ELSEVIER

Sensors and Actuators B 33 (1996) 77–82

SENSORS ACTUATORS B CHEMICAL

Fuzzy neural networks for gas sensing

Dimitrios Vlachos, John Avaritsiotis

National Technical University of Athens, Electrical Engineering Department, Division of Computer Engineering, 157–73 Zographou, Athens, Greece

Accepted 14 May 1996

Abstract

The implementation of a fuzzy neural network with an array of tin oxide based gas sensors for both quantitative and qualitative gas sensing is demonstrated. The architecture of the system is presented with some references to the general theory of fuzzy sets and fuzzy calculus. Experimental results are presented in the case of gas identification between CO, ethanol and methane and in the case of CO detection in different levels of relative humidity. Finally the effect of network parameters to the functionality of the system is discussed, especially in the case of functions evaluating the fuzzy AND and OR operations.

Keywords: Gas sensors; Fuzzy neural network

1. Introduction

Pattern recognition techniques are widely used in gas sensing to extract information when multi component analysis is performed. Generally, the performance of such techniques is affected by four major problems, namely: feature extraction, classification, sensor stability and gas mixture affirmation. Feature extraction is the procedure that transforms the sensor output to qualitative information. Since this transformation is not a linear one on sensor output, suitable numerical tricks must be engaged. Classification, when considered, is the recognition and identification of patterns produced by feature extraction. This procedure can be an easy one when in the representation space, the subspaces referred to the classes produced by the feature extraction are connected. Then, Bayesian surfaces can be used to separate the classes. But normally this is not the situation in gas sensing applications. Sensor stability may dramatically decrease the lifetime of a pattern recognition system. This is caused from the fact that the knowledge built in a system is very closely related to the specific sensor behavior during the learning procedure. Finally, gas mixture affirmation is the procedure of determining the exact gas concentration of the components of gas mixtures. Since the number of component combinations are infinite, only a few of them can be included in the learning procedure.

In the case of classical neural network approaches, the system is very sensitive to the set of data used in the learning procedure. This is due to two reasons: the first is that the transition surfaces cannot be provided (i.e. are randomly selected) and the second is that the system is forced to answer with a 'yes' or 'no'. Moreover, these reasons are responsible for the poor stability of such systems. The use of hybrid networks, including a back propagation network to perform feature extraction and a self organized map to perform classification has been found to improve system characteristics [1–3]. But again frequently calibration of such systems is needed to ensure system efficiency.

A different approach to this problem can be obtained using fuzzy sets and fuzzy calculus. The main advantage of this approach is the flexibility of fuzzy logic which includes all the intermediates answers between the Boolean 'yes' and 'no'. An easy and efficient implementation of fuzzy logic is the fuzzy processor. As in logic processor, neurons that perform the AND and OR operation in their inputs are used. Moreover, neurons are connected through weighted links, which again perform AND and OR operation with the weight. During training, these weights are calculated using a Back Propagation algorithm. Finally, the results of the operation of a fuzzy neural network are presented.

* Corresponding author. Tel.: +30 1 7484621; fax +30 1 7784578.

Elsevier Science S.A.
PII S0925-4005(96)01917-X

2. Fuzzy sets and fuzzy calculus

2.1. Fuzzy sets

In a normal set A whose elements are from a space X, the characteristic function of the set A is defined as

$$\chi_A : X \rightarrow \{0,1\}, \qquad \chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 2, & \text{if } x \notin A \end{cases} \quad (1)$$

A fuzzy set is a generalization of a normal set by forcing the characteristic function to take not only two distinct values (0 and 1) but every real value in [0,1]. Therefore, an element x of X can be 'less' or 'more' a member of the set A. All normal set operations (like intersection, union etc.) can be generalized and applied to fuzzy sets. For example, in normal sets the intersection is defined as

$$\chi_{A\cap B}(x) = \min(\chi_A(x), \chi_B(x)) \quad 92)$$

The same is true for fuzzy sets also. For example, consider two fuzzy sets A and B, and an element $x \in X$. If $\chi_A = 0.6$ and $\chi_B = 0.2$, then $\chi_{A\cap B} = 0.2$. Moreover, this definition of fuzzy sets allows for the introduction of linguistic variables. For example, consider a set A and an element x_0 with $\chi_A(x_0) = 0.81$.Then, by taking the square root of the membership function of every element, we can produce a set B with $\chi_B(x_0) = 0.9$. This is an implementation of the variable '*more*'.

2.2. s and t norms

By a **t** norm we mean a function of two arguments:

$$\mathbf{t}:[0,1] \times [0,1] \rightarrow [0,1]$$

such that
(i) for $x \leq y$ and $w \leq z$ then $x\,\mathbf{t}\,y \leq y\,\mathbf{t}\,z$ (non-decreasing)
(ii) $x\,\mathbf{t}\,y = y\,\mathbf{t}\,x$ (commutative)
(iii) $(x\,\mathbf{t}\,y)\,\mathbf{t}\,z = x\,\mathbf{t}\,(y\,\mathbf{t}\,z)$ (associative) and
(iv) $x\,\mathbf{t}\,0 = 0$ and $x\,\mathbf{t}\,1 = x$.

By an s norm we mean a function of two arguments

$$\mathrm{s}:[0,1] \times [0,1] \rightarrow [0,1]$$

such that:
(I) for $x \leq y$ and $w \leq z$ then $x\,\mathrm{s}\,y \leq y\,\mathrm{s}\,z$ (non-decreasing)
(ii) $x\,\mathrm{s}\,y = y\,\mathrm{s}\,x$ (commutative)
(iii) $(x\,\mathrm{s}\,y)\,\mathrm{s}\,z = x\,\mathrm{s}\,(y\,\mathrm{s}\,z)$ (associative) and
(iv) $x\,\mathrm{s}\,0 = x$ and $x\,\mathrm{s}\,1 = 1$.

It is clear that s and **t** norms are the generalizations of the logical OR and AND operations. This is derived from condition (iv) of the definition of s and **t** norms. Thus, any s (**t**) norm if applied to the set {0,1} is exactly the OR (AND) operation. Moreover, for every **t** norm one can define an s norm by: $x\,\mathrm{s}\,y = 1 - (1-x)\,\mathbf{t}\,(1-y)$ (De Morgan law for fuzzy sets).All s and **t** norms satisfy the inequalities (3) and (4):

$$\max(x, y) \leq x\,\mathrm{s}\,y \leq 1 \quad (3)$$

$$0 \leq x\,\mathbf{t}\,y \leq \min(x, y) \quad (4)$$

These inequalities can be verified very easily using the conditions (iv) of the definition and the fact that s and **t** norms are non-decreasing functions. In Fig. 1, the effect of s and **t** norms is shown in two fuzzy sets A and B.

2.3. Artificial intelligence and fuzzy sets

In the field of artificial intelligence, all the information-processing procedures are symbolic; the symbols are manipulated via a collection of specific syntax rules. At the same time, the numerical information is essential ignored. In the numerical techniques, all the objects are plain numbers. They strive for precision, while their knowledge representation capabilities are non-existent. The artificial intelligence schemes of knowledge representation are powerful and diversified; however they do not cope with any numerical information. Fuzzy sets are placed in between. As a collection of objects, they are described by symbols (like for example *small* or *large*). Simultaneously, these symbols have a certain semantics attached to them that is conveyed by numerical characteristics described by numerical grade of membership. The level of precision as contrasted to generality can easily be modified by changing the number of linguistic labels and modifying their parameters. This enhances an ability to implement principles of incompatibility and efficiently express the tradeoffs existing between achievable levels of precision and relevancy. As an example, consider the case of a simple controller. If the input in the controller (usually the error) is 'big', then it is almost certain that a 'big' action has to be taken by the controller. The above sentence, although it provides a rule that is part of our knowledge and experience, it does not cope with any numerical information because (i) it does not provide a way to decide if a specific input is 'big' and (ii) it does not specify which action is 'big'. Fuzzy calculus solve this problem by assigning fuzzy sets to the linguistic 'big'.

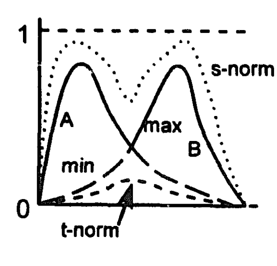

Fig. 1. Effect of s and t norms on fuzzy sets.

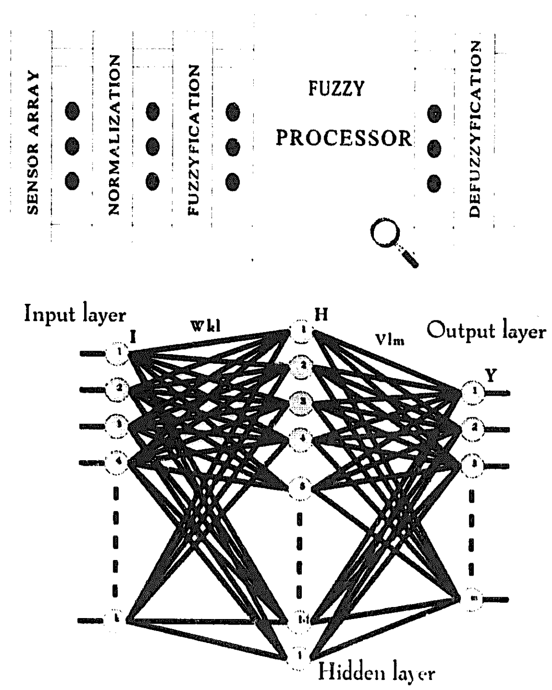

Fig. 2. Fuzzy neural network topology.

Then, it is possible to decide (i) if a specific input belongs to the fuzzy set 'big input' and (ii) which action belongs to the fuzzy set 'big action'. Consequently, the rule 'if (big input) then (big action)' can be applied in every controller. The only thing that changes from application to application is the exact implementation of the fuzzy sets 'big input' and 'big output'. It is obvious that these sets have to be determined by the application specific aspects.

3. Fuzzy neural network topology

As it is shown in Fig. 2, a general fuzzy neural network topology is buildt by

(i) a sensor array as in classical neural networks
(ii) a normalization block which is necessary in the case of fuzzy calculations. Thus, each input value is restricted to the set [0,1].
(iii) a fuzzification block. This is usually obtained by 'softening' the input. A common implementation of fuzzification is

$$\chi_{\mathrm{FUZZ(A)}}(x)=\begin{cases}\left(\dfrac{\chi_{\mathrm{A}}(x)}{2}\right)^{1/2}, & \chi_{\mathrm{A}}(x)<\frac{1}{2}\\ 1-\left(\dfrac{1-\chi_{\mathrm{A}}(x)}{2}\right)^{1/2}, & \text{otherwise}\end{cases} \tag{5}$$

An example of this operation to a fuzzy set is shown in Fig. 3a. The output of this block also produces the complement of each variable so the processor can use it as in usual logic processors.

(iv) a fuzzy processor which is built by a combination of neurons. As shown in Fig. 2, there are three levels of neurons. The input level neurons (I_k) perform an equality test to their input. This test compares the input pattern to a predefined one which can be the zero-grade-air response of the sensor array.

The hidden level nodes (H_l) perform an AND operation to their input. The operation is evaluated with a t-norm. The output level nodes perform an OR operation to their input. The operation is evaluated with an s-norm. The connections from the input layer node k to the hidden layer node l perform an OR operation with the output of input-layer-node k and the connection weight w_{kl}. Finally, the connections between the hidden layer node l and the output layer node m perform an AND operation with the output of hidden-layer-node l and the connection weight v_{lm}.

(v) a defuzzification block in order to produce a considerable output. Defuzzification or contrast intensification (INT), affects an original fuzzy set by suppressing grades of membership lower than 1/2 and elevates values greater than this threshold. An evaluation of INT is given by

$$\chi_{\mathrm{INT(A)}}(x)=\begin{cases}2^{p-1}\chi_{\mathrm{A}}^{p}(x), & \text{if}\ \chi_{\mathrm{A}}<\frac{1}{2}\\ 1-2^{p-1}(1-\xi_{\mathrm{A}}(x))^{p}, & \text{otherwise}\end{cases} \tag{6}$$

By controlling the parameter p, one can make the intensification operation more radical. An example of this operation to a fuzzy set is shown in Fig. 3b.

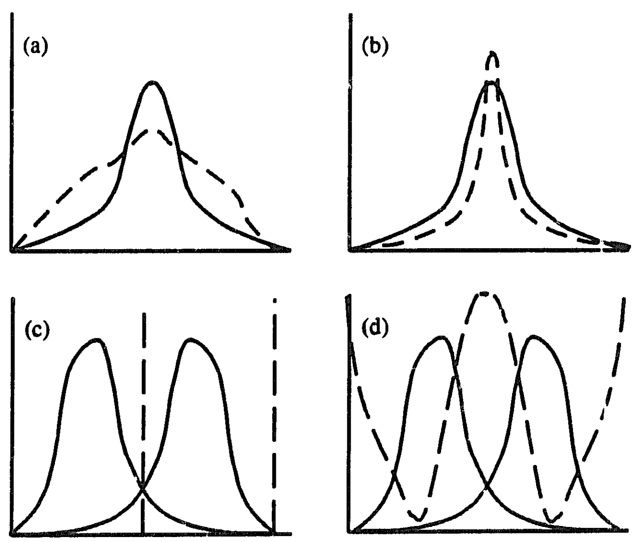

Fig. 3. Effect of different operations on fuzzy sets (continuous lines). (a) fuzzification (discontinuous line), (b) contrast intensification (discontinuous line), © the result of Boolean equality test (discontinuous line) in two fuzzy sets (continuous line) and (d) the result of fuzzy equality test (discontinuous line) in two fuzzy sets (continuous lines).

4. Learning procedure

The training of the fuzzy neural network is evaluated with a back propagation algorithm. Each time a valid input pattern is presented at the input, the output is calculated and compared with the desired one. The connection weights then are modified in order to minimize the error. Let D_m be the desired output and Y_m the calculated one. Then the error E is given by:

$$E=\sum_{m=1}^{M}(Y_m = D_m)=\sum_{m=1}^{M}(1-|Y_m - D_m|) \tag{7}$$

Note here that the equality test in the above equation returns all the intermediates answers between 'yes' and 'no'. In Fig. 3c and d, the results of the Boolean equality-test and the fuzzy equality-test are shown. Moreover, since the dependence of the error function deviation on the connection weights incorporates both the functions that evaluate the s and t norms and their derivatives, the network performance is a substitute for both of them.

5. Experimental results

Two types of experiments were carried out; the first one was for the identification of the active gas between CO, ethanol and methane, and the second one was for the elimination of the humidity interference in CO sensing. In the first experiment six tin oxide based gas sensors were used as the input sensor array. The input layer consists of thirteen nodes (two nodes for each sensor to include the complement and one for bias). The hidden layer consists of nineteen nodes, one of them was used for bias. The network produces three outputs (three neurons at the output layer) one for each gas of interest. The learning data consists of the sensor response in different concentrations of CO, ethanol and methane. Only one gas was present at each time. The network output is shown in Fig. 4.

In the second experiment, again six tin oxide based gas sensors were used as the input sensor array. The hidden layer consists of 17 nodes, one of them was used for bias. The network produces one output (one neuron at the output layer), showing the CO concentration. The learning data consists of the sensor response in different CO concentrations at different levels of relative humidity. The output of the system is shown in Fig. 5.

6. Discussion

6.1. Network performance

In the case of active gas identification, the network quickly learns to distinguish the three gases of interest. Moreover, as it is shown in Fig. 4, the output of the network depends on gas concentration also. This can be considered as an improvement over the classical neural networks, which are trained to answer only with a simple 'yes' or 'no'. Each network output is educated to respond in a specific gas. Ethanol and CO affect the other outputs since the sensor response to these gases is considerably more significant than the sensor response to methane.

In the case of water vapor interference elimination on CO sensing, the one network output was educated to give

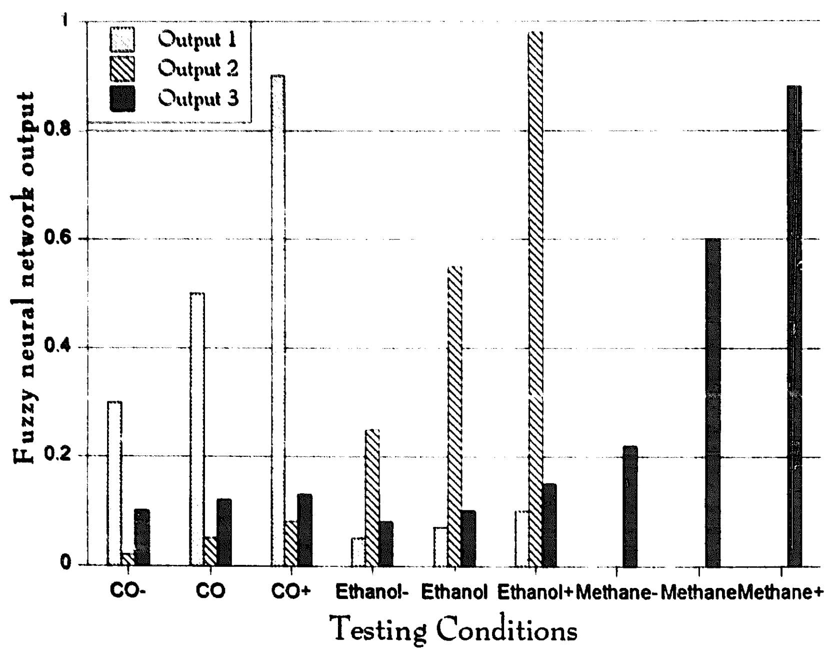

Fig. 4. Fuzzy neural network output for different testing conditions. CO– is 100 ppm CO, CO is 500 ppm CO, CO+ is 1000 ppm CO, Ethanol– is 10 ppm ethanol, Ethanol is 20 ppm ethanol, Ethanol+ is 30 ppm ethanol, Methane– is 500 ppm methane, Methane is 1000 ppm methane and Methane+ is 1500 ppm methane.

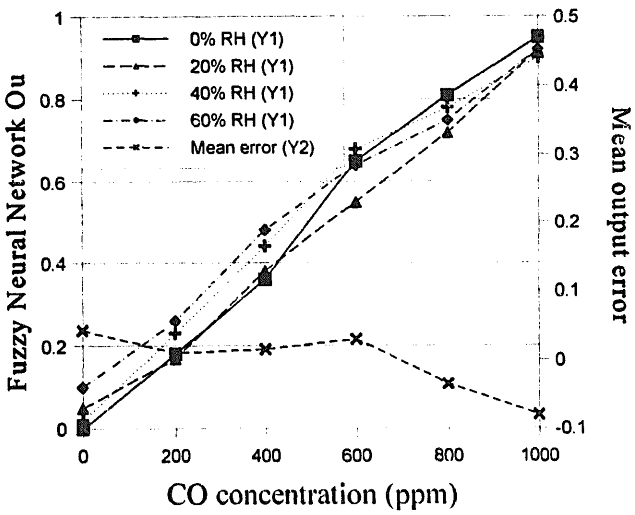

Fig. 5. Fuzzy neural network output and output error for different CO concentrations in various levels of relative humidity.

a linear measure of CO concentration, assigning the zero output to 0 ppm CO concentration and the value 1 to 1000 ppm CO concentration, neglecting the level of relative humidity. Here the learning stage was very long since the network had to suspend the water vapor effect. As shown in Fig. 5, the error in network response (i.e. the difference between the actual value and the desired value) remains close to zero for small concentrations of CO. For concentrations above 800 ppm, the error increases and this can be assigned to the non-linear dependence of tin oxide based gas sensors on CO concentration. Anyway, this can be improved further by increasing the amount of learning data. This is a general remark on fuzzy systems, since the knowledge capacity of such systems is infinite.

6.2. Effect of s and t norm implementation.

Inequalities (3) and (4) are very important in the selection of the specific functions evaluating the s and t norms. Since the sets of all s and t norms are bounded one can define an ordering relationship in these sets. Having in mind inequalities (3) and (4), it is obvious that all s (t) norms are between the max (min) function and the constant 1 (0) function. A min–max configuration (i.e. using the max function for the implementation of s norm and the min function for the t norm) gives very fast learning but is very sensitive to noise because of the abrupt transition of min and max derivatives. On the other hand, the implementation of s and t norms by constant functions (giving 1 and 0, respectively), although very stable, it never learns the intermediate answers. The optimum selection is somewhere in the middle. That is, the functions evaluating the s and t norms must both have 'soft' derivatives in order to make stable networks (i.e. eliminating abrupt transitions observed in min–max configuration) and cover the entire [0,1] space in order to increase the knowledge capacity of the system, which is zero in the case of the implementation of s and t norms by constant functions.

6.3. Comparison with back propagation network

Using a simple back propagation network instead of a fuzzy one, the results from comparison between the two topologies may be summarized in the following:

(i) in the case of gas identification, the back propagation network learns quickly to distinguish between the three gases, as in the case of fuzzy network. The only difference is that the output of the back propagation network does not follow the gas concentration.

(ii) in the case of water vapor interference elimination on CO sensing, it was found impossible to train a back propagation network with a single output. This happens because during training, the network output was forced to take all the intermediate values between 0 and 1, in order to follow the CO concentration. Since the activation of the output neurons are described by a sigmoid function with relative abrupt transition from 0 to 1, the network was driven to local minimums where the output remains constant for every input pattern.

(iii) It was possible to train a back propagation network to recognize the amount of CO concentration at different levels of relative humidity by the following way: A network with five output neurons was used. The first neuron was activated when the concentration of CO was lower than 200 ppm, the second was activated if the concentration of CO was between 200 and 400 ppm and so on.

7. Conclusions

Fuzzy neural networks can be successfully used as pattern recognition systems in gas sensing. Their main advantages compared to back propagation networks are: (i) they do not suffer from abrupt transitions since they are not restricted to discrete answers, (ii) they can be educated very efficiently by increasing the amount of learning data and (iii) they can store a great deal of information that is very difficult to represent with classical neural networks.

References

[1] Y. Shu, Y. Chen and W. Lu, A taste-sensing system with BP and SOM neural network, *Technical Digest of the Fifth International Meeting on Chemical Sensors*, Vol. 2, 1992??, pp. 1041–1044.

[2] C. Di Natale, F.A.M. Davide, A. D'Amico, A. Hierlemann, J. Mitrovics, M. Schweizer, U. Weimar and W. Gopel, A hybrid neural network for classification of binary mixtures, *Technical Digest of the Fifth International Meeting on Chemical Sensors*, Vol. 2, 1992??, pp. 1101–1104.

[3] C. Di Natale, F.A.M. Davide, A. D'Amico, G. Sberveglieri and P.

Nelli, Sensor array based on metal oxide semiconductor thin films for the recognition of vintage years of wine, *Technical Digest of the Fifth International Meeting on Chemical Sensors*, Vol. 2, 1992??, pp. 1094–1096.

Biographies

John Avaritsiotis was born in Greece in 1948. He received his B.Sc. (Hon.) in physics from the Department of Physics of the University of Athens in 1972. His M.Sc. and Ph.D. degrees were obtained from Loughborough University of Technology, UK, in 1974 and 1976, respectively, in the field of thin-film teclmology and fabrication of thinfilm devices. From 1976 to 1979 he was employed as a research fellow in the Thin Film Group of Loughborougll University of Teclmology. From 1980 to 1986 he was appointed lecturer at the Electronics Laboratory of Athens University, and in 1986 he was elected associate professor in the Department of Electrical Engineering of the National Teclmical University of Athens. In 1990 he was elected professor of microelectronics in the same department. He has worked as a technical consultant to various British and Greek industrial firms for the incorporation of new thin-film deposition techniques in their production processes. He has published over 40 teclmical articles in various scientific journals, and has presented more than three dozen papers at international conferences. His research activities have been focused in the field of microelectronic materials and thin film deposition teclmiques for use in the fabrication of semiconductor devices. His present research interests concern study of the electrical properties of thin insulating and semiconducting films deposited by sputtering, on-chip interconnections and MCMs for high-speed circuits, desigh of ASICs, development of fabrication processes for the production of solid-state gas sensors and design and prototyping of smart sensors. Professor Avaritsiotis is a member of the Editorial Advisory Board of the journal *Active and Passive Electronic Components*, a Guest Editor of IEEE Transactions on *Components, Packaging and Manufacturing Technology*, and a senior member of IEEE and a member of IOP and ISHM.

D.S. Vlachos was born in Pireas, Greece, in 1967. He received the B.E. degree in Electrical Engineering from the National Technical University of Athens in 1991. He is currently enrolled in the Ph.D. program in Electrical Engineering at the National Technical University of Athens. Since 1992 he has been working for the National Technical University of Athens, Department of Electrical Engineering, Division of Computer Science, Microelectronics Technology Laboratory as a research engineer. He has published over 20 papers in international conferences and journals. His research interests include modelling and simulation of gas-solid interface interactions, chemisorption theory, fabrication of solidstate gas sensors, neural networks and fuzzy neural networks for gas sensing.