Sensors and Actuators B 61 (1999) 183–190

SENSORS AND ACTUATORS B CHEMICAL

www.elsevier.nl/locate/sensorb

Fuzzy ARTMAP based electronic nose data analysis

Eduard Llobet [a], Evor L. Hines [b,*], Julian W. Gardner [b], Philip N. Bartlett [c], Toby T. Mottram [d]

[a] *Departament d'Enginyeria Electrònica, Universitat Rovira i Virgili, Autovia de Salou s/n, 43006 Tarragona, Spain*
[b] *Electrical and Electronic Engineering Division, School of Engineering, University of Warwick, Coventry CV4 7AL, UK*
[c] *Department of Chemistry, University of Southampton, Southampton SO17 1BJ, UK*
[d] *Silsoe Research Institute, West Park, Silsoe, Bedford MK45 4HS, UK*

Received 5 November 1998; received in revised form 26 July 1999; accepted 26 August 1999

Abstract

The Fuzzy ARTMAP neural network is a supervised pattern recognition method based on fuzzy adaptive resonance theory (ART). It is a promising method since Fuzzy ARTMAP is able to carry out on-line learning without forgetting previously learnt patterns (stable learning), it can recode previously learnt categories (adaptive to changes in the environment) and is self-organising. This paper presents the application of Fuzzy ARTMAP to odour discrimination with electronic nose (EN) instruments. EN data from three different datasets, alcohol, coffee and cow's breath (in order of complexity) were classified using Fuzzy ARTMAP. The accuracy of the method was 100% with alcohol, 97% with coffee and 79%, respectively. Fuzzy ARTMAP outperforms the best accuracy so far obtained using the back-propagation trained multilayer perceptron (MLP) (100%, 81% and 68%, respectively). The MLP being by far the most popular neural network method in both the field of EN instruments and elsewhere. These results, in the case of alcohol and coffee, are better than those obtained using self-organising maps, constructive algorithms and other ART techniques. Furthermore, the time necessary to train Fuzzy ARTMAP was typically one order of magnitude faster than back-propagation. The results show that this technique is very promising for developing intelligent EN equipment, in terms of its possibility for on-line learning, generalisation ability and ability to deal with uncertainty (in terms of measurement accuracy, noise rejection, etc.).

Keywords: Fuzzy ARTMAP; Neural network; Electronic nose; Odour analysis; Intelligent system; Multilayer perceptron

1. Introduction

Electronic noses (ENs) are instruments, comprising an array of chemical sensors with partial selectivity and an appropriate pattern recognition system (PARC), that are capable of recognising simple or complex odours, in an analogue to the human nose [1]. A considerable number of pattern recognition methods have been used to analyse the response produced by sensor arrays. The nature of these techniques can typically be classified using terms such as *parametric* or *non-parametric* and *supervised* or *unsupervised* [2].

A parametric technique assumes that the sensor data can be described by a probability density function (PDF) that a posteriori defines its spread of values. In most cases, the assumption made is that data are normally distributed with a known mean and variance. Non-parametric methods do not assume any PDF for the sensor data and thus apply more generally, typical examples being supervised or unsupervised neural networks.

In a supervised PARC method, a set of known odours are systematically introduced to the EN, which then classifies them according to known descriptors (classes) held in a 'knowledge base'. Then, in a second stage, an unknown odour is tested against the knowledge base and the predicted membership class is given. Unsupervised PARC methods do not need a priori knowledge about class membership because they cluster the different classes using only the response (input) vectors.

Of all the PARC techniques, the back-propagation multilayer perceptron (MLP) neural network, which is a non-linear, non-parametric and supervised method, has been the most widely used. The MLP has been shown to perform well in a variety of applications [3–6]. Standard MLP has a number of drawbacks including the fact that it has a limited capability to compensate for undesirable characteristics of the sensor system (e.g., changes in the

* Corresponding author. Tel.: +44-1203-523-246; fax: +44-1203-418-922; e-mail: e.l.hines@warwick.ac.uk

0925-4005/99/$ - see front matter
PII: S0925-4005(99)00288-9

sensor response due to temperature and moisture variations and drift), learns very slowly, etc. Standard MLPs are trained 'off-line' and are unable to adapt autonomously, in real time, to changes in the environment. Furthermore, the dataset used to train the network may be increased during the development phase by adding new measurements and this would require the network to be re-trained using the complete dataset. This can result in a time consuming and costly process.

One possible way of improving the existing commercial EN instruments is to apply pattern recognition techniques that emulate, more closely, the way that the human olfactive system is understood to work. In particular, a human brain is able to learn many new events without necessarily forgetting events that occurred in the past. If we want an intelligent system capable of adapting 'on-line' to changes in the environment, the system should be able to deal with the so-called 'stability–plasticity dilemma'. That is the system should be designed to have some degree of plasticity to learn new events in a continuous manner and, should be stable enough to preserve its previous knowledge, and to prevent new events destroying the memories of prior learning. Adaptive resonance theory (ART) networks were designed to address the stability–plasticity dilemma, are capable of real-time learning and classification [7], and have been applied with some success to EN data [8].

In this paper we examine the application of Fuzzy ARTMAP [9,10], which is a supervised variant of Fuzzy ART, to process EN data. There are several properties that make Fuzzy ARTMAP a promising pattern recognition method for EN systems.

- *Exhibits fast learning of rare events:* Many traditional learning strategies use forms of slow learning that average over the occurrence of similar events. Fuzzy ARTMAP can rapidly learn a rare event that predicts different consequences than a cloud of similar events in which it is embedded.
- *Suitable for non-stationary environments:* In a non-stationary environment, traditional algorithms tend to loose the memory of old, but still useful knowledge. Fuzzy ARTMAP contains a self-stabilising memory that allows for the accumulation of knowledge in response to a non-stationary environment, until the memory capacity is full (memory can be chosen arbitrarily large).
- *Ability to adjust the scale of generalisation:* In many environments some information may be coarsely defined, whereas other information may be precisely characterised. Fuzzy ARTMAP is able to automatically adjust its scale of generalisation to match the morphological variability of the data. It conjointly maximises generalisation and minimises predictive error using only information that is locally available under incremental learning conditions.
- *Ability to learn many-to-one relationships:* Many-to-one learning combines categorisation of many exemplars into one category, and labelling of many categories with the same name. Individual recognition categories play the role of hidden units in the back-propagation model [11]. Unlike the back-propagation model, Fuzzy ARTMAP discovers, on its own, the number of categorical 'hidden units' that it needs for a specific problem.
- *Ability to deal with uncertainty:* A key element in any measurement system is uncertainty and the fuzzy approach is one way of dealing with it.

In Section 2, a brief review of ART and Fuzzy ARTMAP is given. This is followed by a discussion of the application of Fuzzy ARTMAP to three EN datasets (alcohol discrimination, coffee discrimination and diagnose of ketosis in dairy cattle). In the three cases, the results are discussed and compared with those obtained when optimised MLP networks were used.

2. Adaptive resonance theory

There are two general classes of ART networks: ART1, ART2 and ART3. While ART1 is for classifying binary input patterns, ART2 and ART3 are for analogue patterns. Because Fuzzy ART and Fuzzy ARTMAP are generalisations of ART1 a brief review of this architecture will be given. The reader can find a useful introduction to ART in Ref. [12] and is referred to Ref. [13] for further details.

2.1. ART1

ART1 is formed by two major subsystems: the attentional subsystem and the orienting subsystem. The architecture of the ART1 network is shown in Fig. 1. Two interconnected layers of neurones F1 and F2, which are fully connected both bottom-up and top-down, comprise the attentional subsystem. The links between F1 and F2 are called adaptive filters where the weights represent the long-term memory (LTM) because they remain in the network for an extended period. The application of a single

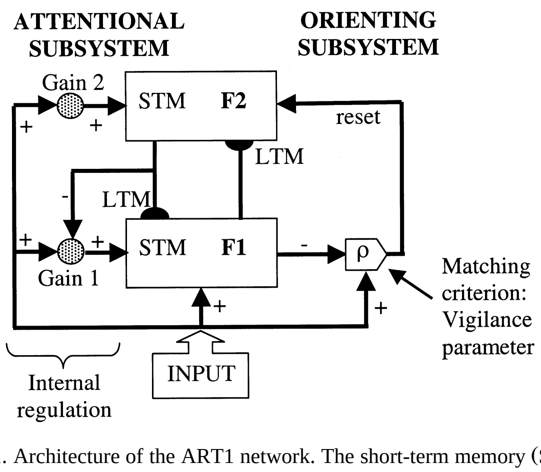

Fig. 1. Architecture of the ART1 network. The short-term memory (STM) patterns are stored in F1 and F2 layers. The long-term memory (LTM) of the system is represented by the adaptive weights of both bottom-up and top-down connections. Excitatory paths are denoted by a plus sign; inhibitory paths are denoted by a minus sign.

input vector leads to patterns of neural activity in both layers F1 and F2. These patterns are known as the short-term memory (STM). The activity in F2 nodes reinforces the activity in F1 nodes due to top-down connections. The interchange of bottom-up and top-down information leads to a resonance in neural activity. As a result, critical features in F1 are reinforced, and have the greatest activity. The orienting subsystem is responsible for generating a reset signal to F2 when the bottom-up input pattern and top-down template pattern mismatch at F1, according to a vigilance criterion. In other words, once it has detected that the input pattern is novel, the orienting subsystem must prevent the previously organised category neurones in F2 from learning this pattern (via a reset signal). Otherwise, the category will become increasingly non-specific. When a mismatch is detected, the network adapts its structure by immediately storing the novelty in additional weights. The vigilance criterion is set by the value of the vigilance parameter. A high value of the vigilance parameter means than only a slight mismatch will be tolerated before a reset signal is emitted. On the other hand, a small value (low vigilance) means that large mismatches will be tolerated. After the resonance check, if a pattern match is detected according to the vigilance parameter, the network changes the weights of the winning node. The ART network stores a weighted part of the present input vector in the LTM, just as any other neural network does. A summarised mathematical model is as follows:

Let: $\boldsymbol{I} = (I_1, \ldots, I_M)$

be the input vector with M components where $I_i = 1$ or 0 (ART1 requires binary inputs);

$\boldsymbol{X} = (X_1, \ldots, X_M)$

be the vector of F1 nodes where $X_i = 1$ or 0.

Let w_{Ji} be the top-down weight from the winning node J in the F2 layer (F2 is a competitive layer) to a node i in the F1 layer, and let z_{iJ} be the corresponding bottom-up weight. Assuming fast learning (e.g., weight update equations reach their asymptotic values before the next training vector is presented), the weights take the values:

$$w_{ji} = \begin{cases} 1 & \text{if } i \in \boldsymbol{X} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$z_{iJ} = \begin{cases} \dfrac{L}{L - 1 + |\boldsymbol{X}|} & \text{if } i \in \boldsymbol{X} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $|\boldsymbol{X}|$ is the cardinality of $\boldsymbol{X}$, and L is a network parameter ($L > 1$).

If $\boldsymbol{W}_J$ is the vector of top-down weights from the winning neurone in F2 and $\boldsymbol{Z}_J$ is the vector of bottom-up weights, Eqs. (1) and (2) can be rewritten as follows:

$$\boldsymbol{W}_J = \boldsymbol{X} \tag{3}$$

$$\boldsymbol{Z}_J = \frac{L\boldsymbol{X}}{L - 1 + |\boldsymbol{X}|} = \frac{L\boldsymbol{W}_J}{L - 1 + |\boldsymbol{W}_J|}. \tag{4}$$

A given node j in the F2 layer gets the following input from the F1 layer:

$$t_j = \sum_{i-1}^{i=M} z_{ij} I_i. \tag{5}$$

Using Eq. (4), Eq. (5) can be rewritten in terms of the top-down weights:

$$t_j = \left(\frac{L}{L - 1 + |\boldsymbol{W}_j|} \right) \sum_{i=1}^{i=M} w_{ji} I_i. \tag{6}$$

The summation term in Eq. (6) is the number of common 1s the vectors $\boldsymbol{W}_j$ and $\boldsymbol{I}$ have in corresponding positions. Then, Eq. (6) can be rewritten as:

$$t_j = \left(\frac{L|\boldsymbol{I} \cap \boldsymbol{W}_j|}{L - 1 + |\boldsymbol{W}_j|} \right) \tag{7}$$

where $|\boldsymbol{I} \cap \boldsymbol{W}_j|$ is the cardinality of the intersection set of $\boldsymbol{W}_j$ and $\boldsymbol{I}$.

Category choice is made by selecting the neurone in F2 with the maximum value for t_j. Thus, the Choice Function can be defined as:

$$T_j(\boldsymbol{I}) = \frac{|\boldsymbol{I} \cap \boldsymbol{W}_j|}{L - 1 + |\boldsymbol{W}_j|}. \tag{8}$$

A given node i in the F1 layer is active only if both top-down weight $\boldsymbol{W}_{Ji}$ from the winning F2 node and the input to node i are non-zero:

$$X = |\boldsymbol{I} \cap \boldsymbol{W}_J|. \tag{9}$$

The winning node J in the F2 layer is reset by the orienting subsystem if:

$$\frac{|\boldsymbol{I} \cap \boldsymbol{W}_J|}{|\boldsymbol{I}|} < \rho \tag{10}$$

where ρ is the vigilance parameter. Once a node is reset, it remains inactive for the duration of the trial.

2.2. Fuzzy ART

Fuzzy ART is a generalisation of Eqs. (8)–(10). The generalisation is achieved by using fuzzy set theory operations rather than binary set theory operations. An overview of fuzzy set theory and fuzzy neural networks can be found in Ref. [14].

Input nodes can take values between 0 and 1 (analogue patterns). The wining neurone in the F2 layer (e.g., wining category) is the one for which the choice function attains its maximum value. The new choice function, derived from Eq. (8) is as follows:

$$T_j(\boldsymbol{I}) = \frac{|\boldsymbol{I} \wedge \boldsymbol{W}_j|}{L - 1 + |\boldsymbol{W}_j|} \tag{11}$$

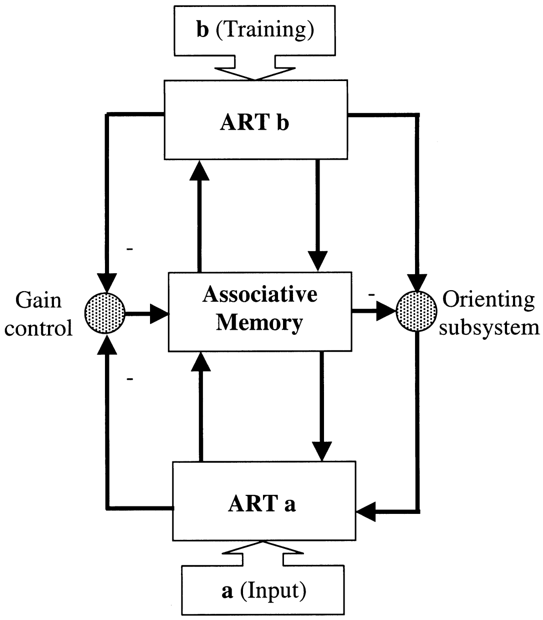

Fig. 2. Architecture of the Fuzzy ARTMAP network. It consists basically of two ART modules interconnected by an associative memory (or map field) and some internal control structures that regulate learning and information flow. Inhibitory paths are denoted by a minus sign; other paths are excitatory.

where $\boldsymbol{I} \wedge \boldsymbol{W}_j$ is the equivalent operation in fuzzy set theory of the intersection of $\boldsymbol{W}_j$ and $\boldsymbol{I}$ in standard set theory. Thus:

$$\boldsymbol{I} \wedge \boldsymbol{W}_j = \left(\min\left(I_1, w_{j1}\right), \ldots, \min\left(I_i, w_{ji}\right), \ldots, \min\left(I_M, w_{jM}\right)\right) \tag{12}$$

and the expression for the cardinality is as follows:

$$|X| = \sum_{i=1}^{M} |X_i|. \tag{13}$$

The winning category is reset by the vigilance subsystem if:

$$\frac{|\boldsymbol{I} \wedge \boldsymbol{W}_J|}{|\boldsymbol{I}|} < \rho. \tag{14}$$

When a node in F2 is first committed (e.g., because novelty has been detected), a fast commit equation is used:

$$\boldsymbol{W}_J^{(\text{new})} = \boldsymbol{I}. \tag{15}$$

Once a node has been committed, a slow recode equation is used to change the weights towards the spatial position of the actual input vector:

$$\boldsymbol{W}_J^{(\text{new})} = \left(\boldsymbol{I} \wedge \boldsymbol{W}_J^{(\text{old})}\right) + (1-\beta)\boldsymbol{W}_J^{(\text{old})} \quad 0 \le \beta \le 1. \tag{16}$$

The slow recode of committed nodes prevents noisy data from erroneously recoding them. Small values of β cause the system to base its results on a long-term average of its experience, while values of β near 1 allow adaptation to a rapidly changing environment.

2.3. Fuzzy ARTMAP

Fuzzy ARTMAP in its most general form, includes two Fuzzy ART modules ($\mathrm{ART_a}$ and $\mathrm{ART_b}$) whose F2 layers are linked by an inter-ART associative memory referred to as ‘match tracking system’. The Fuzzy ARTMAP architecture is shown in Fig. 2. The reader is referred to Ref. [9] and to Ref. [11] for a detailed review of this architecture. During supervised learning $\mathrm{ART_a}$ receives a stream of input patterns $\{a^M\}$ and $\mathrm{ART_b}$ also receives a stream of patterns $\{b^M\}$, where b^M is the correct prediction given a^M. When a prediction by $\mathrm{ART_a}$ is not confirmed by $\mathrm{ART_b}$, inhibition of the inter-ART associative memory activates a match tracking process. This increases $\mathrm{ART_a}$ vigilance by the minimum amount needed for the system either to activate an $\mathrm{ART_a}$ category that matches the $\mathrm{ART_b}$ category or to learn a new $\mathrm{ART_a}$ category. Fuzzy ARTMAP carries out supervised learning like back-propagation. But unlike back-propagation, Fuzzy ARTMAP is self-organising, self-stabilising and suitable for real-time learning. Section 3 deals with the application of Fuzzy ARTMAP to three different EN datasets.

3. Application of fuzzy ARTMAP to EN data

3.1. Datasets

Three EN datasets were analysed. The first two datasets were collected using an EN that consisted of an array of 12 commercial metal oxide semiconductor gas sensors (Figaro Engineering, Japan). The third dataset was collected using an EN consisting of six metal oxide sensors (Figaro and FIS, Japan). Table 1 is a list of the sensors used.

The first dataset consisted of samples taken from the headspace of simple alcohols: 5 ppm in air of methanol, ethanol, butan-1-ol, propan-2-ol and 2-methyl-butanol. The process was repeated to provide eight identical samples of each of the five classes, making a set of 40 input vectors

Table 1
List of the metal oxide semiconductor sensors used

Sensor	Manufacturer	Alcohol/ coffee analysis	Cow’s breath analysis
TGS 812	Figaro	✓	
TGS 711	Figaro	✓	
TGS 813	Figaro	✓ (2)	
TGS 814	Figaro	✓	
TGS 824	Figaro	✓	
TGS 815	Figaro	✓	
TGS 882	Figaro	✓	
TGS 816	Figaro	✓ (2)	
TGS 817	Figaro	✓	
TGS 880	Figaro		✓
NFIN43	FIS		✓
NFI1813	FIS		✓
TGS 825	Figaro	✓	✓
STAQ1A	FIS		✓
TGS 822	Figaro		✓

[3]. The second dataset consisted of samples taken from the headspace of three different roasted coffees. Two of the coffees were of the same roasting level but different blends, while the third was a different roast level but of the same blend as one of the first two coffees. The process was repeated to provide 30 samples of the first two coffees and 29 of the third, making a total of 89 input vectors [15]. The third dataset consisted of samples of the breath of cows. The aim was to diagnose ketosis in cows, a disease that is characterised by the accumulation of ketone bodies in an organism. The cows were classified into three different categories: healthy, ketotic and sub-clinical ketotic. These three categories were previously established by analysing blood samples from the cows studied. The measurement of the breath samples was repeated to provide 136 measurements (75 of ketotic, 46 of healthy and 15 of sub-clinical cows) [16].

3.2. Data pre-processing

The choice of the data pre-processing algorithm has been shown to affect the performance of the pattern recognition stage. The choice should depend upon the underlying sensor principle and the nature of interfering signals. When metal oxide semiconductor sensors are used, it has been shown that the fractional change in conductance $(G^{\mathrm{odour}} - G^{\mathrm{air}})/G^{\mathrm{air}}$ helps both to linearise the sensor response with concentration and to reduce its temperature sensitivity, this has been shown to improve the performance of neural network based techniques [5]. For the alcohol and coffee data sets the fractional change in sensor conductance was used. For the cow data set, in addition to the fractional conductance change, we experimented with the rise time of the conductance defined as $t_{\mathrm{r}} = 0.6(t|_{G=G^{\mathrm{odour}}} - t|_{G=G^{\mathrm{air}}})$ as well because it has been shown that, in some cases, information from sensor dynamics can help odour recognition [6]. In all the cases studied, the data were also normalised to set their range to [0, 1] when Fuzzy ARTMAP networks were used.

The Fuzzy ARTMAP networks were simulated using the NeuralWorks Professional II/Plus software from NeuralWare, USA [17].

It is important to note that in order to facilitate the comparison between the Fuzzy ARTMAP results and those previously published, relating to the use of MLPs, we have employed the same neural network evaluation strategy. Thus, the datasets can be regarded as a benchmark for the Fuzzy ARTMAP approach.

3.3. Results

3.3.1. Alcohol data

The alcohol dataset was divided into eight test folds containing one measurement of each alcohol (five measurements per fold). Given a test fold, the neural network was trained using the remaining 35 measurements and then tested with the five test vectors, one from each class. The network had 12 inputs (one per sensor in the array) and five outputs, since a 1-of-5 code was used to code the five different alcohols. For Fuzzy ARTMAP, 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 only refine these categories if necessary. The recode rate (β in Eq. (16)) was set to 0.5. This value allows the established categories to be modified if there is a persistent attempt to do so (slow recode). The values of the choice parameter ($L-1$ in Eq. (11)) and error tolerance (if the output error is greater than the tolerance, then a reset signal is triggered) were varied. It was found that 0.1 and 0.01, respectively, were optimal for this experiment. It was found that Fuzzy ARTMAP was able to accurately classify 100% of the alcohol samples. During the training process six nodes (that is categories) were committed. A back-propagation network with 12 input, seven hidden and five output neurones [3] was also able to reach a 100% success rate in alcohol classification. However, while the back-propagation network required typically 10,000 training cycles to obtain these results, the Fuzzy ARTMAP required only 50 training iterations. Thus, the time necessary to train the network was dramatically reduced. Other types of ART based neural networks [8], and techniques for automating the design of MLP, such as constructive algorithms [18] and genetic algorithms [19] were explored previously. However, the results did not show any significant advantage compared with those of Fuzzy ARTMAP.

3.3.2. Coffee data

The analysis of coffees is a more difficult problem to solve than that of the alcohols because the headspace of coffee forms a complex odour. The coffee dataset was divided into five test folds containing 18 measurements each (six measurements per coffee category). Given a test fold, the neural network was trained using the remaining 71 measurements and then tested with the 18 test vectors, six from each class. The network has 12 inputs and three outputs since a 1-of-3 code was used to code the three different coffees. Table 2 shows the results of the classification of coffee data sets. Only three samples were misclassified, leading to an accuracy of 97% in coffee classifi-

Table 2
Results of analysing the coffee data set using Fuzzy ARTMAP after 80 training iterations
71 patterns were used for training and 18 patterns tested in each fold (17 in the 5th fold). Baseline vigilance = 0, recode rate = 0.5, choice parameter = 0.1 and error tolerance = 0.001.

Fold	Patterns misclassified	Accuracy (%)
1	0	100
2	1	94
3	1	94
4	0	100
5	1	94
Total	3	97

cation. The number of training iterations was typically equal to 80 and the number of committed nodes after training was 15. This result compares favourably to the 81% accuracy obtained with a back-propagation neural network [4]. Other approaches that were investigated previously such as different ART based neural networks [8] or self-organising maps [20] did not show significant advantages compared with back-propagation.

3.3.3. Cattle diagnosis

This was a much more difficult problem to solve because the breath samples were collected in the field (actually a barn!) over a 2-week period. This implies that the background air was subject to considerable environmental variations in terms of ambient temperature, humidity, etc. The cow dataset was divided into 4 test folds containing 34 measurements each (19 corresponding to ketotic cows, 12 of healthy cows and three of sub-clinical cows). Given a test fold, the neural network was trained using the remaining 102 measurements and then tested with the 34 test vectors. The network has six inputs and three outputs since a 1-of-3 code was used to code the three different diagnoses. Table 3 shows the results of the classification when the rise time of the sensor conductance was input to the Fuzzy ARTMAP network. The number of training iterations was 200 and the number of committed nodes was 36. The use of the dynamic information led to a slightly better accuracy (79%) compared with a 75% accuracy when the fractional change in conductance was used. These results also compare favourably to the 56% accuracy obtained with a four-input, five-hidden and three-output back-propagation neural network or to the 68% accuracy obtained with a four-input, five-hidden and one-output network, which is significantly higher than the a priory probability of 56% for ill cows [16].

Table 4 summarises the performances of Fuzzy ARTMAP and back-propagation networks for the analysis of the three datasets.

Table 3
Results of analysing the cow's breath data set using Fuzzy ARTMAP after 200 training iterations

102 patterns were used for training and 34 patterns tested in each fold. Baseline vigilance = 0, recode rate = 0.5, choice parameter = 0.1 and error tolerance = 0.01.

Fold	Patterns misclassified (three categories)[a]	Accuracy (%)	Patterns misclassified (two categories)[b]	Accuracy (%)
1	5	85	3	91
2	7	79	4	88
3	9	74	7	80
4	8	77	5	85
Total	29	79	19	86

[a] Healthy, sub-clinical and ketotic.
[b] Healthy and ill (this last category includes sub-clinical and ketotic).

Table 4
Comparison of the performance of Fuzzy ARTMAP and back-propagation neural networks on alcohol, coffee and cow's breath data

Algorithm	Accuracy (%)			Training iterations		
	Alcohol	Coffee	Cow	Alcohol	Coffee	Cow
Fuzzy ARTMAP	100	97	79	50	80	200
Back-propagation	100	81	68	10,000	10,000	1000

3.4. Significance test

A t-test was performed to assess if Fuzzy ARTMAP was performing significantly better than the MLP in terms of the total number of patterns correctly classified for coffees and cattle diagnosis. The null hypothesis H_0 demonstrated that there was no significant difference between the mean number of patterns misclassified by the Fuzzy ARTMAP and the MLP. The null hypothesis H_0 was clearly rejected at 5% significance level ($t = 7.48$ for coffees and $t = 3.22$ for cattle diagnosis) because the critical t values at 5% significance level for 4 and 3 df are less at 2.13 and 2.35, respectively.

4. Summary and conclusions

Fuzzy ARTMAP neural networks have been applied to the classification of alcohol, coffee and cow's breath patterns, gathered with EN instruments. An accuracy of 100% on alcohol, 97% on coffee and 79% on cow's breath was achieved. It was found that these performances compared favourably with those achieved previously with back-propagation trained MLPs (100%, 81% and 68%, respectively) and other techniques. Furthermore, the training time of Fuzzy ARTMAP was found to be typically an order of magnitude faster than back-propagation. There are several properties of Fuzzy ARTMAP networks that may explain these promising results.

- Because they are self-organising, the number of training patterns and the number of training iterations needed to match, or exceed, the performance of MLP is lower. Thus, calibration and training times in the development of an instrument may be significantly reduced.
- The number of committed nodes in the F2 layer has a similar meaning to the hidden neurones in the MLP approach. While in Fuzzy ARTMAP the nodes are committed automatically, the MLP requires the number of hidden neurones to be optimised by means of a trial–error procedure. We have explored techniques for automating the design of MLP but none of them have been found to have such a significant effect as the use of Fuzzy ARTMAP.
- Fuzzy ARTMAP is able to learn new patterns without forgetting older ones, as long as the memory of the

system is not full. Furthermore, it is able to slowly recode previous categories, to be able to adjust for example for the long-term drift of the sensors. This may explain why it outperforms MLP in the analysis of the cow's breath samples. These samples were measured over a period of 2 weeks and were therefore likely to be subjected to significant variations in the temperature and moisture conditions.

· Fuzzy set processing is one way of attempting to deal with uncertainty, which is a key element in any measurement system, and this is an inherent feature in Fuzzy ARTMAP. For example, the system can deal with noise in the input patterns via the error tolerance parameter, which is used by the match tracking system for triggering a category reset, and this can be optimised for every specific application.

The fact that Fuzzy ARTMAP is able to perform on-line learning without forgetting previous learnt patterns and the fact that it can deal with uncertainty in the data makes this approach very promising for the development of the next generation of intelligent EN systems.

Acknowledgements

Alcohol and Coffee datasets were gathered by the Warwick-Southampton Electronic Nose Group. Cattle and veterinary support and analysis were provided by Paul Dobbelaar, Faculty of Veterinary Medicine, University of Utrecht. Dr. R. Elliot-Martin and C.J. Allen gathered the experimental data and operated the equipment. The Cattle diagnosis experiments were funded by the Biotechnology and Biological Sciences Research Council (No. A01037). Dr. E. Llobet gratefully acknowledges financial support (grant ref. PR1997-0138) from the Spanish Ministry of Education and Culture during his stay at the University of Warwick.

References

[1] J.W. Gardner, P.N. Bartlett, A brief history of electronic noses, Sensors and Actuators B 18/19 (1994) 211–220.

[2] J.W. Gardner, E.L. Hines, Pattern analysis techniques, in: E. Kress-Rogers (Ed.), Handbook of Biosensors and Electronic Noses. Medicine, Food and the Environment, CRC Press, Boca Raton, 1997, pp. 633–652.

[3] J.W. Gardner, E.L. Hines, M. Wilkinson, The application of artificial neural networks in an electronic nose, Meas. Sci. Technol. 1 (1990) 446–451.

[4] J.W. Gardner, P.N. Bartlett, Pattern recognition in odour sensing, in: P.N. Bartlett, J.W. Gardner (Eds.), Sensors and Sensory Systems for an Electronic Nose, Kluwer, Dordrecht, 1991.

[5] J.W. Gardner, E.L. Hines, H. Tang, Detection of vapours and odours from a multisensor array using pattern recognition: Part 2. Artificial neural networks, Sensor and Actuators B 9 (1992) 9.

[6] E. Llobet, J. Brezmes, X. Vilanova, J.E. Sueiras, X. Correig, Qualitative and quantitative analysis of volatile organic compounds using transient and steady-state responses of a thick-film tin oxide gas sensor array, Sensors and Actuators B 41 (1997) 13–21.

[7] S. Grossberg, Adaptive pattern classification and universal recoding: II. Feedback, expectation, olfaction and illusions, Biol. Cybernet. 23 (1976) 187–202.

[8] J.W. Gardner, E.L. Hines, C. Pang, Detection of vapours and odours from a multisensor array using pattern recognition: self-organising adaptive resonance techniques, Measurement and Control 29 (1996) 172–177.

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

[10] G. Carpenter, S. Grossberg, J. Reynolds, A fuzzy ARTMAP nonparametric probability estimator for nonstationary pattern recognition problems, IEEE Trans. Neural Networks 6 (1995) 1330–1336.

[11] G.A. Carpenter, S. Grossberg, Fuzzy ARTMAP: a synthesis of neural networks and fuzzy logic for supervised categorization and nonstationary prediction, in: R.R. Yager, L.A. Zadeh (Eds.), Fuzzy Sets, Neural Networks, and Soft Computing, Van Nostrand-Reinhold, New York, 1994, pp. 126–165.

[12] D. Wienke, L. Buydens, Adaptive resonance theory based neural networks — the 'ART' of real-time pattern recognition in chemical process monitoring?, Trends Anal. Chem. 14 (1995) 398–406.

[13] G.A. Carpenter, S. Grossberg, A massively parallel architecture for a self-organizing neural pattern-recognition machine, Computer Vision Graphics and Image Processing 37 (1987) 116–165.

[14] C.T. Lin, C.S.G. Lee, Neural Fuzzy Systems, Prentice Hall, NJ, 1996.

[15] J.W. Gardner, H.V. Shurmer, T.T. Tan, Application of an electronic nose to the discrimination of coffees, Sensors and Actuators B 6 (1992) 71–75.

[16] J.W. Gardner, P.N. Bartlett, E.L. Hines, F. Molinier, T.T. Mottram, Prediction of the health of dairy cattle from breath samples using a neural network with parametric model of the dynamic response of an array of semiconducting gas sensors, Proceedings of IEE: Science, Measurement and Technology 146 (1999) 102–106.

[17] NeuralWare, NeuralWorks Professional II/Plus, Reference Guide, Pittsburg, USA, 1995.

[18] E.L. Hines, C.C. Gianna, J.W. Gardner, Neural network based electronic nose using constructive algorithms, in: P.J.G. Lisboa, M.J. Taylor (Eds.), Techniques and Applications of Neural Networks, Ellis Horwood, New York, 1993, pp. 137–154.

[19] A.A. Fekadu, E.L. Hines, J.W. Gardner, Genetic algorithm design of neural network based electronic nose, in: R.C.R. Albrecht, N.C. Steele (Eds.), Artificial Neural Nets and Genetic Algorithms, Springer-Verlag, New York, 1993, pp. 691–698.

[20] E.L. Hines, J.W. Gardner, C.E.R. Potter, Olfactory feature maps for an electronic nose, Measurement and Control 30 (1997) 262–268.

Eduard Llobet graduated in telecommunication engineering from the Universitat Politècnica de Catalunya (UPC) (Barcelona, Spain) in 1991, and received his PhD in 1997 from the same university. His thesis focused on selectivity enhancement of metal oxide chemical sensors via the study of sensor dynamics. He has been a Lecturer in the Electronic Engineering Department at the Universitat Rovira i Virgili (Tarragona, Spain) since 1992, and is currently a visiting fellow at the School of Engineering, University of Warwick (Coventry, UK). His main areas of interest are in the fabrication, and modelling, of semiconductor chemical sensors and in the application of intelligent systems to complex odour analysis. Dr. Llobet is a member of the Institute of Electrical and Electronic Engineers.

Evor L. Hines is a Senior Lecturer in the School of Engineering at Warwick. He has been involved in the publication of about 100 technical papers in areas including radar meteorology, computer vision, robotics and Intelligent Systems Engineering. Most of these papers have been concerned with the application of Intelligent Systems.

Julian W. Gardner (BSc PhD DSc CEng FIEE MIEEE) is Chair in Electronic Engineering in the School of Engineering at Warwick. His research interests are microsensors, microsystems technology, electronic noses, intelligent sensors and multivariate data processing methods. He has previously spent 5 years in industry working first at AEA Technology and later at Molins Advanced Technology Unit on instrumentation. At Molins he developed a novel opto-electronic sensor that has been packaged in the UK and US for implementation on high speed packaging machinery. In 1989 he received the Esso Centenary Education Award sponsored by the Royal Society and Fellowship of Engineering to pursue his research interests. He has published over 200 technical papers and is an author of four books in Nanotechnology (1991), Electronic Noses (1992 and 1999) and Microsensors (1994). He held an Alexander von Humboldt Fellowship in Germany in 1994. He currently heads the Sensors Research Laboratory in the Centre for Nanotechnology and Microengineering at Warwick University and is Chairman of the IEE Professional Group J1 committee on Measurements and Instrumentation.