

# Gas Identification Based on Committee Machine for Microelectronic Gas Sensor

Minghua Shi, Amine Bermak, *Senior Member, IEEE*, Sofiane Brahim Belhouari, and Philip C. H. Chan, *Senior Member, IEEE*

**Abstract**—Gas identification represents a big challenge for pattern recognition systems due to several particular problems such as nonselectivity and drift. The purpose of this paper is twofold: 1) to compare the accuracy of a range of advanced and classical pattern recognition algorithms for gas identification for the in-house sensor array signals and 2) to propose a gas identification ensemble machine (GIEM), which combines various gas identification algorithms, to obtain a unified decision with improved accuracy. An integrated sensor array has been designed with the aim of identifying combustion gases. The classification accuracy of different density models is compared with several neural network architectures. On the gas sensors data used in this paper, Gaussian mixture models achieved the best performance with higher than 94% accuracy. A committee machine is implemented by assembling the outputs of these gas identification algorithms through advanced voting machines using a weighting and classification confidence function. Experiments on real sensors' data proved the effectiveness of the system with an improved accuracy over the individual classifiers. An average performance of 97% was achieved using the proposed committee machine.

**Index Terms**—Classification, density models, ensemble methods, gas sensor array, neural networks, pattern recognition.

## I. INTRODUCTION

**G**AS identification has raised extensive attention in the past twenty years. The ability to monitor and precisely measure leakages of combustible and explosive gases is crucial in preventing the occurrence of accidental explosions. Accordingly, the development of sensors combined with pattern recognition systems that can selectively detect and determine various kinds of combustible gases is urgently needed. Different types of sensors can be used for detecting combustible gases [1]. Microelectronic gas sensors have attracted widespread attention in the past decade. In this paper, we use the micro-hotplate (MHP)-based SnO<sub>2</sub> thin-film sensors, which offer a number of interesting features and are particularly attractive for their practical interest [2]. Indeed, these devices feature high sensitivity, lower power consumption, and compactness and compatibility with semiconductor technology. However, thin-film sensors suffer from a number of shortcomings such as nonselectivity, nonlinearities of the sensor's response, and long-term drift. Pattern recognition algorithms combined with

a gas sensor array have been traditionally used to address these issues [3]. In fact, a gas sensor array permits to improve the selectivity of the single gas sensor and shows the ability to classify different odors. An array of different gas sensors is used to generate a unique signature for each odor. After a preprocessing stage, the resulting feature vector is used to solve a given classification problem, which consists of identifying an unknown sample as one from a set of previously learned gases. Significant work has been devoted to design a successful pattern analysis system for machine olfaction [3]. Various kinds of flexible pattern recognition algorithms have been used for classifying chemical sensor data. Most notably, neural networks have been exploited, in particular, multilayer perceptrons (MLPs), radial basis functions (RBFs), and self-organizing maps (SOMs) [3]. Other methods based on the class-conditional density estimation have been used, such as the quadratic and  $K$ -nearest neighbors (KNN) classifiers.

Recently, a new family of semiparametric methods based on mixture distributions such as Gaussian mixture models (GMMs) has been successfully applied for a number of applications such as speech recognition [4] and image retrieval [5]. Despite their great potential as classifiers, density-model-based approaches have not been exploited for machine olfaction and electronic nose applications. In this paper, we will present a gas classification approach based on a committee machine that combines class-conditional density estimation using different density models and discriminant functions as well as probabilistic principal component analysis (PPCA). First, the ability of the proposed models to perform gas identification will be compared using an experimentally obtained data set. An integrated sensor array has been designed with the aim of identifying combustion gases with a special interest in detecting low to medium concentrations of CO gas in the presence of high concentration of other interfering combustion gases. Data collected from the microelectronic gas sensor constituting two chips with four sensors each would first undergo a preprocessing stage before being fed to the classifier. A total of five classifiers (i.e., MLP, RBF, KNN, GMM, and PPCA) will be compared using the same gas data set. Ensembles techniques have drawn considerable attention in recent years [6]. In fact, a set of learning machines increases classification accuracy with respect to a single machine. The second aim of this paper is to introduce a construction of a committee machine for gas identification to obtain a highly accurate classifier using novel weighting and confidence functions.

This paper is organized as follows: Section II presents the integrated micromachined gas sensor technology and the

Manuscript received June 15, 2004; revised March 4, 2006. This work was supported by the Research Grant Council of Hong Kong under Competitive Earmarked Research Grant HKUST 6162/04E.

The authors are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (HKUST), Kowloon, Hong Kong (e-mail: ebermak@ust.hk).

Digital Object Identifier 10.1109/TIM.2006.880956

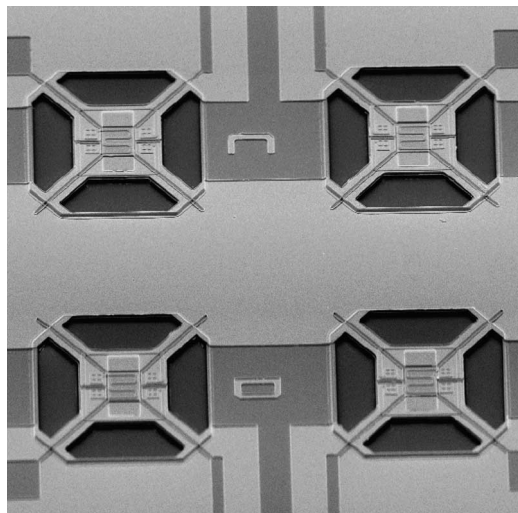


Fig. 1. Microphotograph of the integrated gas sensor array.

experimental characterization of the integrated gas sensor array. Section III describes the pattern recognition procedure using classical and advanced classifiers and their identification performances. Section IV presents the proposed ensemble classifier: a gas identification ensemble machine (GIEM). Based on the analysis of the GIEM performance, a new weighting strategy used to build a committee machine referred to as GIEM-NW is proposed, allowing to reach performances of up to 97%. Section V concludes this work.

## II. SENSOR TECHNOLOGY AND DATA DESCRIPTION

Microelectronic gas sensors based on tin oxide films are extensively used in gas detection applications. When heated at high temperature levels (300 °C), the tin-oxide-based sensors are sensitive to specific gases. To reach such high temperatures, a microelectronic structure referred to as the MHP was developed [7]. The thermally isolated hotplate was fabricated using surface silicon micromachining technique, which provides excellent manufacturing yield compared to its bulk micromachined counterpart as well as improved thermal characteristics, which are essential in integrated gas sensor applications. This also provides isolation required when additional signal processing circuitry is to be implemented on the same die.

In a previous paper [2], the stability of the sensor as well as its sensitivity were experimentally characterized, and it was found that the sensor exhibits a stability figure of up to 500 cycles and 1 ppm sensitivity to CO gas. Based on this sensor structure, an integrated gas sensor array, including four individually controllable units, was developed. Fig. 1 shows a microphotograph of the manufactured chip including four sensors on a single chip featuring different characteristics. The sensor array would provide four distinct responses, which could be seen as a fingerprint or a signature corresponding to a given gas mixture, which can then be exploited by a pattern recognition system in order to build a selective detection system, as will be described in the next section. Distinct responses are obtained from the two chips, which integrate a total of eight sensors, by using different sensing films with chip 1 integrating sensors S1–S4

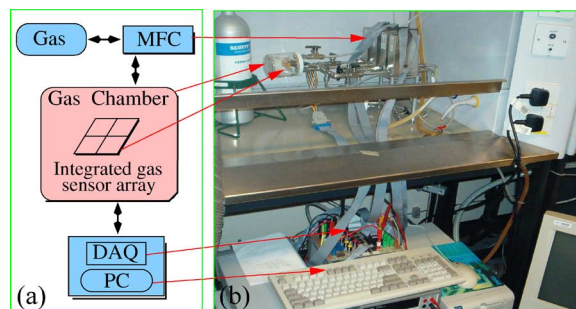


Fig. 2. Experimental setup used to characterize the sensors. MFC stands for mass flow controller, and the DAQ is the data acquisition board used to acquire the signals from the sensor array. (Color version available online at <http://ieeexplore.ieee.org>.)

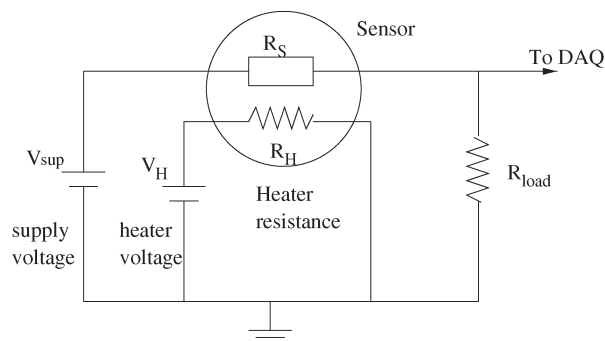


Fig. 3. Electronic circuit used to monitor the sensor response.

and chip 2 integrating sensors S5–S8. The sensing films of sensors S1 and S5 are based on Au/SnO<sub>2</sub>, those of sensors S2 and S6 are based on Pt/Cu(0.16 wt.%)–SnO<sub>2</sub>, and, finally, those of the remaining four sensors S3, S4, S7, and S8 are based on Pt/SnO<sub>2</sub>. It should be noticed that the two chips are identical; however, the operating temperature is set differently in order to tune the selectivity of the two chips to different gases (chip 1 operates at 260 °C, whereas chip 2 operates at 300 °C).

Extensive measurements have been done using an experimental setup consisting of a special sensor chamber equipped with gas pumps and mass flow controllers (MFCs) as well as a data acquisition board as shown in Fig. 2. This figure also shows a photograph of the characterization platform used. Gases used in the experiment are methane, carbon monoxide, hydrogen, and two binary mixtures of methane and carbon monoxide as well as hydrogen and carbon monoxide, respectively. Vapors were injected into the gas chamber, with a diameter of 90 mm and a reaction volume of 100 cm<sup>3</sup>, at a flow rate determined by the MFCs. The gas concentrations in the sensor chamber are adjusted by selecting the correct flow rate for different gases. Input signals generated by the data acquisition board and used to control the MFC are pulse signals corresponding to different concentrations. Fig. 3 shows the schematic of the circuit used to monitor the sensor response and the operating temperature. Two external voltages are applied to the sensor. One is used to preheat ( $V_H$ ) the sensor, whereas the second supply ( $V_{Sup}$ ) is used to read out the output voltage through a voltage divider. The output is then processed by the data acquisition (DAQ) board. The power consumption was measured at about 200 mW per chip, when the operating temperature is 300 °C.

TABLE I  
SENSOR ARRAY CHARACTERISTICS AND CHARACTERIZATION  
SETUP PARAMETERS

Process	In-house $4\mu m$ .
Sensor array	$2 \times 2$
Power	$200mW$ at $300^\circ C$
Sensitivity	$14 \Omega$ at $55$ ppm CO.
Heater Resistance	$12K\Omega$ at $300^\circ C$
Tested gases	CO, CH <sub>4</sub> , H <sub>2</sub> ,
Concentrations	CO : 25 – 200, CH <sub>4</sub> : 500 – 4000, H <sub>2</sub> : 500 – 2000

Table I summarizes the chip features as well as the characterization setup parameters and the concentration ranges used in the experiment. The sensors' output is voltage measurement in the form of exponential-like curves. Fig. 4 shows the output from the two sensor array chips (left and right figures). The figure illustrates how it is possible to use an array of sensors in order to obtain different response signatures. In order to further illustrate the selectivity problem, Fig. 5 reports the polar representation of the normalized steady-state responses for the sensor array exposed to different gases. Although some differences between the patterns can be seen, the shapes appear to be quite similar, and the sensors present a lack of selectivity.

The relationship between the gas concentration and the output voltage was carefully analyzed. The temporal response as well as the steady-state voltage decreases as a function of increasing concentrations for the eight microelectronic gas sensors. We notice that our sensors exhibit nonlinear relationship between its input (gas concentration) and output (voltage). Hence, a normalization procedure is necessary in order to reduce the influence of concentrations and nonlinearities. In gas discrimination applications, the identification must be based on signature pattern, and not on the concentration-dependent amplitudes [8], [9]. Thus, normalization has been previously employed to reduce the concentration dependence within the data space. Normalization is also useful to set the range of values for sensors' output to [0, 1] range in order to avoid the data pattern with larger signal magnitude to dominate in the data space. Data preprocessing is achieved by a baseline subtraction and Euclidean normalization. The raw data generated by the eight sensors are transformed into a characteristic feature vector expressed by

$$x_i = \frac{v_{i,ss} - v_{i,b}}{\sum_{i=1}^8 (v_{i,ss} - v_{i,b})^2} \quad (1)$$

where  $i$  is the sensor number,  $v_{i,b}$  is the baseline voltage of sensor  $i$  before gas exposure, and  $v_{i,ss}$  is the steady-state voltage observed for the  $i$ th sensor during the exposure of the gas.

A gas data set of 220 patterns was used for training the proposed algorithm and estimating the identification performance.

### III. INDIVIDUAL CLASSIFIERS' DESCRIPTION AND PERFORMANCE

#### A. Algorithmic Review

The goal of a pattern classifier is to generate a class label  $C_{\hat{k}}$  for an unknown feature vector  $\mathbf{x} \in R^d$  from a discrete set of previously learned classes  $\{C_1, C_2, \dots, C_c\}$ . For each class  $C_k$ , we can consider the posterior probability of class membership  $\varphi(C_k|\mathbf{x})$ . It can be demonstrated that the decision rule that gives the smallest misclassification error consists in assigning  $\mathbf{x}$  to class  $C_{\hat{k}}$  as

$$C_{\hat{k}} = \arg \max_k [\varphi(C_k|\mathbf{x})] = \arg \max_k [\varphi(\mathbf{x}|C_k)\varphi(C_k)] \quad (2)$$

where  $\varphi(\mathbf{x}|C_k)$  is the class-conditional density, and  $\varphi(C_k)$  is the prior probability. In the absence of prior knowledge,  $\varphi(C_k)$  can be approximated by the relative frequency of examples in the data set. One way to build a classifier is to develop a model (such as MLP and RBF) that estimates the posterior probabilities directly, where the boundaries are learned from data. An alternative way is to estimate the class-conditional densities by using representation models for how each pattern class populates the feature space. In this approach, classifier systems are built by considering each of the classes in turn and estimating the corresponding class-conditional densities  $\varphi(\mathbf{x}|C_k)$  from the data. We briefly present the algorithms used in this paper: GMMs, PPCA [10], KNN, MLP, and RBF neural networks [11].

1) *KNN*: The most widely used method of nonparametric density estimation is the KNN. To classify a pattern  $\mathbf{x}$ , we find the closest  $K$  examples in the data set and select the predominant class  $C_{\hat{k}}$  among those  $K$  neighbors. Despite the simplicity of the algorithm, it often performs very well and is an important benchmark method. However, one drawback of KNN is that all the training data must be stored, and a large amount of processing is needed to evaluate the density for a new input pattern.

2) *Density Models*: Density estimation can be carried out using mixture distributions. In a mixture model, a probability density function is expressed as a linear combination of basis functions. A model with  $M$  components is described as a mixture distribution [12] by

$$\varphi(\mathbf{x}) = \sum_{j=1}^M \varphi(j)\varphi(\mathbf{x}|j) \quad (3)$$

where  $\varphi(j)$  is the mixing coefficient, and the parameters of the component density functions  $\varphi(\mathbf{x}|j)$  vary with  $j$ . The method for training mixture model is based on maximizing the data likelihood. The log likelihood of the data set  $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ , which is treated as an error, is defined by

$$l = \sum_{i=1}^n \log \varphi(\mathbf{x}_i). \quad (4)$$

A specialized method is commonly used to produce optimum parameters, known as the expectation-maximization (EM)

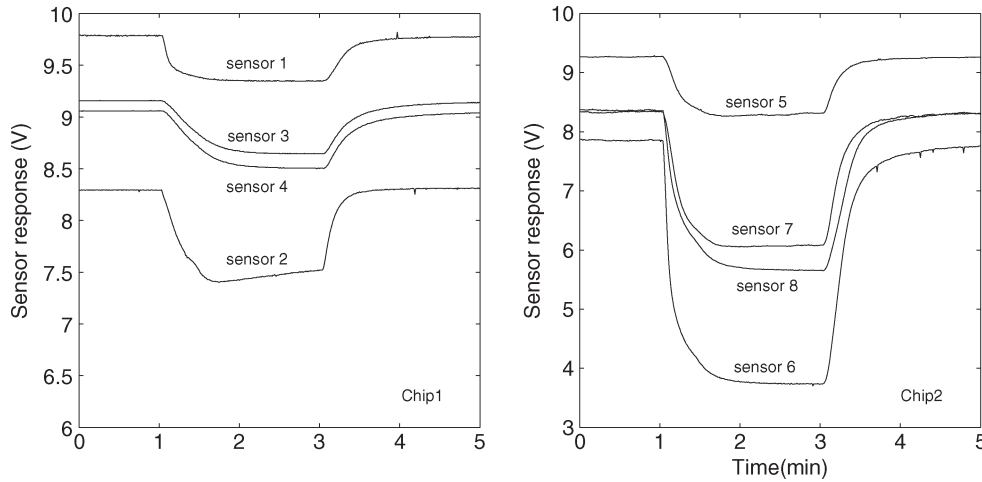


Fig. 4. Typical response of the sensor array. The figure on the left shows the response of chip 1, whereas the figure on the right shows the response of chip 2.

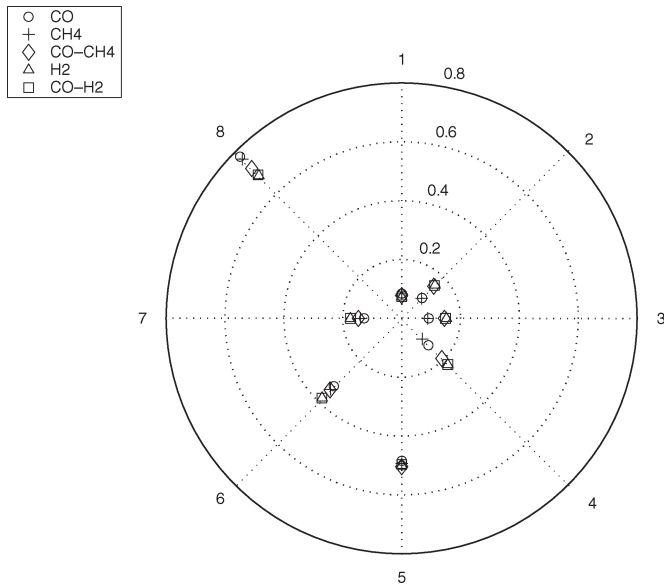


Fig. 5. Polar representation of the array response exposed to different gases. Measurement type: 200 ppm CO (circles), 500 ppm CH<sub>4</sub> (plus signs), mixture 25–500 ppm CO–CH<sub>4</sub> (diamonds), 500 ppm H<sub>2</sub> (triangles), and mixture 25–500 ppm CO–H<sub>2</sub> (squares).

algorithm [11]. In GMM, each mixture component is defined by a Gaussian parametric distribution in  $d$ -dimensional space by

$$\varphi(\mathbf{x}|j) = \frac{1}{(2\pi)^{d/2} |\Sigma_j|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_j)^T \Sigma_j^{-1} (\mathbf{x} - \boldsymbol{\mu}_j) \right\}. \quad (5)$$

The parameters to be estimated are the mixing coefficients  $\varphi(j)$ , the covariance matrix  $\Sigma_j$ , and the mean vector  $\boldsymbol{\mu}_j$ . PPCA is a Gaussian model with certain constraints. The distribution of the observed data is given by [10]

$$\varphi(\mathbf{x}|\mathbf{W}, \sigma) = \frac{1}{(2\pi)^{d/2} |\mathbf{C}|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \mathbf{C}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \quad (6)$$

where  $\mathbf{C} = \mathbf{W}\mathbf{W}^T + \sigma^2\mathbf{I}$ . The covariance matrix is the sum of two terms: One is diagonal in a  $q$ -dimensional subspace

spanned by the first  $q$  principal components and the other is spherical.

A mixture of PPCA has the same form as (3), where each component density function is given by a PPCA. Hence, the training of such a model can be done in the maximum-likelihood framework using the EM algorithm.

3) *Neural Networks*: The posterior densities can be estimated via discriminant analysis, where the boundaries are directly learned from data. This supervised classification is based on the learning of the input pattern–class label mapping. The simplest discriminant functions are linear in the features. One way to generalize the discriminant function, so as to permit a much larger range of possible decision boundaries, is to transform the input vector  $\mathbf{x}$  using a set of predefined nonlinear basis functions  $\phi_j(\mathbf{x})$  and represent the output as a linear combination of these functions. MLP and RBF are the two most popular types of neural network architectures used for electronic nose applications [3].

### B. Classification Results

Let us now study the classification capabilities of each selected classifier and compare their performances. Experiments are based on our gas sensors data of 220 patterns. Each pattern consists of eight steady-state responses. Since the data set that we used was small, generalization performances were estimated by using a tenfold cross-validation approach. The 220 gas data patterns are mixed randomly and divided into ten groups (22 data patterns in each group). Density models are built by considering each of the class in turn and estimating the corresponding class-conditional densities  $\varphi(\mathbf{x}|C_k)$  from the data. For discriminant functions, a softmax output activation function is used to ensure that the outputs lay in the [0, 1] range and summed to one. A comparative study of different dimensionality reduction techniques has been conducted in [13]. The best results were found using the PCA projection. Therefore, the inputs to each classifier are the projections of the data using PCA. The selected classifiers are built on projected data sets with different numbers of principal components. The parameters of each mixture models were adapted to the

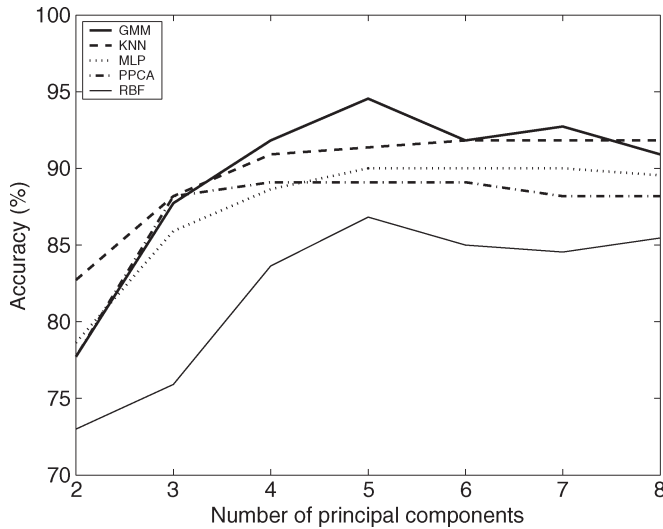


Fig. 6. Accuracy as a function of the number of principal components.

training data in the maximum-likelihood framework using EM algorithm. All parameters (for example, the number of hidden units) were chosen according to the validation stage. Fig. 6 gives the classification performance of the trained classifiers for different numbers of principal components. This figure shows that the best performance is achieved when projecting to five principal components and the addition of components actually degrades the classifier performance. The most accurate discriminant function is the MLP, whereas the most accurate density model is the GMM. This points out to an important result, which suggests that higher generalization performance can be obtained by using feature reduction as preprocessing for increasing the ratio of the number of training samples over the number of features. This is mainly true for density model classifiers as they are based on statistical approaches. Since only a limited number of examples are typically available for gas sensors applications, there is an optimal number of feature dimensions beyond which the classifier performance starts to degrade.

#### IV. PROPOSED ENSEMBLE CLASSIFIER

Ensembles are well established as a method for obtaining highly accurate classifiers by combining different algorithms. We present a GIEM, which is a novel approach for assembling the outputs of various gas identification algorithms to obtain a unified decision with enhanced accuracy.

##### A. Architecture of the Committee Machine

A number of researchers have applied ensemble methods to improve the performance of neural networks [14]. The basic idea of a committee machine is to combine a mixture of experts and to effectively make use of the results produced by each expert within the ensemble. Fig. 7 provides the architecture of our GIEM system with five algorithms (i.e., MLP, GMM, RBF, KNN, and PPCA). By combining the result of each classifier, we can arrive at a final result with improved performance. Each classifier gives its result  $R$  and the confidence  $Cf$  for

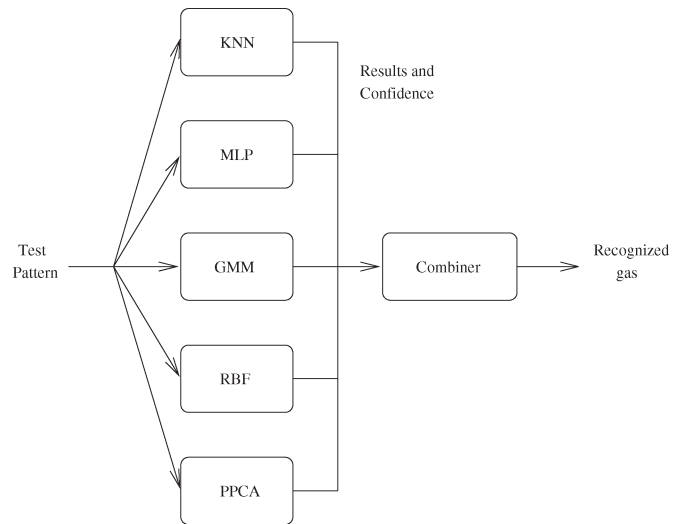


Fig. 7. Ensemble machine architecture.

the result to the combiner. We introduce the use of confidence as a weighted vote for the combiner to avoid affecting the final decision by the result of individual expert featuring low confidence. In order to find confidences of various algorithms, we adopted three different approaches.

- 1) KNN: To classify a pattern  $\mathbf{x}$ , we find the closest  $K$  examples in the data set and select the predominant class  $C_{\hat{k}}$  among those  $K$  neighbors. Thus, the confidence of the result is defined as the number of data point of the result class divided by  $K$ , i.e.,

$$Cf = \frac{K_{\hat{k}}}{K}. \quad (7)$$

- 2) Density models: We compute  $\varphi(C_k|\mathbf{x})$  by modeling the class-conditional density  $\varphi(\mathbf{x}|C_k)$  (i.e., we train a model for each class) and then applying Bayes' rule to compute the following posterior distribution:

$$\varphi(C_k|\mathbf{x}) = \frac{\varphi(\mathbf{x}|C_k)\varphi(C_k)}{\sum_{l=1}^c \varphi(\mathbf{x}|C_l)\varphi(C_l)}. \quad (8)$$

The class  $C_{\hat{k}}$  with the maximum posterior probability is chosen as the result, and the corresponding probability value is considered as the classification confidence of the density models.

- 3) Discriminant functions: We choose a binary vector of size  $c$  for the classifier output. The target class is set to one, and the others are set to zero. The class  $C_{\hat{k}}$  with output value closest to 1 is chosen as the result, and the output value is chosen as the classification confidence of the discriminant function.

Once the classification result of each individual classifier is obtained, the corresponding decision of each classifier is weighted. The idea of introducing a weighting function is to increase the impact of the best classifiers in making the final decision. The weight  $W$  is therefore arbitrarily chosen as the average classification accuracy of each individual classifier. After collecting the results  $R$ , confidences  $Cf$ , and the average accuracies that correspond to the weights  $W$ , from the five

TABLE II  
GAS IDENTIFICATION RESULTS USING GIEM (IN PERCENT)

Set	MLP	KNN	GMM	RBF	PPCA	GIEM
1	95.4	95.4	86.4	86.4	86.4	<b>100</b>
2	<b>100</b>	<b>100</b>	95.4	90.9	95.4	<b>100</b>
3	90.9	86.4	95.4	86.4	90.9	95.5
4	90.9	95.4	<b>100</b>	95.4	95.4	<b>100</b>
5	90.9	90.9	95.4	95.4	90.9	95.5
6	86.4	86.4	90.9	77.3	77.3	86.5
7	90.9	95.4	95.4	77.3	90.9	<b>100</b>
8	86.4	81.8	90.9	77.3	90.9	<b>91</b>
9	86.4	90.9	<b>100</b>	90.9	86.4	95.5
10	81.8	90.9	95.4	90.9	86.4	95.4
Average	90	91.4	<b>94.5</b>	86.8	89	<b>96</b>

algorithms, the combiner assembles the results by calculating the score  $S$  of each class as follows:

$$S_k = \sum_{i=1}^5 W(i) * Cf_k(i) * R_k(i). \quad (9)$$

The class with the highest score would be selected as the recognized class of our GIEM. Equation (9) illustrates the strategy used in our GIEM in order to make the final classification decision. A classifier that presents a high confidence in making a given classification and performs well on average as compared to the other classifiers will be the dominant classifier in the decision process.

### B. Classification Accuracy

To evaluate the performance of the proposed committee machine and individual algorithms, we used the tenfold cross-validation method for the experiment. Table II reports the accuracy of the trained classifiers for different sets. It is clear from this table that GIEM gives superior performances. This is clearly shown in sets 1 and 7, where none of the classifiers has 100% accuracy, but our ensemble machine achieves it. It is, however, important to note that in sets 6 and 9 of Table II, the lower performance of MLP, RBF, and PPCA did affect the final performance of GIEM. This is particularly due in some cases to the lower confidence of GMM while providing the correct classification. This can occur particularly when the number of classes increases in our classification problem. Let us now analyze one pattern of set 9 shown in Table III in which the GMM and KNN produce the correct classification (class 5) while MLP, RBF, PPCA, and GIEM produce a misclassification (class 4). Let us closely analyze the different values of con-

fidence and weights as well as the corresponding scores expressed by (9). We can note from Table III that the difference in performance between the various classifiers is quite low (7.7%). Using the classification performance as a weight coefficient ( $W$  in column 2 of the table) for each classifier will therefore give very little weighting coefficient to the best classifier over the worst one (0.945 versus 0.868). This explains the little impact of GMM on the decision for the test pattern illustrated in this misclassification example. In order to overcome this problem, we propose a new weighting function defined by the following equation:

$$NW_i = \frac{P_i - P_w}{P_b - P_w} \quad (10)$$

where  $NW_i$  is the weight of a given classifier  $i$  with a given performance  $P_i$ .  $P_i$  is defined as the average classification performance of each individual classifier in the committee machine, which is obtained from software simulation results shown in the last row of Table II.  $P_b$  and  $P_w$  are the best and the worst classification performance within the ensemble, respectively. These are defined as  $P_b = \max_{\{1, \dots, n\}} P_i$  and  $P_w = \min_{\{1, \dots, n\}} P_i$ , respectively. Using this new weighting scheme, the impact of each classifier is normalized with respect to its performance within the ensemble. In addition, the worst classifier ( $P_i = P_w$ ) is removed from the ensemble because its corresponding weight will be equal to zero. The last row of Table III illustrates the new scores using the new ensemble referred to as GIEM-NW for the same test pattern. Using the new weighting scheme, a higher score is now obtained for the correct class (class 5) as compared to class 4 for GIEM. The overall performance of the committee machine was evaluated using GIEM-NW. Table IV illustrates the results that are compared with GIEM. An average performance of 97% is now reached using GIEM-NW. The result also demonstrates that with the use of confidence and normalized weighting function, a poor result from individual classifiers would not affect the ensemble result. This robustness property can be seen in all sets of Table IV.

### V. CONCLUSION

In this paper, we presented a gas identification method based on either class-conditional density estimation or discriminant classifier. We conducted comparative experiments on five algorithms for the problem of classifying combustion gases. Data were obtained from an array of eight microelectronic gas sensors using an experimental setup. Comparison of various classifiers revealed that GMM presents the best classification performance (over 94%) among all tested classifiers. A significant dependence of the classifiers' performance on the number of principal components was found. The classification performance was therefore optimized as function of the number of components. The obtained result suggests that higher generalization performance can be obtained by using feature reduction as preprocessing for increasing the ratio of the number of training samples over the number of features, particularly for density model classifiers.

TABLE III  
COEFFICIENT VALUES ( $C_f$  AND  $W$ ) FOR EACH INDIVIDUAL CLASSIFIER FOR ONE TEST PATTERN WITHIN SET 9

Classif	W	NW	Class									
			1 <sup>st</sup>		2 <sup>nd</sup>		3 <sup>rd</sup>		4 <sup>th</sup>		5 <sup>th</sup>	
			$R$	$C_f$	$R$	$C_f$	$R$	$C_f$	$R$	$C_f$	$R$	$C_f$
MLP	0.9	0.416	0	0	0	0	0	0	1	0.8211	0	0.1787
KNN	0.914	0.597	0	0	0	0	0	0	0	0	1	1
GMM	0.945	1	0	0	0	0	0	0	0	0.4581	1	0.5419
RBF	0.868	0	0	0.1245	0	0.0138	0	0	1	0.5579	0	0.4289
PPCA	0.89	0.286	0	0	0	0	0	0	1	0.9532	0	0.0468
Score of GIEM			0		0		0		<b>2.0716</b>		1.4261	
Score of GIEM-NW			0		0		0		0.6142		<b>1.1389</b>	

TABLE IV  
GAS IDENTIFICATION RESULTS USING GIEM-NW (IN PERCENT)

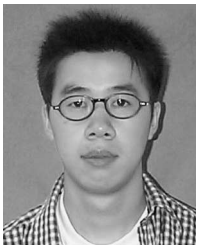
Set	MLP	KNN	GMM	RBF	PPCA	GIEM	GIEM-NW
1	95.4	95.4	86.4	86.4	86.4	<b>100</b>	<b>100</b>
2	<b>100</b>	<b>100</b>	95.4	90.9	95.4	<b>100</b>	<b>100</b>
3	90.9	86.4	95.4	86.4	90.9	<b>95.5</b>	<b>95.5</b>
4	90.9	95.4	<b>100</b>	95.4	95.4	<b>100</b>	<b>100</b>
5	90.9	90.9	95.4	95.4	90.9	95.5	<b>96.3</b>
6	86.4	86.4	90.9	77.3	77.3	86.5	<b>91.3</b>
7	90.9	95.4	95.4	77.3	90.9	<b>100</b>	<b>100</b>
8	86.4	81.8	90.9	77.3	90.9	<b>91</b>	<b>91</b>
9	86.4	90.9	<b>100</b>	90.9	86.4	95.5	<b>100</b>
10	81.8	90.9	95.4	90.9	86.4	95.4	<b>95.9</b>
Average	90	91.4	<b>94.5</b>	86.8	89	96	<b>97</b>

In a second study, this paper proposes a gas identification system based on ensemble machine referred to as GIEM. The proposed committee machine combines five heterogeneous classifiers, and the results of each classifier are combined using a novel confidence and weighting function. A particularly robust result was obtained when using a normalized weighting function referred to as GIEM-NW. The confidence and weighting functions permit to enhance the classification performance by inhibiting poor individual classifiers from affecting the final result. The GIEM-NW achieves a 97% accuracy, which outperforms all other individual classifiers.

REFERENCES

- [1] D.-D. Lee, "Hydrocarbon gas sensors," in *Chemical Sensor Technology*, vol. 5. Tokyo, Japan: Kodansha Ltd., 1994, pp. 79–99.
- [2] P. C. H. Chan, G. Yan, L. Sheng, R. K. Sharma, Z. Tang, J. K. O. Sin, I.-M. Hsing, and Y. Wang, "An integrated gas sensor technology using surface micro-machining," *Sens. Actuators B, Chem.*, vol. 82, no. 2/3, pp. 277–283, Feb. 2002.
- [3] R. Gutierrez-Osuna, "Pattern analysis for machine olfaction: A review," *IEEE Sensors J.*, vol. 2, no. 3, pp. 189–202, Jun. 2002.
- [4] Y. Zhang, M. Alder, and R. Togneri, "Using Gaussian mixture modeling in speech recognition," in *Proc. ICASSP Conf.*, 1994, vol. 1, pp. 613–616.
- [5] N. Vasconcelos and A. Lippman, "Feature representations for image retrieval: Beyond the color histogram," in *Proc. IEEE ICME Conf.*, 2000, vol. 2, pp. 899–902.

- [6] M. Pardo and G. Sberveglieri, "Learning from data: A tutorial with emphasis on modern pattern recognition methods," *IEEE Sensors J.*, vol. 2, no. 3, pp. 189–202, Jun. 2002.
- [7] G. Yan, L. Sheng, Z. Tang, J. Wu, P. C. H. Chan, and J. K. O. Sin, "A low power CMOS compatible integrated gas sensor using maskless tin oxide sputtering," *Sens. Actuators B, Chem.*, vol. 49, no. 1/2, pp. 81–87, Jun. 1998.
- [8] R. Polikar, R. Shinar, V. Honavar, L. Udupa, and M. D. Porter, "Detection and identification of odorants using an electronic nose," in *Proc. IEEE ICASSP*, 2001, vol. 5, pp. 3137–3140.
- [9] E. L. Hines, E. Llobet, and J. W. Gardner, "Electronic noses: A review of signal processing techniques," *Proc. Inst. Elect. Eng.—Circuit Devices and Systems*, vol. 146, no. 6, pp. 297–310, Dec. 1999.
- [10] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *J. R. Stat. Soc., B*, vol. 61, no. 3, pp. 611–622, 1999.
- [11] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Clarendon, 1995.
- [12] D. M. Titterton, A. F. M. Smith, and U. E. Makov, *Statistical Analysis of Finite Mixture Distributions*. New York: Wiley, 1985.
- [13] S. Brahim-Belhouari, A. Bermak, G. Wei, and P. C. H. Chan, "Gas identification algorithms for microelectronic gas sensor," in *Proc. IEEE Conf. IMTC*, Como, Italy, 2004, pp. 675–678.
- [14] M. Su and M. Basu, "Gating improves neural network performance," in *Proc. IEEE Conf. IJCNN*, 2001, vol. 3, pp. 2159–2164.



**Minghua Shi** received the B.S. degree in electronic engineering from the East China University of Science and Technology, Shanghai, China, in 2002. He is currently working toward the Ph.D. degree at the Hong Kong University of Science and Technology (HKUST), Kowloon, Hong Kong, under the supervision of Prof. A. Bermak.

He joined the Department of Electrical and Electronic Engineering, HKUST, in September 2002. His research interests are related to hardware implementation of pattern recognition algorithms for electronic nose applications.



**Amine Bermak** (M'99–SM'04) received the M.Eng. and Ph.D. degrees in electronic engineering from Paul Sabatier University, Toulouse, France, in 1994 and 1998, respectively.

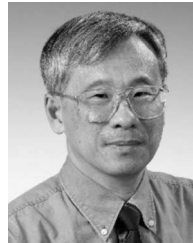
During his Ph.D. years, he was part of the Microsystems and Microstructures Research Group, French National Research Center LAAS-CNRS, where he developed a three-dimensional very large scale integration (VLSI) chip for artificial neural network classification and detection applications. He joined the Advanced Computer Architecture Research Group, York University, York, U.K., where he was working as a Post doctoral researcher on VLSI implementation of correlation matrix memory neural network for vision applications in a project funded by British Aerospace. In November 1998, he joined Edith Cowan University, Perth, Australia, first as a Research Fellow working on smart vision sensors and then as a Lecturer and a Senior Lecturer in the School of Engineering and Mathematics. He is currently an Assistant Professor with the Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology (HKUST), Kowloon, Hong Kong, where he is also serving as the Associate Director of the Computer Engineering Program. His research interests are related to VLSI circuits and systems for signal, image processing, sensors, and microsystems applications. He has published extensively on the aforementioned topics in various journals, book chapters, and refereed international conferences.

Dr. Bermak was the recipient of the Bechtel Foundation Engineering Teaching Excellence Award and the IEEE Chester Sall Award in 2004. In 2005, he received the Best Paper Award at the "5th IEEE International Workshop on System-On-Chip for Real-Time Applications." He is a member of the IEEE CAS Committee on Sensory Systems.



**Sofiane Brahim Belhouari** received the Diploma in electrical engineering from the Polytechnic Institute of Algiers, Algiers, Algeria, in 1993 and the Ph.D. degree in automatic control and signal processing from the University of Paris XI, Orsay, France, in 2000.

After his Ph.D., he held Postdoc positions at the Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland, and the Hong Kong University of Science and Technology (HKUST), Kowloon, Hong Kong. His main research interests are data analysis, statistical signal processing, and pattern recognition.



**Philip C. H. Chan** (SM'95) was born in Shanghai, China, and raised in Hong Kong. He received the B.S. degree in electrical engineering from the University of California, Davis, where he graduated with highest honors and departmental citation, and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, under Prof. C. T. Sah.

He stayed at the University of Illinois initially as an IBM Postdoctoral Fellow and later as a Visiting Assistant Professor in electrical engineering.

He joined Intel Corporation, Santa Clara, CA, in 1981, as a Senior Engineer in the Technology Development Computer-Aided Design Department. Later, he became a Principal Engineer and Senior Project Manager. He had the corporate responsibility for circuit simulation tools, very large scale integration device modeling and process characterization. In 1990, he transferred to the Design Technology Department, Microproducts Group, where he led a team of engineers that defined and developed a computer-aided design system to design multichip module products. This effort led to the first functional 486-based multichip module at Intel. He joined the Hong Kong University of Science and Technology (HKUST), Kowloon, Hong Kong, in April 1991 as a Reader. He became a Professor in 1997. He served as the Director of Undergraduate Studies, the founding Director of Computer Engineering Programme, the Associate Dean of Engineering, and the Acting Head and then Head of the Department of Electrical and Electronic Engineering until 2002. He was also the Director of the Microelectronic Fabrication Facility, which supports all the microelectronic related research at HKUST. He is also the Director of the Advanced Electronic Packaging Laboratory, where he initiated the research in flip-chip technology. He became the Dean of Engineering in September 2003. His research interests include microelectronics devices, circuits, integrated sensors and electronic packaging.