Meas. Sci. Technol. **10** (1999) 538–548. Printed in the UK PII: S0957-0233(99)01657-4

Non-destructive banana ripeness determination using a neural network-based electronic nose

Eduard Llobet†, Evor L Hines‡, Julian W Gardner‡ and Stefano Franco‡

† Department of Electronic Engineering, Universitat Rovira i Virgili, Autovia de Salou s/n, 43006 Tarragona, Spain
‡ Division of Electrical and Electronic Engineering, School of Engineering, University of Warwick, Coventry CV4 7AL, UK

Received 8 February 1999, in final form and accepted for publication 26 March 1999

Abstract. An electronic nose based system, which employs an array of inexpensive commercial tin-oxide odour sensors, has been used to analyse the state of ripeness of bananas. Readings were taken from the headspace of three sets of bananas during ripening over a period of 8–14 days. A principal-components analysis and investigatory techniques were used to define seven distinct regions in multisensor space according to the state of ripeness of the bananas, predicted from a classification of banana-skin colours. Then three supervised classifiers, namely Fuzzy ARTMAP, LVQ and MLP, were used to classify the samples into the observed seven states of ripeness. It was found that the Fuzzy ARTMAP and LVQ classifiers outperformed the MLP classifier, with accuracies of 90.3% and 92%, respectively, compared with 83.4%. Furthermore, these methods were able to predict accurately the state of ripeness of unknown sets of bananas with almost the same accuracy, i.e. 90%. Finally, it is shown that the Fuzzy ARTMAP classifier, unlike LVQ and MLP, is able to perform efficient on-line learning in this application without forgetting previously learnt knowledge. All of these characteristics make the Fuzzy-ARTMAP-based electronic nose a very attractive instrument with which to determine non-destructively the state of ripeness of fruit.

Keywords: non-destructive testing, fruit ripeness, electronic nose, Fuzzy ARTMAP, neural networks, tin-oxide sensors

1. Introduction

One of the most important objectives in the food industry is that of achieving a uniform quality both of raw materials and of the final product. One of the main concerns of the fruit industry is the systematic determination of fruit ripeness under harvest and post-harvest conditions, because a variability in ripeness is perceived by consumers as a lack of quality.

Most of the traditional methods that have been used to assess fruit ripeness are destructive and thus cannot be so readily applied. Some of these methods rely on the measurement of fruit firmness, which is correlated to ripeness, by using a penetrometer [1] or an impact force [2]. Other strategies include measuring levels of chemical species and parameters that are correlated to ripeness such as pH, titratable acidity [3], soluble-solids contents (i.e. sugars) [4] and ethylene contents [5].

The measurement of these chemical species and parameters generally requires once again the destruction of the fruit together with complex analytical techniques, such as gas–liquid chromatography (ethylene) and titration (acidity) [6].

More recently, non-destructive methods of determining fruit ripeness have been proposed. These methods include

(i) nuclear magnetic resonance (NMR) [7] and proton magnetic resonance (PMR) [8] to determine levels of soluble solids,
(ii) vision systems to determine skin colour of fruit (colour has been shown to be correlated to sugar contents) [9] and
(iii) acoustic methods to determine fruit firmness [10].

However, there are problems inherent to these approaches. NMR and PMR techniques rely upon expensive equipment. Acoustic methods present the problem of how to reproducibly couple the acoustic impedances of the emitter and the fruit to be tested. Fruit colour is, in some cases, only weakly correlated to fruit ripeness.

An attractive and alternative strategy for determining the state of ripeness of fruit consists of sensing the aromatic volatiles emitted by fruit using electronic artificial nose systems [11, 12]. Electronic-nose systems appear to be very promising for non-destructively determining fruit ripeness

0957-0233/99/060538+11$19.50

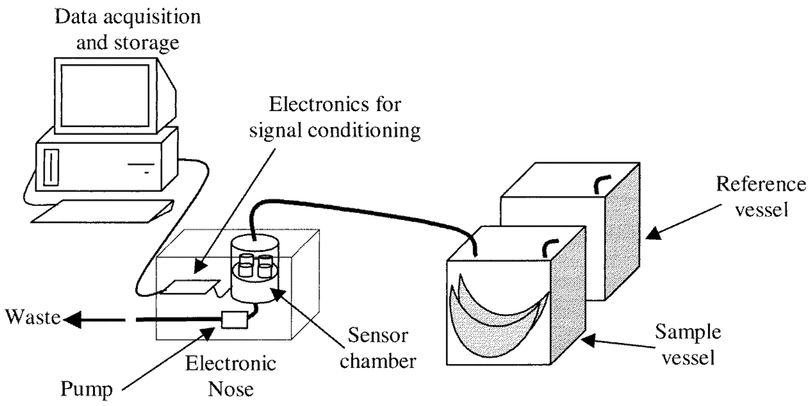


Figure 1. The experimental set-up used to analyse the headspace of bananas.

for a number of reasons. The main ones are that they are based on inexpensive, non-specific solid-state sensors, which are sensitive to ethylene (the ripening hormone in climacteric fruit, such as apples, peaches, bananas, etc.) and to other volatile compounds emitted by fruit during ripening. Furthermore, once an electronic nose has been 'trained', it does not require a skilled operator and can obtain the results in a few seconds.

In electronic-nose equipment, the sensor signals are processed by a pattern-recognition engine that allows the system to analyse complex aromas. Neural networks have been used extensively to perform pattern recognition and good results have been reported for the classification of foodstuffs, such as beverages [13], coffees [14], fish and meat [15, 16]. The back-propagation-trained multilayer perceptron (MLP) paradigm is the most popular pattern-recognition method in aroma analysis today. However, it has recently been shown [17] that, in some cases, the Fuzzy ARTMAP paradigm [18–20] outperforms MLP in aroma-classification problems. Another promising technique is learning vector quantization (LVQ), which is a supervised technique based on the self-organizing-map (SOM) paradigm [21]. In this paper, we report on the use of an electronic nose employing an array of four tin-oxide sensors, in combination with a pattern-recognition engine (Fuzzy ARTMAP, LVQ and MLP neural networks), to classify the ripeness of bananas.

2. Experimental procedures

2.1. Materials

Three sets of bananas were purchased in turn from a local supermarket. For each set and without any additional manipulation, the bananas were placed into standard plastic food vessels (of volume 5 l). Figure 1 shows the experimental set-up. The vessels had two small holes in their covers, to allow the headspace to be analysed with the electronic-nose equipment. The ambient conditions (temperature and humidity) of the room in which the bananas were kept were monitored throughout the test procedure. Each time that a new set of bananas was analysed, new plastic vessels were employed.

2.2. Test procedures

The sensor system comprises four tin-oxide odour sensors from four different manufacturers (see table 1) housed in a sensor chamber. The electrical conductances of the sensors vary in the presence of reducing/oxidizing gases. In general, the conductance of sensors was found to increase as the bananas ripened. A thin plastic tube was connected from the input to the sensor chamber to one of the holes in the cover of the appropriate plastic vessel. A diaphragm pump (Vacuum Pump Manufacturing Co Ltd, UK) was used to facilitate sampling of the headspace of the vessel. The headspaces of the vessel containing the bananas (seven bananas) and the reference vessel were sampled in sequence as follows.

(i) For the banana vessel, a sample measurement typically took 10 min to complete. The flow rate was 2 l min^{-1}. The air removed from the fruit vessel by the pump was replaced by air from the room.

(ii) For the reference vessel, the tube from the input to the sensor chamber was connected to an empty plastic vessel (the reference vessel) and air from the room was then pumped into the sensor chamber from the vessel. In this way the sensors were allowed to return to their baseline over a period of some 50 min after sampling the headspace of the banana vessel. This was to make sure that the electronic-nose system was responding to the banana aromas rather than to any residual smell of the plastic vessel.

One measurement comprises taking, alternately, a headspace sample from the banana vessel and one from the reference vessel.

During the measurements, a sample of each sensor resistance was taken every 100 ms and stored in a data file for subsequent processing. The acquisition and storage system was controlled using LabVIEW© software (National Instruments Inc) [22].

2.3. The use of banana-skin colour as a class reference

Skin colour has been shown to be related to pigment changes within the skin of the banana [23]. Furthermore, skin colour has been shown to be correlated quite well to the sugar contents, development of flavour and textural characteristics of the pulp. The ripeness of bananas can be assessed by

Table 1. Commercially available metal oxide sensors used in the electronic nose.

Sensor	Manufacturer	Sensitive to
TGS 822	Figaro Engineering Inc, Japan	Volatile organic compounds
SP-11(P1)	FIS, Japan	Hydrocarbons, combustible gases
MGS 1100	Motorola, USA	Carbon monoxide
AAS14	Capteur Ltd, UK	Volatile organic compounds

comparing the colour of the skin with standardized colour charts [23]. We therefore elected to use skin colour as a scale for the non-destructive classification of the various states of ripeness of the bananas. Twice a day, a picture of the bananas was taken as a visual record of the various states of ripening. Using the test procedures described above, three data sets were gathered as follows:

(i) set 1: 49 measurements were performed with the electronic nose over a period of 8 days;
(ii) set 2: 35 measurements were performed over seven consecutive days; and
(iii) set 3: 91 measurements were gathered over a period of 14 consecutive days.

The bananas used for collecting each data set were initially as fresh as possible, i.e. as green as possible, but exhibited slight differences in their state of ripeness (according to their colour). They would also have experienced all sorts of variations prior to arrival in our laboratory. However, all bananas of a given set ripen simultaneously. The numbers of days, samples, etc. were limited by practical circumstances. Additionally, for example, no data were collected overnight (for the three data sets) and over the weekend (for the first data set).

3. Data analysis

3.1. Signal pre-processing

The choice of the data pre-processing algorithm has been shown elsewhere to affect the performance of the pattern-recognition stage. In this case a difference model (i.e. the static change in sensor resistance) was used: $\Delta R = R^{air} - R^{odour}$. The complete banana data set was then normalized, by dividing each ΔR by the maximum value, to set their range to [0, 1]. This is necessary when the Fuzzy ARTMAP neural network is used; and, for consistency, the same normalization was used for the MLP and LVQ. All of the neural networks were simulated using the NeuralWorks Professional II/Plus software from NeuralWare Inc, USA [24].

3.2. Clustering of data

The use of PCA, SOM and Fuzzy-cluster analysis to assess clustering within the data sets is now discussed. Several classification methods were applied to verify that the categories established by each method were not arbitrary. The objective of using these methods was to establish classes according to the state of ripeness of the bananas.

PCA is a linear method that has been shown to be effective for discriminating between the responses of an electronic nose to simple and complex odours [25]. The method consists of expressing the response vectors in terms of

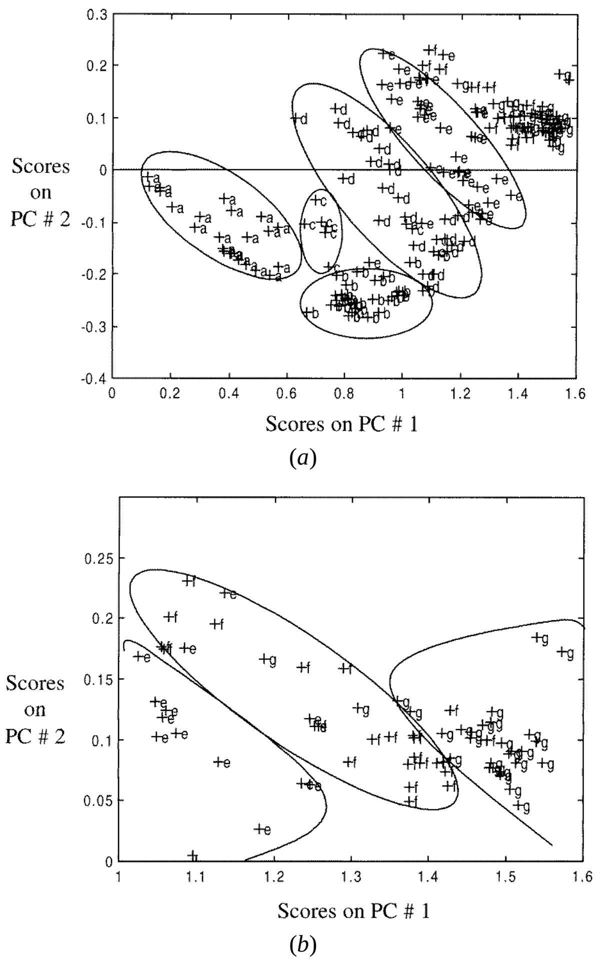


Figure 2. (*a*) Results of the principal-components analysis of the response of a four-element tin-oxide sensor based electronic nose to the bananas' aromas. Seven clusters or categories appear. These correspond to seven states of ripeness. (*b*) A detailed expansion of an area of (*a*), showing the boundaries corresponding to categories e, f and g.

a linear combination of orthogonal vectors. Each orthogonal (principal) vector accounts for a certain amount of variance in the data, with a decreasing degree of importance.

PCA was used to investigate how the response vectors from the sensor array cluster in multisensor space. The objective of this analysis was to establish simple categories for the state of ripeness of the bananas. The results of the PCA, using the data normalized as described in section 3.1, are shown in figures 2(*a*) and (*b*). Figure 2(*b*) is an enlargement of an area in figure 2(*a*) to better show the clusters formed. Two principal components were kept, which accounted for 99.7% of the variance in the data (PC 1 and PC 2 explained 88.2% and 11.5% of the variance, respectively). Seven categories appear to be evident. These categories are summarized in table 2. Most of the variance

Table 2. Categories of the state of ripeness of the bananas for the three data sets studied. The numbers in the table indicate the day on which the measurements were performed. Data set 1 was gathered towards the end of September, data set 2 in October and data set 3 in November.

Category	Set 1	Set 2	Set 3
a		1, 2	1, 2
b	1–3	3	
c		4, 5	
d	4	6, 7	3
e	6–8		4, 5
f			6–8
g			9–14

in the data is explained by considering the two first principal components, which implies that the sensor responses were highly correlated; this is generally the case with metal-oxide gas sensors. The loadings associated with figure 2 were (0.5195, −0.5705) (0.302, 0.2032) (0.5502, −0.3313) and (0.6530, 0.7236) for sensors 1, 2, 3 and 4, respectively. The loadings for PC 1 of sensors 1, 3 and 4 were rather similar. However, since the loadings for PC 2 were quite different, none of the sensors was omitted.

Table 3 shows the correspondence between the categories established and banana-skin colour determined from the photographs. In the context of table 3, it must be borne in mind that there will be some effect due to the fact that the bananas in the three sets had been purchased at different times. Therefore, they will be in different states of initial ripeness.

According to tables 2 and 3, the ripening process varied from one set to another. In particular, bananas in the third set ripened faster than did bananas in the first and second sets. This may be due to the fact that humidity and temperature were not constant in the laboratory during the period within which the experiments were performed. Since there is a reasonable correlation between categories and skin colours, it can be assumed that the categories established by PCA are consistent with there being different states of ripeness.

From left to right in figure 2(*a*), the categories appear ordered according to increasing ripeness. However, the inter-category boundaries are complex in shape (see figure 2(*b*)) and some patterns that belong to different categories have very similar scores. This effect may have been magnified by our experimental approach, which was focused on practicality of implementation. The occurrence of complex boundaries suggests that a nonlinear classification method is needed in order to obtain a good performance in terms of pattern recognition, rather than linear PCA.

The clusters corresponding to categories a, d and e exhibit a significant spread in a direction which is almost perpendicular to the direction of increasing ripeness (see figure 2(*a*)). Generally the samples in these categories with lower values of the scores in PC 2 belong to data sets 1 and 2, and the samples with higher values of the scores in PC 2 belong to the third data set. This suggests that the spread may be due to drift in the sensor response (the three data sets were gathered sequentially over a period of some two and a half months). The categories established by the PCA cluster together measurements gathered with different banana sets.

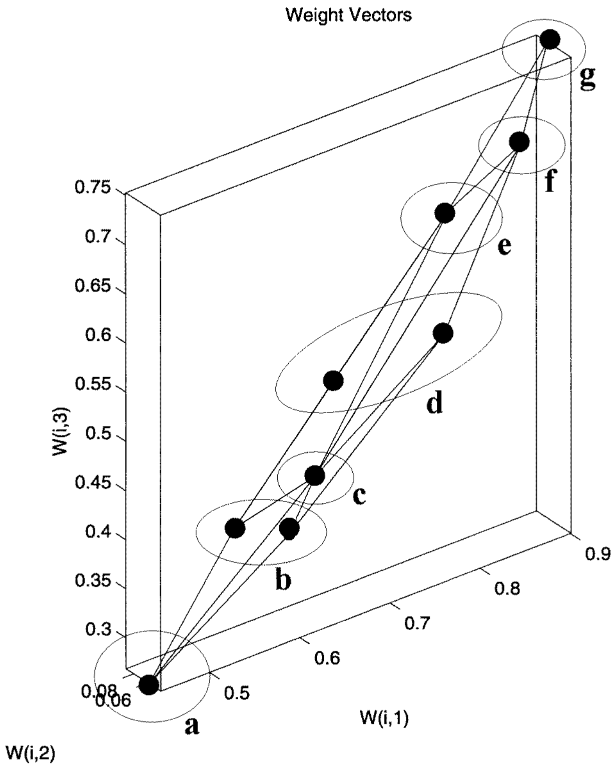


Figure 3. The SOM network after training with the three banana data sets together. Each neurone (or group of neurones) represents a state-of-ripeness category, which is consistent with the results obtained by applying PCA.

Furthermore, these categories exhibit a good consistency with banana-skin colour over the three banana sets. It can be concluded that the clustering is due to changes in the state of ripeness of fruit rather than to changes due to sensor drift.

A SOM network is able to accumulate statistical information about the environment without any supplementary information other than that provided by the sensors [26, 27]. A 3 × 3 square SOM network was created and trained using the three banana data sets together. Once the training process had finished, each sample was associated with one of the neurones in the network. Samples that are associated with the same neurone have been identified as belonging to the same category. Figure 3 shows the SOM after the training process. The neurones in figure 3 are labelled according to the category that they represent. The labels are consistent with those derived from the PCA.

Fuzzy *c*-means clustering (FCM) is a data clustering algorithm in which each data point belongs to a cluster according to its degree of membership [28]. FCM splits a data set into *c* fuzzy groups and finds a cluster centre for each group such that a dissimilarity function is minimized. In FCM, an initial estimate of the number of clusters is needed before the algorithm can be applied. The results obtained by applying FCM are shown in figure 4. *c* was selected to be equal to 7, to establish seven categories. The results show that category a was split into two clusters (a_1 and a_2) and categories f and g have merged into the same cluster.

In conclusion, it was found that the categories established and the number of patterns belonging to each category were very similar for the three techniques. Furthermore, these categories represent particular states during the ripening

Table 3. Correspondences between categories in table 1 and banana-skin colour.

Category	Set 1	Set 2	Set 3
a		Green–yellow	Green–yellow
b	Yellow, trace of green	Yellow, trace of green	
c		Completely yellow	
d	Yellow–green tips	Yellow flecked with brown	Yellow–green tips
e	Yellow flecked with brown		Yellow flecked with brown
f			Yellow flecked with brown
g			Yellow, blackening

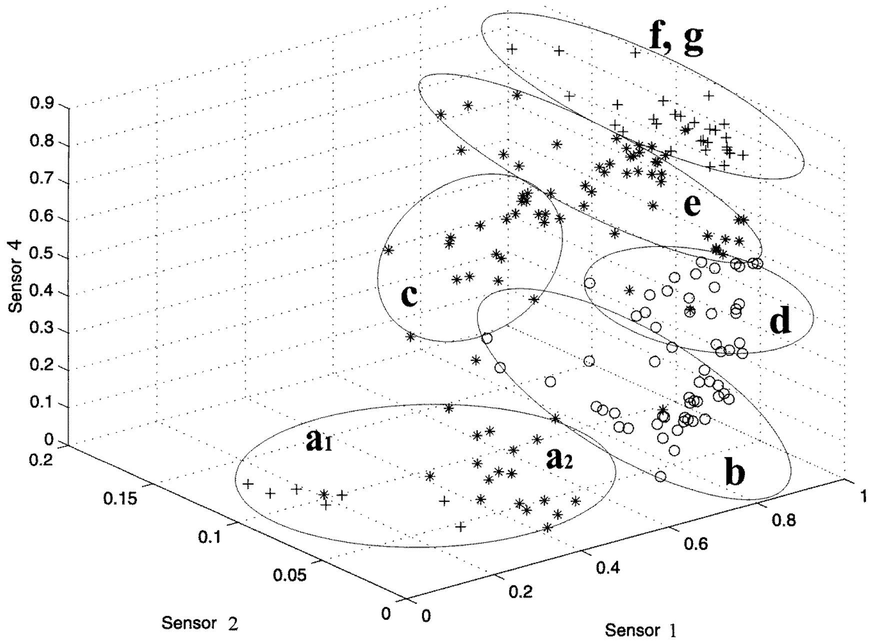


Figure 4. Results of the fuzzy c-means clustering (with c equal to 7). Each cluster represents a state-of-ripeness category. The results are consistent with those of PCA and SOM.

of fruit, according to standardized skin-colour charts [23]. Therefore, the seven categories established by PCA were used as the basis on which to apply supervised neural-network classifiers. The results of the analysis are presented and discussed in the following sub-sections.

3.3. Neural networks

The data sets were analysed using the Fuzzy ARTMAP, LVQ and the back-propagation trained multilayer perceptron paradigms. Fuzzy ARTMAP and LVQ carry out supervised learning, as does the back-propagation MLP. However, unlike back-propagation, Fuzzy ARTMAP and LVQ are self-organizing and self-stabilizing. Fuzzy ARTMAP has also been shown to be suitable for incremental learning [18].

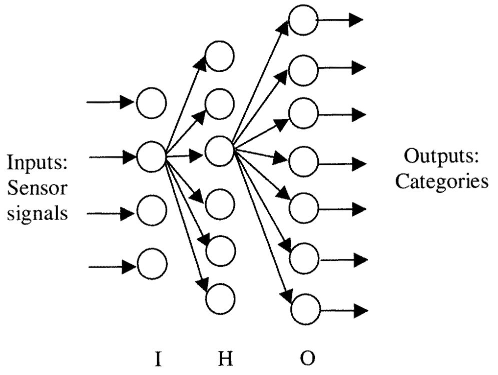
Fuzzy ARTMAP generally performs better at learning than does MLP, for example when there is an uneven number of patterns of each category in the training set. Other properties of Fuzzy ARTMAP which make this method promising for electronic nose systems, include the following facts [19, 20].

(i) Fuzzy ARTMAP can rapidly learn a rare event that predicts different consequences from those for a cloud of similar events within which it is embedded.
(ii) It contains a self-stabilizing memory that allows the accumulation of knowledge in response to a non-stationary environment, that is until the memory capacity is full (the memory capacity can be chosen arbitrarily large).
(iii) Fuzzy ARTMAP is able to adjust its scale of generalization automatically to match the morphological variability of the data. It conjointly maximizes generalization and minimizes prediction error using only information that is locally available under incremental learning conditions.
(iv) Individual recognition categories play the role of hidden units in the back-propagation model [20]. Unlike the back-propagation model, Fuzzy ARTMAP discovers, on its own, the number of categorical 'hidden units' that it needs for a specific problem. Conversely, in the cases of the MLP and the LVQ, the optimal numbers of neurones in the hidden layer and in the competitive layer, for a given application, are usually found by trial and error.

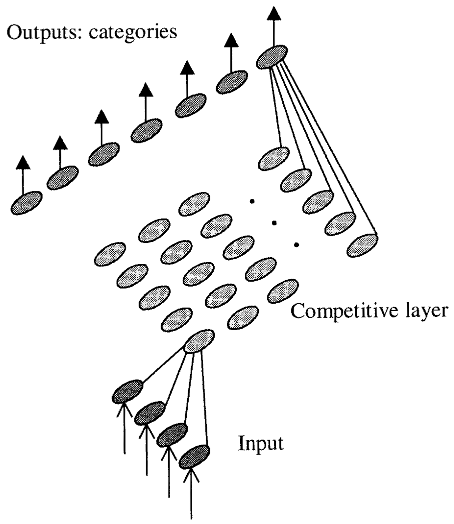
Figure 5 shows the topologies of Fuzzy ARTMAP, LVQ and MLP neural networks.

3.4. Analysis of the three banana data sets together

For this analysis, the data (175 response vectors from the three banana data sets) were divided into four folds. Three folds consisted of 131 training vectors and 44 test vectors, whereas the fourth one consisted of 132 training vectors



a)

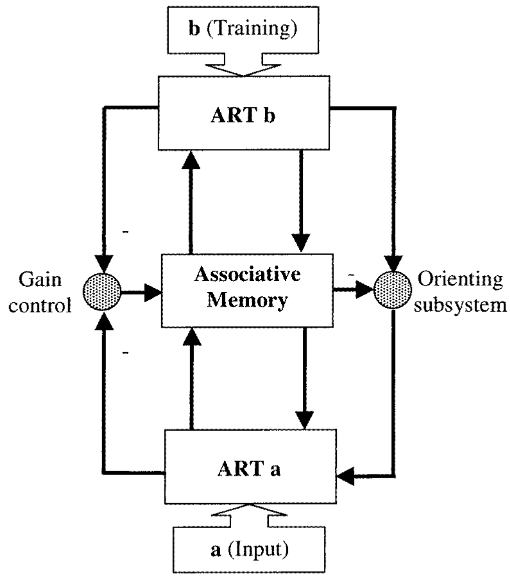


b)

Figure 5. Architectures of the networks studied. (*a*) MLP. For simplicity, only the connections from a neurone in the input layer (I) to the hidden layer (H) and a neurone in the hidden layer to the output (O) are shown. (*b*) LVQ. The inputs are connected to every neurone in the competitive layer. The neurones in the competitive layer are inter-connected in a rectangular mesh (not shown) and connected to one of the output neurones (categories). (*c*) Fuzzy ARTMAP network. It basically consists of two ART modules interconnected by an associative memory (or map field) and some internal control structures that regulate learning and information flow. Inhibitory paths are denoted by a minus sign; other paths are excitatory.

and 43 test vectors. Test vectors were selected at random without replacement. Since the number of vectors per category was uneven (e.g. eight replicates in category c and 35 replicates in category e), each category was represented both in training sets and in testing sets in proportion to its number of replicates. Table 4 shows the composition of the test folders.

The networks used had four inputs (one per sensor in the array) and seven outputs, since a one-of-seven code was used to code the seven different ripeness categories.



(*c*)

Figure 5. (Continued)

Table 4. Composition of the test folders, indicating the number of replicates per category.

Category	Fold 1	Fold 2	Fold 3	Fold 4
a	6	6	6	6
b	7	7	8	7
c	2	2	2	2
d	7	7	6	6
e	9	9	8	9
f	6	6	6	6
g	7	7	8	7
Total tested	44	44	44	43
Total for training	131	131	131	132

3.4.1. For Fuzzy ARTMAP. The network had four input and seven output neurones. The number of neurones in the category layer was set to 50 (a maximum of 50 nodes could be committed). However, the number of committed nodes after each training process was always below the maximum.

(i) The baseline vigilance was set to 0. This is the recommended value for the vigilance since it allows for very coarse categories and the match-tracking system will refine these categories only if necessary.
(ii) The recoding rate was set to 0.5. This value allows the established categories to be modified if there is a persistent attempt to do so (slow recoding).
(iii) The values of the choice parameter and error tolerance (if the output error is greater than the tolerance, then a reset signal is triggered) were varied. Values of 0.1 and 0.01, respectively, were found to be optimal for this experiment.

Fuzzy ARTMAP was able to correctly classify 90.3% of the response vectors. During the training process 25 nodes were committed.

3.4.2. For LVQ. The networks had four input and seven output neurones and a variable number of nodes in the competitive layer.

Table 5. Results of the state of ripeness classification with Fuzzy ARTMAP, LVQ and MLP, in terms of numbers of patterns correctly classified (numbers of patterns to be classified) and overall performance.

Method	Fold 1	Fold 2	Fold 3	Fold 4	Overall
Fuzzy ARTMAP	38 (44)	41 (44)	39 (44)	40 (43)	158 (175)
LVQ	40 (44)	41 (44)	39 (44)	41 (43)	161 (175)
MLP	36 (44)	36 (44)	34 (44)	40 (43)	146 (175)

(i) Initially the network was trained with a learning rate equal to 0.06 and the conscience factor was equal to 1. With this last option, the class winner is always moved towards the input vector (if it is in the right class) or moved away from the input vector (if it is in the wrong class).
(ii) In the second stage, once a 'relatively good' solution had been found, this solution was further refined by modifying the boundaries between zones where misclassifications occurred. The learning rate was set to 0.03.
(iii) The number of neurones in the competitive layer was changed. It was found that 35 was the optimal number of neurones for this application.

LVQ was able to correctly classify 92% of the response vectors.

3.4.3. For MLP. A back-propagation network (with a learning rate equal to 0.3 and a momentum term equal to 0.4) with four input, six hidden and seven output neurones was able to attain an 83.4% success rate in classification.

3.4.4. Summary. The back-propagation network required typically 40 000 training cycles, LVQ required 4000 iterations and Fuzzy ARTMAP required only 150 training iterations. Thus, the time necessary to train Fuzzy ARTMAP is more than an order of magnitude smaller than those for the other two. The numbers of patterns correctly classified are summarized in table 5. Table 6 shows the confusion matrices for Fuzzy ARTMAP, LVQ and MLP techniques.

3.5. Assessment of network performance

3.5.1. Student's t test. A t test was performed to assess whether Fuzzy ARTMAP and LVQ were performing significantly better than the MLP in terms of the total number of patterns correctly classified. The null hypothesis H_0 demonstrated that there was no significant difference between the mean numbers of patterns misclassified by the Fuzzy ARTMAP and LVQ and the MLP. The hypothesis H_0 was rejected at 5% significance level ($t = 2.45$ for Fuzzy ARTMAP; $t = 4.49$ for LVQ) because the critical t value at 5% significance level and three degrees of freedom is 2.35.

3.5.2. Sensitivity and specificity. Other metrics that are helpful in evaluating the performance of trained neural networks are the sensitivity and specificity [29]. Here the sensitivity of a neural network indicates the likelihood of a category being detected given that it is present. The specificity of a network is the likelihood that the absence of a category will be detected, given that it is absent. These performance metrics are defined as follows:

$$\text{Sensitivity} = \frac{n_{TP}}{n_{TP} + n_{FN}} \tag{1}$$

$$\text{Specificity} = \frac{n_{TN}}{n_{TN} + n_{FP}} \tag{2}$$

where number n subscripts are: TP (true positive) means that the recognition of a category coincides with the actual category according to the target standard; FN (false negative) means the non-recognition of a category according to the target standard; TN (true negative) is the non-recognition of a category in agreement with the target standard and FP (false positive) means that a category is recognized incorrectly. Ideally, both the sensitivity and the specificity should be equal to unity and their values must be in the range 0–1. Table 7 shows the values of these parameters for the trained Fuzzy ARTMAP, LVQ and MLP networks. Even though the specificity of all the networks are very similar (Fuzzy ARTMAP and LVQ are slightly better than the MLP), the sensitivities of Fuzzy ARTMAP and LVQ are clearly better than that of the MLP, especially for categories c and f.

3.6. Generalizability

The objective here was to study how effectively the networks were able to predict the state of ripeness of one of the banana data sets, when the set bad not been used for training the original neural network; in other words, to assess whether the networks were able, once they had been trained, to acquire the knowledge necessary to classify the state of ripeness of any new bananas. This was performed in two stages.

3.6.1. Stage 1. The patterns in sets 2 and 3 were used for training (126 patterns, in which the seven categories were represented) and the patterns in set 1 were used to test the networks, 49 patterns corresponding to classes b (21 patterns), d (eight patterns) and e (20 patterns). This was a difficult problem to solve because the three categories in the training set (b, d and e) were represented by only eight, 18 and 15 patterns, respectively. Thus, for classes b and e, there were more patterns in the test set than there were in the set used for training.

Under these conditions, Fuzzy ARTMAP attained a performance of 61.2% in the classification of the test patterns, LVQ gave a performance of 67.3% and the MLP provided a success rate of 57.1%. The MLP gave an especially poor performance in the identification of class e (only two out of 20 patterns were correctly classified). These results are shown in table 8.

3.6.2. Stage 2. The patterns in sets 1 and 3 were used to train the networks (140 patterns). All the categories, except category c, were represented in the training set. The patterns in set 2 were used for testing. Set 2 contained 35 patterns corresponding to the classes a (ten), b (eight), c (eight) and d (nine). In the training set, the categories a, b and d were represented by 14, 21 and 17 patterns, respectively.

Table 6. Confusion matrices of the banana-ripeness-classification problem with Fuzzy ARTMAP, [LVQ] and (MLP).

	Actual category						
Predicted	a	b	c	d	e	f	g
a	**23 [23] (24)**		(1)	1			
b		**27 [28] (28)**	(1)	1 (4)	1 [2] (1)		
c			**5 [7] (1)**				
d	1 [1]	2 [1] (1)	1 (2)	**23 [23] (22)**	(2)		1
e			2 [1] (2)	1 [2]	**33 [31] (28)**	2 [1]	
f			(1)		1 [2] (4)	**21 [21] (14)**	2 [1]
g						1 [2] (10)	**26 [28] (29)**

Table 7. Sensitivities and specificities of the trained Fuzzy ARTMAP, [LVQ] and (MLP) networks for every ripeness category.

Performance metric	Category						
	a	b	c	d	e	f	g
Sensitivity	0.958 [1.000] (1.000)	0.931 [0.966] (0.966)	0.625 [0.875] (0.125)	0.885 [0.885] (0.846)	0.943 [0.886] (0.800)	0.875 [0.875] (0.583)	0.897 [0.965] (1.000)
Specificity	0.993 [0.993] (0.993)	0.986 [0.986] (0.959)	1.000 [1.000] (1.000)	0.966 [0.986] (0.966)	0.964 [0.972] (0.986)	0.980 [0.980] (0.967)	0.993 [0.986] (0.932)

There were no patterns corresponding to category c in the training set.

This led to performances of 68.6% in the classification of the test patterns for Fuzzy ARTMAP and the MLP and 74.3% for the LVQ. The performance was increased to nearly 90% both for Fuzzy ARTMAP and for MLP and to 96% for LVQ if only the classification of the patterns in the test set that belonged to categories represented in the training set was considered (see table 8).

The results from stages 1 and 2 suggest the conclusion that all the networks have the ability to apply the previously learnt knowledge to the classification of new, unknown, sets of bananas. Even though the numbers of samples correctly classified by MLP, Fuzzy ARTMAP and LVQ were very similar, the two latter methods are preferred since they classified seven out of the eight patterns of class c as either class b or class d. These are the nearest categories, in terms of the state of ripeness, to the true category which were present in the training set. On the other hand, the MLP classified only three out of eight patterns of class c into the nearest available categories.

3.7. Incremental learning

The fact that Fuzzy ARTMAP is reported to be able to perform on-line learning without forgetting previously learnt patterns makes this approach very attractive compared with the standard MLP (standard MLPs are trained ‘off-line’) or the LVQ. This is important because, from a practical standpoint, in terms of an instrument being used in industry, the data set used to train the network may be increased during the development phase by adding new measurements. This would require the network to be re-trained using the complete data set in the case of the MLP. This can result in a time-consuming and costly process. The evaluation of the abilities of Fuzzy ARTMAP, LVQ and MLP in terms of incremental learning was performed separately as follows.

3.7.1. Fuzzy ARTMAP. In order to study the ability of Fuzzy ARTMAP to learn new knowledge without forgetting previous knowledge, a Fuzzy ARTMAP network was trained with the first banana data set (see table 2). Then the value of the recoding parameter β was set to a small value (0.1). In this way, a previously learnt category will be slowly re-coded and a significant change will occur only if a persistent attempt to modify the category is made during new learning. Afterwards, the network was further trained using only the patterns in the second banana data set (see table 2) and tested with the patterns from the first and second data sets. The aim was to check that Fuzzy ARTMAP was able to learn the patterns in the second data set without forgetting the patterns in the first one. Finally, the network was further trained using the patterns in the third data set and tested with the patterns in the first, second and third data sets.

It was found that the Fuzzy ARTMAP performed very well in this experiment. After each new learning process, the previous knowledge was not significantly degraded and the overall performance in learning remained almost constant in the range [94.6%, 98%]. These results are summarized in table 9.

3.7.2. MLP. The same process as that described above was repeated using a MLP. Although the network was able to learn the new patterns in each successive learning process, the previously learnt knowledge was significantly degraded (see table 9). In particular, when a previously learnt category is not present during a new learning cycle, the knowledge about this category is forgotten (e.g. category e during the learning of set 2 and categories b and c during the learning of set 3). A poor performance in the classification of the state of ripeness is provided by the MLP (63.4%) after the incremental learning.

3.7.3. LVQ. The process was repeated again using LVQ. The network was able to learn the new patterns in each

Table 8. Results of the study of the generalizability of Fuzzy ARTMAP, [LVQ] and (MLP) networks in ripeness classification, in terms of patterns correctly classified/numbers of patterns.

Training/tested sets	Category a	b	c	d	e	Overall performance	Restricted set performance[b]
2, 3/1		16/21 [15/21] (18/21)		8/8 [8/8] (8/8)	6/20 [10/20] (2/20)	30/49 [33/49] (28/49)	
1, 3/2	8/8 [8/8] (8/8)	7/9 [9/9] (7/9)	0/8[a] [0/8][a] (0/8)	9/10 [9/10] (9/10)		24/35 [26/25] (24/35)	24/27 [26/27] (24/27)

[a] Fuzzy ARTMAP and LVQ classify seven out of eight patterns of class c as class b or class d.
[b] Restricted-set performance when the results in classification of class c, which is not in the training set, are not considered.

Table 9. Incremental learning on the three banana data sets with Fuzzy ARTMAP, [LVQ] and (MLP). For Fuzzy ARTMAP, the recoding rate was fixed to $\beta = 0.1$. Data are expressed as numbers of patterns correctly classified/total numbers of patterns in the category.

Learning/tested sets	Category a	b	c	d	e	f	g	Performance (%)
1/1		21/21 [20/21] (20/21)		7/8 [8/8] (8/8)	20/20 [19/20] (18/20)			98.0 [95.9] (93.8)
2/1 and 2	10/10 [9/10] (9/10)	28/29 [22/29] (27/29)	7/8 [8/8] (8/8)	16/17 [15/17] (15/17)	19/20 [18/20] (0/20)			94.6 [85.7] [70.2]
3/1, 2 and 3	24/24 [21/24] (23/24)	28/29 [28/29] (0/29)	7/8 [8/8] (0/8)	25/26 [14/26] (25/26)	33/35 [23/35] (15/35)	23/24 [23/24] (21/24)	29/29 [23/29] (27/29)	96.0 [80.0] (63.4)

successive learning process and, in some cases, was able to keep the previously learnt knowledge. When no patterns of a previously learnt category were present in a new learning process, the knowledge of the category was not degraded (see category c in table 9). On the other hand, when some new patterns of a previously learnt category were present, the weights of the neurones corresponding to this category are recoded. This results in a degradation of the overall performance (see categories d and e in table 9).

3.7.4. Summary. In overall terms the performance of Fuzzy ARTMAP (an average of 96.0%) compares very favourably with those of the MLP and LVQ, as shown in table 9.

3.8. Noise rejection

To analyse how the different networks were able to deal with uncertainty, which is a key factor in any measurement system, various levels of Gaussian noise (1, 5, 10 and 20%) were added to the sensor responses and the classification process was repeated. Figure 6 shows the accuracies in classification attained by the three networks in the presence of noise. It was found that noise affected all the networks in a similar way. However, it was found that, in this application, the performances of Fuzzy ARTMAP and the MLP were less resilient to low values of noise than was the performance of LVQ (see figure 6).

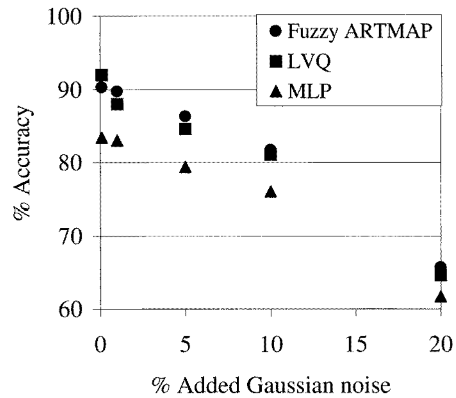


Figure 6. Accuracies in classification of the state of ripeness of bananas attained by Fuzzy ARTMAP, LVQ and MLP in the presence of 0, 1, 5, 10 and 20% of added Gaussian noise.

Finally, the incremental learning process was repeated when 1, 5, 10 and 20% of Gaussian noise had been added to the sensor signals. It was found that Fuzzy ARTMAP performed well, even in the presence of noise, and that it clearly outperformed both LVQ and MLP. These results are summarized in table 10.

4. Discussion

From table 6, we conclude that there are two main reasons for the classification performances of the Fuzzy ARTMAP and LVQ techniques being superior to that of MLP.

Table 10. Incremental learning on the three banana data sets with Fuzzy ARTMAP, [LVQ] and (MLP). Data are percentage accuracies in classification when 1, 5, 10 and 20% of Gaussian noise was added to the sensor responses.

	Gaussian noise added (%)			
Learning/tested sets	1	5	10	20
1/1	100 [98.0] (91.8)	98.0 [95.9] (81.6)	100 [95.9] (79.2)	100.0 [89.8] (77.6)
2/1 and 2	97.6 [84.5] (69.0)	95.2 [85.7] (67.9)	91.7 [83.3] (65.3)	97.6 [78.5] (63.1)
3/1, 2 and 3	98.3 [80.6] (64.5)	96.0 [76.6] (50.8)	92.0 [74.8] (46.8)	90.9 [72.0] (44.0)

(i) Fuzzy ARTMAP and LVQ are able to adapt themselves to the distribution in databases for which the number of patterns per category is uneven. Thus, whereas Fuzzy ARTMAP and LVQ are able to classify most of the patterns corresponding to class c, the MLP classifies only one pattern correctly. MLP is less able to adapt to the uneven distribution of samples.

(ii) Fuzzy ARTMAP is able to adjust its scale of generalization (increasing the vigilance parameter) to match the morphological variability of the patterns and thus achieves a better performance than does MLP in the separation of the classes e, f and g (see figure 2(*b*)).

(iii) In the LVQ algorithm, when a relatively good solution has been found, this solution can be further refined by modifying the boundaries between zones where misclassifications occur.

Furthermore, there is another property of Fuzzy ARTMAP and LVQ networks that may explain their performance.

(iv) Because they are self-organizing, the number of training patterns and the number of training iterations needed to match, or exceed, the performance of MLP is lower.

5. Conclusions

Odour patterns from three different sets of bananas were gathered with an electronic nose instrument. First, seven different ripeness categories were identified with the help of principal-components analysis, SOM and Fuzzy clustering of the sensor responses. This observation is in good agreement with the seven known classes of banana ripeness from skin-colour analysis. Then, Fuzzy ARTMAP and LVQ neural networks were applied to the classification of the state of ripeness of bananas. An accuracy of 90.3% was reached in the classification using Fuzzy ARTMAP (92% was obtained using LVQ). It was found that these performances compared favourably with that achieved with back-propagation trained MLPs (83.4%). Furthermore, the sensitivities and specificities of the trained Fuzzy ARTMAP and LVQ networks to the established categories were higher than those of the MLP.

Finally, the training time of Fuzzy ARTMAP was found to be typically more than an order of magnitude less than those for back-propagation MLP and LVQ.

The generalizability of the trained networks to the prediction of the state of ripeness of new, unknown, bananas was investigated. It was found that the networks had a good performance, providing 90% accuracy in the classification of patterns belonging to previously trained categories. If a new category for which the network had not been trained occurred during testing, Fuzzy ARTMAP and LVQ associated these patterns with classes which were the nearest to the actual state of ripeness and which were already known.

Fuzzy ARTMAP was shown to perform slightly better than LVQ in the presence of noise. Finally, Fuzzy ARTMAPs superior ability to perform incremental learning without forgetting previously learnt patterns was demonstrated in this application, when it significantly outperformed LVQ and MLP, even in the presence of added noise.

All these characteristics make the Fuzzy ARTMAP network very attractive for pattern classification in the context of real instruments. However, further work to assess the long-term reliability of the system is needed. In particular, the effects of sensor drift on its accuracy should be investigated. The use of a commercial headspace autosampler would help to ameliorate any inherent limitations in the sampling process and thus improve system performance, resulting in a system that is more suitable for commercial applications to continuous monitoring of fruit ripeness. Nevertheless, we believe that a Fuzzy ARTMAP-based electronic nose provides an attractive means of identifying the ripeness of commercial fruit.

Acknowledgments

The help provided by Mr Frank Courtney with the experimental set-up and the development of LabVIEW programs for data acquisition and storage by Mr James Covington and Mr Reza Tamadoni are gratefully acknowledged. Dr E Llobet gratefully acknowledges financial support (grant PR1997-0138) from the Spanish Ministry of Education and Culture during his stay at the University of Warwick.

References

[1] Horton B D 1992 Ripening patterns within a peach as indicated by force and soluble solids concentration *J. Am. Soc. Horticultural Sci.* **117** 784–7

[2] Reyes M U, Paull R E, Williamson M R and Gautz L D 1996 Ripeness determination of ‘Solo’ papaya (*Carica papaya* L.) by impact force *Appl. Engng. Agriculture* **12** 703–8

[3] Kalt W and McDonald J E 1996 Chemical composition of lowbush blueberry cultivars *J. Am. Soc. Horticultural Sci.* **121** 142–6

[4] Vangdal E 1980 Threshold values of soluble solids in fruit determined for the fresh fruit market *Acta Agriculturae Scand.* **30** 445–8

[5] McGlasson W B 1983 Ethylene and fruit ripening *Hortscience* **18** 626

[6] Pare J R J and Belanger J M R (eds) 1997 *Instrumental Techniques in Food Analysis* (Amsterdam: Elsevier)

[7] Cho S I, Krutz G W, Gibson H G and Haghighi K 1990 Magnet console design of an NMR-based sensor to detect ripeness of fruit *Trans. ASAE* **33** 1043–50

[8] Wai W K, Stroshine R L and Krutz G W 1995 Modified Hahn echo pulse sequence for proton magnetic resonance (1H-MR) measurements of percent soluble solids of fruits *Trans. ASAE* **38** 849–55

[9] Ward G and Nussinovitch A 1996 Peel gloss as a potential indicator of banana ripeness *Lebensmittel-Wiss.-Technol.* **29** 289–94

[10] Galili N, Shmulevich I and Benichou N 1998 Acoustic testing of avocado for fruit ripeness evaluation *Trans. ASAE* **41** 399–407

[11] Benady M, Simon J E and Miles G E 1995 Fruit ripeness determination by electronic sensing of aromatic volatiles *Trans. ASAE* **38** 251–7

[12] Simon J E, Hetzroni A, Bordelon B, Miles G E and Charles D J 1996 Electronic sensing of aromatic volatiles for quality sorting of blueberries *J. Food Sci.* **61** 967–70

[13] Gardner J W, Hines E L and Tang H C 1992 Detection of vapours and odours from a multisensor array using pattern-recognition techniques: 2. Artificial neural networks *Sensors Actuators* B **9** 9–15

[14] Gardner J W, Shurmer H V and Tan T T 1992 Application of an electronic nose to the discrimination of coffees *Sensors Actuators* B **6** 71–5

[15] Schweizer-Berberich P M, Vaihinger S and Göpel W 1994 Characterization of food freshness with sensor arrays *Sensors Actuators* B **18** 282–90

[16] Winquist F, Hornsten E G, Sundgren H and Lündstrom I 1993 Performance of an electronic nose for quality estimation of ground meat *Meas. Sci. Technol.* **4** 1943–50

[17] Llobet E, Hines E L, Gardner J W, Bartlett P N and Mottram T T 1999 Fuzzy ARTMAP based electronic nose data analysis *Sensors Actuators* B submitted

[18] Carpenter G, Grossberg S, Markuzon N, Reynolds J and Rosen D 1992 Fuzzy ARTMAP: a neural network architecture for incremental supervised learning of multidimensional maps *IEEE Trans. Neural Networks* **3** 698–713

[19] Carpenter G, Grossberg S and Reynolds J 1995 A Fuzzy ARTMAP nonparametric probability estimator for nonstationary pattern recognition problems *IEEE Trans. Neural Networks* **6** 1330–6

[20] Carpenter G A and Grossberg S 1994 Fuzzy ARTMAP: a synthesis of neural networks and fuzzy logic for supervised categorization and nonstationary prediction *Fuzzy Sets, Neural Networks, and Soft Computing* ed R R Yager and L A Zadeh (New York: Van Nostrand Reinhold) pp 126–65

[21] Kohonen T 1989 *Self-Organization and Associative Memory* 3rd edn (Heidelberg: Springer) pp 185–209

[22] National Instruments Corporation 1998 *LabVIEW Version 5 Data Manual* (Austin, TX: National Instruments)

[23] Von Loesecke H W 1950 *Bananas: Chemistry, Physiology, Technology* (New York: Interscience)

[24] NeuralWare Inc 1995 *NeuralWorks Professional II/Plus, Reference Guide* (Pittsburgh: NeuralWare)

[25] Gardner J W 1991 Detection of vapours and odours from a multisensor array using pattern recognition. Part 1: principal components and cluster analyses *Sensors Actuators* B **4** 108–16

[26] Kohonen T 1982 Self organised formation of topologically correct feature maps *Biol. Cybernet.* **43** 59–69

[27] Dinatale C, Davide F and D'Amico A 1995 Pattern-recognition in gas sensing: well-stated techniques and advances *Sensors Actuators* B **23** 111–18

[28] Jang J S R, Sun C T and Mizutani E 1997 *Neuro-Fuzzy and Soft Computing. A Computational Approach to Learning and Machine Intelligence* (Upper Saddle River, NJ: Prentice Hall) pp 423–33

[29] Eberhart R C, Dobbins R W and Hutton L V 1990 *Neural Network PC Tools: A Practical Guide* (New York: Academic) pp 161–76