detect sensor redundancy. In the second stage, an ANN algorithm is applied to different sets of the acquired data to determine those experimental samples and pre-processing algorithms which work better in the identification task.

2. Experimental

2.1. Measured species

The aromatic species that were tested were cinnamon, red pepper, thyme, pepper and nutmeg, all of them in powdered form. To carry out measurements, these species were introduced in a test chamber kept at room temperature and ambient humidity.

2.2. Sensor array and experimental set-up

The sensor array used consisted of nine commercially available Taguchi gas sensors, listed in Table 1. The sensors were operated as suggested in [16]. To measure the sensor conductance (or resistance), the voltage in a grounded resistor in series with the sensor was monitored. The circuit conditioning is shown in Fig. 1. The relationship between the sensor resistance and the monitored voltage is expressed by the following equation:

$$R_s = \frac{V_{sup} R_G}{V_{R_G}} - R_G$$

To reach the working temperature, the sensors were heated by applying a 5 V DC voltage to their heater resistance (R_H).

The experimental set-up is shown in Fig. 2. The electrical response of the nine gas sensors was mon-

Table 1
Array components, main applications and detection ranges

Model	Main applications	Typical detection ranges (ppm)
TGS 800	Air quality control	1–10 (Cigarette smoke, gasoline exhaust...)
TGS 813	Combustible gas detection	500–10,000 (General combustible gases)
TGS 822	Solvent vapour detection	50–5,000 (Organic solvents)
TGS 824	Toxic gas detection	30–300 (Ammonia)
TGS 825	–	5–100 (H_2S)
TGS 830	Halocarbon gas detection	100–3,000 (R-22, R-113)
TGS 842	Methane detection	500–10,000 (Low sensitivity to interfering gases)
TGS 880	Cooking control	(Volatile gases and water vapour from food)
TGS 882	–	50–5,000 (Alcohol vapour from food)

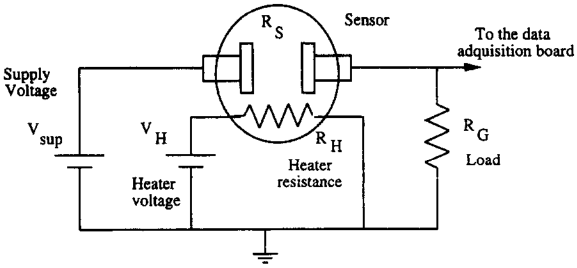

Fig. 1. Electronic circuit to monitor the sensor response.

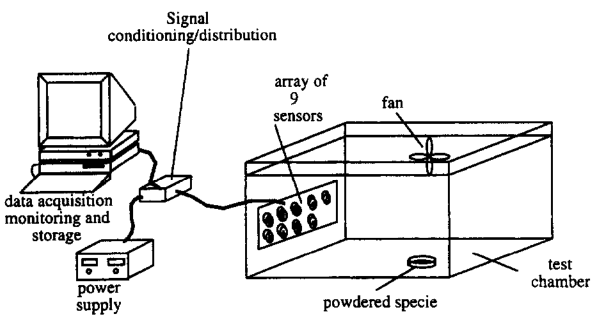

Fig. 2. Experimental set-up (the test chamber is not to scale).

itored by an acquisition data board (PC-LabCard© PCL-818, 12 bit ADC) installed in an IBM compatible Personal Computer (with a written-in-house data acquisition program). The experimental data was displayed in real-time on the computer screen and stored as text files on disk for later processing. Excel was used for graphical presentation. MATLAB™, PLS_Toolbox [17] and Neuralex were used for data analysis.

2.3. Procedures

The experiment was started when the resistance of the gas sensors remained stable to a high value (typ. 200 kΩ–1 MΩ). A fan homogenised the air in the test chamber. A little amount of the species to be tested was introduced into the test chamber. Then, the data acquisition started. The response of the sensors was acquired and stored every 5 s as described above for one hour. Because the main application of such a system would be in environments where neither temperature nor humidity are controlled, no special care was taken to monitor and control these parameters inside the test chamber even though the behaviour of tin dioxide gas sensors is heavily influenced by environmental conditions.

The total number of experiments carried out were 50, ten for each species, giving a total of 50 different sets of values collected by each sensor.

3. Results and discussion

3.1. Data conditioning

Experimental measurements showed that aroma evaporation was a very slow process. For that reason the complete stabilisation of some of the signals from the sensors lasted up to half an hour since the placement of the species inside the test chamber. Fig. 3 shows that after a sharp initial transition the sensor response changed slowly. So, in order to determine the optimal measurement time, three different time durations (750 s, 1500 s and 3500 s) were considered to evaluate the discrimination performance of the system. Samples at 1500 s after the introduction of the species inside the test chamber gave the best overall discrimination results.

Another problem that arose was to determine the pre-processing algorithms that had to be applied to the sensor signals before they were used as the input of the processing algorithms. Four different pre-processing algorithms were studied. Preliminary studies showed that the normalised electrical resistance increment gave better results than the normalised conductance. The resistance normalisation is described by the following equation:

$$\Delta R_n = \frac{R_f - R_0}{R_0}$$

Figs. 4 and 5 show two polar plots, where the signal of sensor TGS-800 has been normalised. The radar plots are proportional to the normalised electrical resistance of each sensor. Fig. 4 compares the patterns of measurements for different species while Fig. 5 compares the patterns of different measurements for the same species. It can be seen that each species has a clearly different pattern and that different measurements of the same species have only slight differences between their patterns, probably due to changing environmental conditions and poor sensor repeatability.

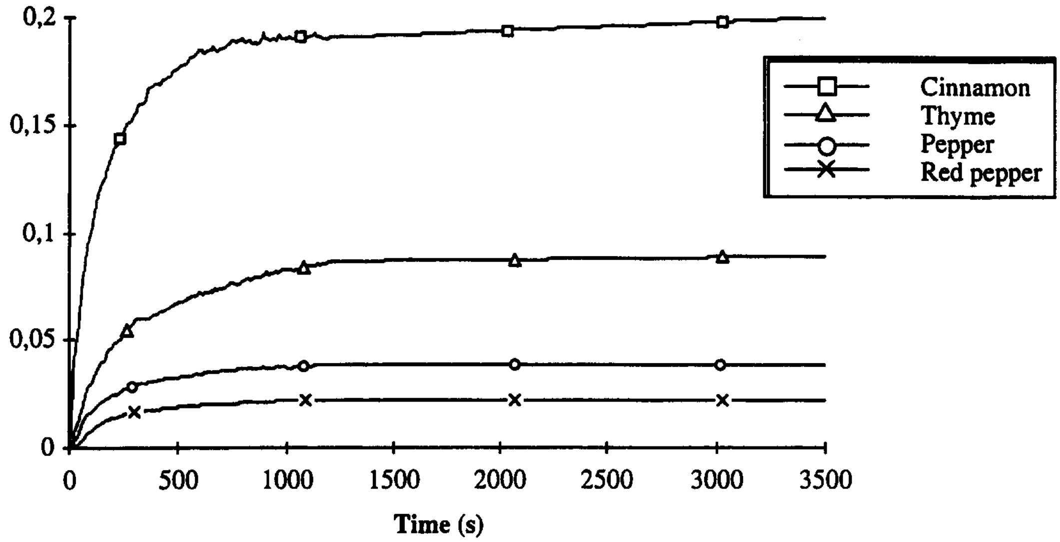

Fig. 3. Typical sensor response for the tested species.

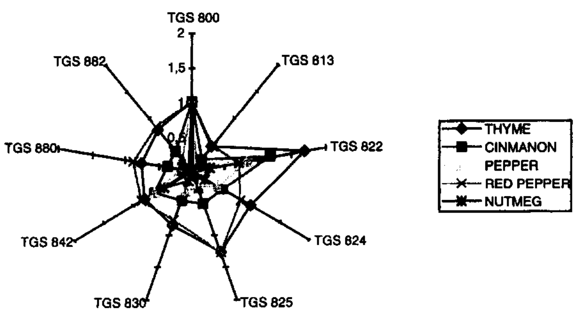

Fig. 4. Radar plot for five different species normalized to the electrical resistance increment of sensor TGS-800

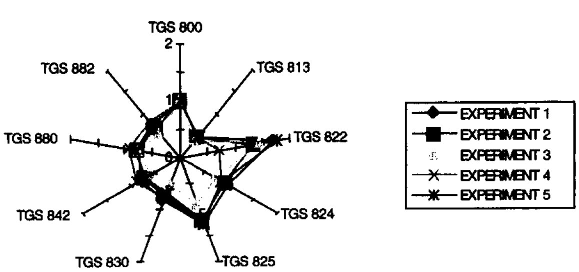

Fig. 5. Radar plot for five different measurements of the same species normalized to the electrical resistance increment of sensor TGS-800

3.2. PCA analysis

Once the input data was acquired, and before the pre-processed signals were fed to a feedforward ANN, a PCA confirmed the discrimination ability of the sensor array.

In the response matrix of the array of measurements each column was associated to a sensor while each row was related to a different experiment. Table 2 shows the percentage of this matrix captured by the PCA model. The first two principal components contained more than 98% of the total variance. These two first principal components concentrated the most relevant information to classify the species. In Fig. 6 the

Table 2
Percent variance captured by the PCA model

# P.C.	Eigenvalue	% Var.	% Tot. Var.
1	2.5738	95.1575	95.1575
2	0.0816	3.0186	98.1761
3	0.0409	1.5105	99.6866
4	0.0036	0.1321	99.8187
5	0.0018	0.0651	99.8838
6	0.0016	0.0604	99.9442
7	0.0010	0.0360	99.9802
8	0.0005	0.0170	99.9972
9	0.0001	0.0028	100

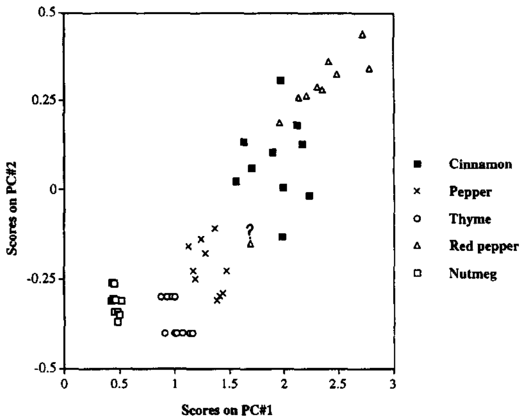

Fig. 6. Scores plot for the five studied species keeping the two first principal components.

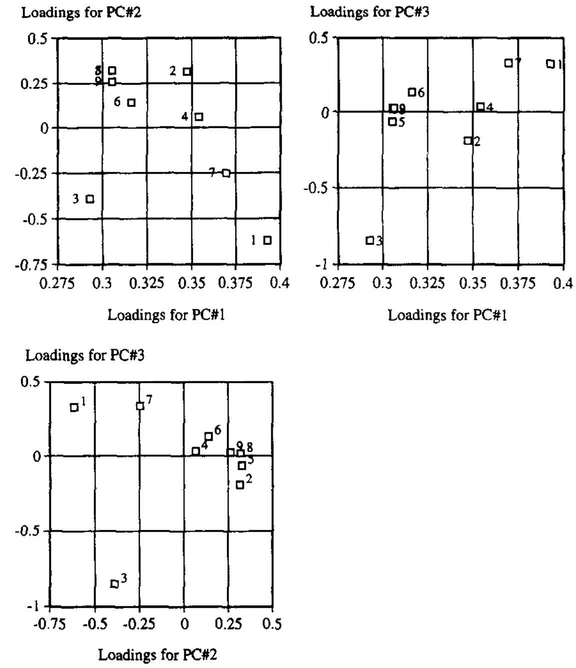

Fig. 7. Loading plots of the array components: (a) PC#1 vs. PC#2, (b) PC#1 vs. PC#3, (c) PC#2 vs. PC#3.

scores for PC#1 and PC#2 are shown. Five different classes can be easily differentiated in this Figure. There is a clear separation from one to another. Nevertheless between classes A (red pepper) and B (cinnamon) there is a slight overlapping. One measurement with red pepper (labelled with a question mark in Fig. 6) can be identified as an outlier. The use of three principal components did not improve the separation between the overlapped classes.

The PCA analysis also helped to detect redundancy between the input signals. Three sensors were found to give redundant information which led to a reduction in the number of sensors that formed the array. In Fig. 7 the sensor loadings are plotted when the first three PC have been retained. One can see clearly in Fig. 7 a), b) and c) how loadings for sensors 5 (TGS 822), 8 (TGS 825) and 9 (TGS 882) are very close. Since similar loadings imply colinearity and redundancy, these three sensors were giving similar information. Thus, removing sensors 8 and 9 from the array was considered before applying an ANN to classify the 50 measured samples.

3.3. Neural network processing algorithms

To evaluate the network performance two different strategies were carried out. These strategies evaluated different concepts:

1. To evaluate the method of using ANN to identify species, the 'leave one out' approach was used. This method trains 50 different networks (all with the same structure) using 49 training vectors out of the 50 measured samples and evaluating with the remaining one.
2. To evaluate discrimination ability of a particular network, 5 randomly selected vectors of the 10 available for each species were used to train a network. The remaining 25 were used as test vectors.

The tested ANN were feedforward fully-connected networks composed by seven (one neurone for each pre-processed sensor signal taking out redundant sensors) or nine input neurones (for the non-reduced array), a hidden layer with a number of neurones between 15 and 30 and an output layer with five neurones, one for each aromatic species. An ideal output would yield a 1 for the neurone assigned to the species sampled and a 0 for the rest of the output neurones. Fig. 8 represents the network structure. The training algorithm used was a modified backpropagation, with an adaptive learning rate of 0.01 and a momentum term of 0.8.

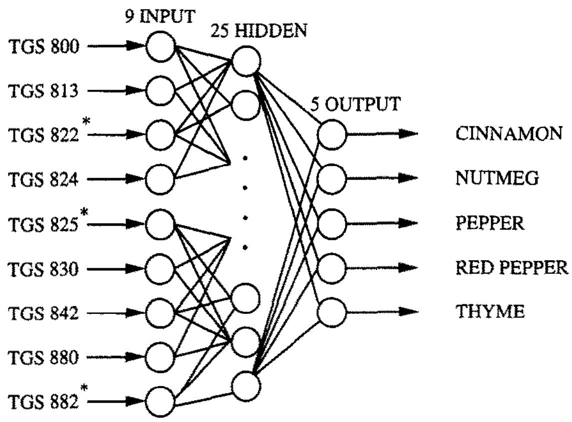

Fig. 8. Diagram of the feed-forward artificial network evaluated.

Table 3
Species classification success rate (in %). Comparison between the complete (ANN: 9 input, 25 hidden and 5 output neurones) and the reduced (ANN: 7 input, 20 hidden and 5 output neurones) sensor array and asscociated ANN. The evaluation was performed both with 25 training and 25 evaluating vectors and with the leave one out approach

Test method	Network structure	
	Complete	Reduced
25+25	98	100
'Leave one out'	98	98

With the array dimension reduced, a network consisting of seven inputs and five outputs was trained and evaluated. The number of neurones in the hidden layer were varied in order to obtain the best results. With 25 training and 25 evaluating vectors a 100% identification success rate was reached. On the other hand, evaluating the method using the 'leave one out' approach led to a 98% identification success rate (1 error: a pepper measurement was classified as cinnamon).

These results are almost identical to the ones obtained using the nine sensor array, which is in good agreement with the PCA conclusions, where redundancy between sensors 5, 8 and 9 was found. Table 3 compares the results obtained between the nine sensor array and the reduced seven input system.

4. Conclusions

An 'Electronic nose' capable of discriminating among five aromatic species (thyme, cinnamon, pepper, red pepper and nutmeg) has been introduced. A fully-connected feedforward, backpropagation-trained ANN processes the signals of a nine sensor array to perform the classification of the species. The array is composed of nine commercially available tin oxide semiconductor gas sensors. These sensors have overlapped sensibilities and they are sensible to a wide spectrum of reducing gases.

The measurement and conditioning of the input signals have been studied and it has been found that the normalised resistance changes of the sensor, measured 1500 s after species exposition gives the overall best results.

The PCA has pointed out the existence of redundancy in the sensor array. Three redundant sensors were detected and two of them were removed. In that case, a network with a 7 input, 20 hidden and 5 output structure, trained with 25 samples and evaluated with 25 test patterns shows a 100% recognition success rate. The same network structure was evaluated with the leave-one-out approach, giving a 98% success rate.

These results show that proper electronic processing leads to excellent aroma recognition using inexpensive and non-selective chemical sensors.

Acknowledgements

This work has been partially funded by the Comisión Interministerial de Ciencia y Tecnología (CICYT), project No. MAT96-1143-C02-01.

References

[1] J.W. Gardner and P.N. Barlett, Sensors Actuators, B 18,19 (1994) 211–220.
[2] C. di Natale, F. Davide and A. d'Amico, Sensors Actuators, B 23 (1995) 111–118.
[3] T. Moriizumi, T. Nakamoto, M. Wilkinson, in: J.W. Gardner and P.N. Bartlett (Eds.), Sensors and Sensory Systems for an Electronic Nose, Kluweer, Amsterdam, 1992, pp. 217–236.
[4] P.E. Keller, R.T. Kouzes, L.J. Kangas, World Congress on Neural Networks, San Diego, USA, 1994, vol.I, 269–272.
[5] B. Bott and T.A. Jones, Sensors Actuators, 9 (1986) 19–25.

[6] M. Fard, M.Sc. Thesis, University of Warwick, 1985.
[7] W.P. Carey, K.R. Beebe and B.R. Kowalski, Anal. Chem., 59 (1986) 860–866.
[8] H. Abe, T. Yoshimura, S. Kanayama, Y. Takahashi, Y. Miyashita and S. Sasaki, Anal. Chim. Acta, 194 (1987) 1–9.
[9] H.V. Shurmer, J.W. Gardner and H.T. Chan, Sensors Actuators, 18 (1989) 361–371.
[10] C. di Natale and A. d'Amico, Sensors Actuators, B 24,25 (1995) 801–804.
[11] P.T. Moseley, J.O.W. Norris, D.E. Williams, Techniques and Mechanisms in Gas Sensing, Adam Hilger, Bristol, 1991.
[12] N. Taguchi, US Patent 3 631 436, 1971.
[13] W. Göpel (Ed.), Sensors, a Comprehensive Survey, vol. 2, Chemical and Biochemical Sensors, VCH, Weinheim, 1991.
[14] A. Hierlemann and W. Göpel , Sensors Actuators, B 26,27 (1995) 126–134.
[15] J. Brezmes, E. Llobet, X. Vilanova, X. Correig, VI Chemometrics in Analytical Chemistry, CAC'96, Tarragona, Spain, 1996.
[16] Figaro Gas Sensor Technical Reference, 1992, Figaro Engineering Inc. USA: P.O. Box 357, 1000 Skokie Blvd. Room 160 Wilmette, Illinois 60091 USA. Tel: (1) 708-256-3546.
[17] B.M. Wise, PLS_Toolbox Version 1.4, Eigenvector Technologies, West Richland, 1995.