World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:2, No:1, 2008

Gas Detection via Machine Learning

Walaa Khalaf, Calogero Pace, and Manlio Gaudioso

***Abstract*—**We present an Electronic Nose (ENose), which is aimed at identifying the presence of one out of two gases, possibly detecting the presence of a mixture of the two. Estimation of the concentrations of the components is also performed for a volatile organic compound (VOC) constituted by methanol and acetone, for the ranges 40-400 and 22-220 ppm (parts-per-million), respectively.

Our system contains 8 sensors, 5 of them being gas sensors (of the class TGS from FIGARO USA, INC., whose sensing element is a tin dioxide (SnO_2) semiconductor), the remaining being a temperature sensor (LM35 from National Semiconductor Corporation), a humidity sensor (HIH–3610 from Honeywell), and a pressure sensor (XFAM from Fujikura Ltd.).

Our integrated hardware–software system uses some machine learning principles and least square regression principle to identify at first a new gas sample, or a mixture, and then to estimate the concentrations. In particular we adopt a training model using the Support Vector Machine (SVM) approach with linear kernel to teach the system how discriminate among different gases. Then we apply another training model using the least square regression, to predict the concentrations.

The experimental results demonstrate that the proposed multiclassification and regression scheme is effective in the identification of the tested VOCs of methanol and acetone with 96.61% correctness. The concentration prediction is obtained with 0.979 and 0.964 correlation coefficient for the predicted versus real concentrations of methanol and acetone, respectively.

***Keywords*—**Electronic nose, Least square regression, Mixture of gases, Support Vector Machine.

I. Introduction

THE paper deals with the problems of detection and recognition of binary mixtures of gases as well as with the estimation of their concentrations.

The detection of volatile organic compounds (VOCs) has become a serious task in many fields, because the fast evaporation rate and toxic nature of VOCs could be dangerous at high concentration levels in air and working ambient for the health of human beings. In fact, the VOCs are also considered as the main reason for allergic pathologies, skin and lung diseases.

In this paper, to identify the type of gas we use the Support Vector Machine (SVM) approach, which was introduced by Vapnik [12] as a classification tool. The SVM method strongly relies on statistical learning theory. Classification is based on the idea of finding the best separating hyperplane (in terms of classification error and separation margin) of two points–sets in the sample space (which in our case is the Euclidean eight–dimension vector space). Our classification approach includes the possibility of adopting kernel transformations within the SVM context. We adopt a multisensor scheme and useful information is gathered by combining the outputs of the different sensors.

The use of just one sensor does not allow in general to identify the gas. In fact the same sensor output may correspond to different concentrations of many different gases. On the other hand by combining the information coming from several sensors of diverse types we identify the gas and estimate its concentration. In this paper we present the description of the system shown in Fig. 1, producing the details of its construction, a brief theoretical overview of the mathematical models for classification, and the results of our experiments.

II. Electronic Nose

An electronic nose combines an array of gas sensors, whose response constitutes an odor patterns [10]. A single sensor in the array should not be highly specific in its response but should respond to a broad range of compounds, so that different patterns are expected to be related to different odors. To achieve high recognition rate, several sensors with different selectivity patterns are used and pattern recognition techniques must be coupled with the sensor array [9]. Our system consists of five different types of gas sensors, which are from the TGS class of FIGARO USA INC.

The sensing element is a tin dioxide (SnO_2) semiconductor. In particular three of them are of TGS–822 type, one of the TGS–813 type, and the last one is TGS–2600 type. In addition three auxiliary sensors are used: they are a temperature sensor (LM35 from National Semiconductor Corporation), a humidity sensor (HIH-3610 from Honeywell), and a pressure sensor (XFAM from Fujikura Ltd.). Humidity and temperature changes have a strong effect on most sensors.

Manuscript received November 30, 2007. This work has been partially supported by the Italian "Ministero dell'Università e della Ricerca Scientifica" under PRIN Project "Numerical Methods for Global Optimization and for some classes of Nonsmooth Optimization problems"(2005017083.002).

Walaa Khalaf Dipartimento di Elettronica, Informatica e Sistemistica, Università della Calabria, 87036 Rende (CS), Italia (corresponding author to provide phone: 0039-0984-494695; fax: 0039-0984-494671; e-mail: walaa@deis.unical.it).

Calogero Pace Dipartimento di Elettronica, Informatica e Sistemistica, Università della Calabria, 87036 Rende (CS), Italia (e-mail: cpace@unical.it).

Manlio Gaudioso Dipartimento di Elettronica, Informatica e Sistemistica, Università della Calabria, 87036 Rende (CS), Italia (e-mail: gaudioso@deis.unical.it).

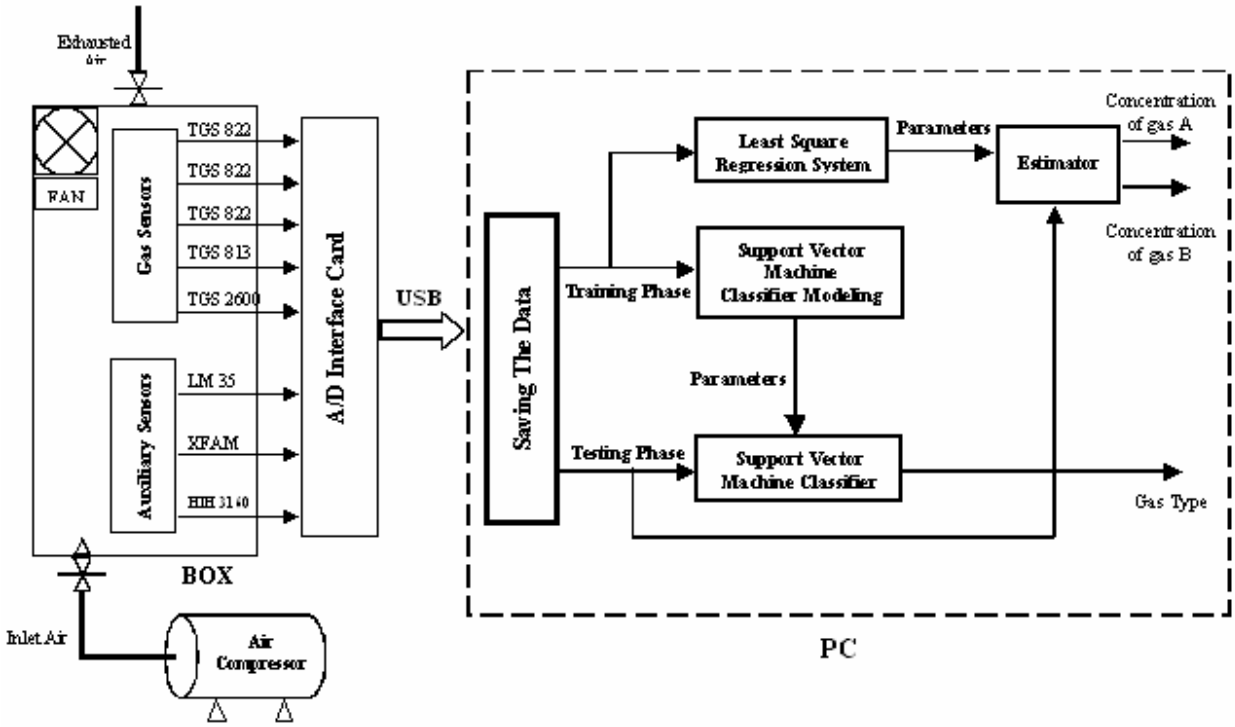

Fig. 1 Block diagram of the system

The gas sensors and the auxiliary sensors are put inside a box with effective volume of 3000 cm^3. Inside the box we put also a fan to let the solvent drops evaporate easily. All sensors are connected with an analog to digital converter (A/D) and then, by means of an interface card, to a PC, where the training for both classification and estimation of the concentration is implemented.

III. Support Vector Machine

Support Vector Machines (SVM) are a set of related supervised learning methods used for classification and regression. They belong to a family of generalized linear classifiers. This family of classifiers has both the abilities of minimizing the empirical classification error and maximizing the geometric margin as shown in Fig. 2. In fact a SVM is also known as maximum margin classifier [5]. In this section we summarize the main features of SVM. Detailed surveys can be found in [2][4][7][12].

As shown in Fig. 2, geometrically the optimal separating hyperplane for two point sets can be found by constructing the convex hulls of the two classes and then equally dividing the shortest distance between the hulls. The orthogonal line will be the optimal hyperplane that separates the two classes. The equation of this hyperplane is defined by:

$$f(x) = \mathrm{w}^T \mathrm{x} + b = 0 \tag{1}$$

and it can be demonstrated that, for a proper normalization, the margin is equal to $\|\mathbf{w}\|^{-1}$. Therefore maximizing the margin is equivalent to minimizing $\|\mathbf{w}\|$. The advantage of this maximum margin criterion is the robustness against noise and uniqueness of the solution.

In many practical cases the data are not linearly separable, then the hyperplane tries to both maximize the margin and minimize the sum of classification errors at the same time. The error ξ_i of a point $(\mathbf{x}_i, y_i)$, $y_i \in \{-1, +1\}$ (with respect to target margin γ and for a hyperplane f) is defined as:

$$\xi_i = \xi((\mathrm{x}_i, y_i), f(\mathrm{x}_i), \gamma) = \max(0, \gamma - y_i f(\mathrm{x}_i)) \tag{2}$$

ξ_i is called margin slack variable, which measures how much a point fails to have margin γ. If y_i and $f(\mathbf{x}_i)$ have different signs the point $\mathbf{x}_i$ is misclassified and

$$\xi_i = \max(0, \gamma - y_i f(\mathrm{x}_i)) > \gamma > 0 \tag{3}$$

The error ξ_i is greater than **zero** if the point $\mathbf{x}_i$ is correctly classified but with margin smaller than γ. Finally, the more $\mathbf{x}_i$ falls in the wrong region, the bigger is the error. The cost function to be minimized is:

$$\frac{1}{2}\|\mathrm{w}\|^2 + C\sum_i \xi_i \tag{4}$$

where C is a positive constant, which determines the trade off between accuracy on the training set and margin width. Therefore, this constant can be regarded as a *regularization parameter*. When C has a small value, the optimal separating hyperplane tends to maximize the distance with respect to the closest point, whereas for large values of C, the optimal separating hyperplane tends to minimize the noncorrectly classified points. If the original patterns are not linearly separable, they must be mapped by means of appropriate kernel functions to a higher dimensional space called *feature*

space. A linear separation in the feature space corresponds to a nonlinear separation in the original input space [13].

Kernels are a special class of functions that permit the inner products to be calculated directly in the feature space, without explicitly applying the mapping. The family of kernel functions adopted in machine learning range from simple linear and polynomial mappings to sigmoid and radial basis functions [8]. A radial basis function kernel maps the training data into an infinite dimensional space. In this paper the linear kernel is used.

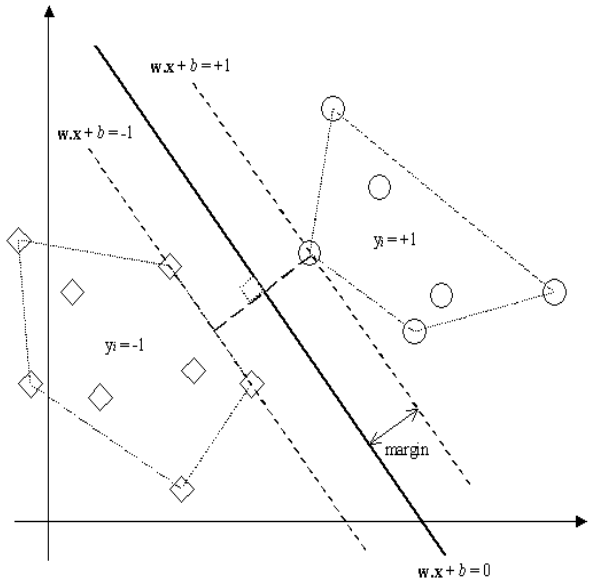

Fig. 2 How to find the optimal hyperplane

IV. Experiments and Results

In our experiments we used two different types of gases, methanol and acetone, at different concentrations. Mixtures between them were also considered. Thus we came out with a 3-classes classification problem. The data set for these gases was made up of samples in R^8 with each sample corresponding to the outputs of the sensors for a given couple (gas, concentration). In the first analysis, we used a SVM with a linear kernel function, and we applied a multiclass classification by using the LIBSVM-2.82 package [3].

The optimal regularization parameter *C* was found out experimentally by minimizing the leave-one-out cross-validation error over the training set. In *k*–fold cross-validation, the data is divided into *k* subsets (where *k* equals the data set size). Then the program is trained *k* times; each time leaving out one of the subsets from training set, and considers this omitted subset as a testing sample to compute the error. The results of classification are shown in Table I with respect to different values of *C*.

TABLE I
Multiple *C* Values vs. Classification Rate

C value	Classification Rate %
1	22.03
10	86.44
50	91.52
100	93.22
200	91.52
300	94.92
400	**96.61**
700	94.92
1000	94.92

We used 20 gas samples for methanol, 24 gas samples for acetone, and 16 gas samples for mixture between them. Each experiment was repeated twice. By using linear Kernel we got 96.61% classification correctness, making 60 cross–validations. After we finished the classification process, the next step was to estimate the concentration of the classified gas (in case of binary mixture), whereas if it is a pure gas we can use our previous method, which is in [6].

To this aim, we have used the least square regression. We built an approximation of the bilinear type of the response (sensor resistance versus concentration) for each sensor. Then we used this approximation to find the concentration for each gas. We came out in our experiments with five coefficients (α's) for the first gas type, five coefficients (β's) for the second gas type, and five coefficients (δ's) for the interaction between the two gases.

They were obtained by solving the following minimization problem with respect to α, β, and δ

$$\min_{\alpha,\beta,\delta} \sum_{j=1}^{M} \sum_{i=1}^{N} \left(S_{ji} - \alpha_j C_{Ai} - \beta_j C_{Bi} - \delta_j C_{Ai} C_{Bi} \right)^2 \tag{5}$$

where N is the number of gas samples (in this work N = 60), M is the number of sensors (M = 5 in our case). We indexed by *j* the sensor and by *i* the sample, *S* is the sensor response in volt or resistance, $\mathbf{C}_A$ and $\mathbf{C}_B$ are the gases concentrations in ppm (parts-per-million). When we solved the above optimization problem, we got the best values for α, β, and δ.

Such values were used to obtain an estimate of the concentration of the gases in new samples.

We adopted again the least squares approach and came out with the solution of the following problem:

$$\min_{C_A, C_B} \sum_{j=1}^{M} (S_j - \alpha_j C_A - \beta_j C_B - \delta_j C_A C_B)^2 \tag{6}$$

TABLE II
EXPERIMENTAL RESULTS

ORIGINAL QUANTITY (PPM)		ESTIMATED VALUES (PPM)		RELATIVE ERROR %	
ACETONE	METHANOL	ACETONE	METHANOL	ACETONE	METHANOL
22	40	17.653	55.287	19.76	38.22
22	120	21.296	115.882	3.20	3.43
22	200	18.776	219.891	14.65	9.94
22	400	29.407	395.362	33.67	1.16
66	40	64.681	36.545	1.99	8.64
66	120	69.495	105.778	5.29	11.85
66	200	75.000	129.285	13.64	35.36
66	400	68.281	468.624	3.45	17.16
110	40	131.431	34.765	19.48	13.09
110	120	117.081	85.873	6.43	28.44
110	200	60.933	236.195	44.61	18.09
110	400	124.953	418.685	13.59	4.67
220	40	207.053	59.456	5.88	48.64
220	120	244.340	108.906	11.06	9.24
220	200	235.446	165.702	7.02	17.15
220	400	302.208	417.796	37.37	4.45

where of course the two unknowns are the concentrations $\mathbf{C}_A$ and $\mathbf{C}_B$, and the values of the parameters (α_j, β_j, and δ_j), j=1,…,M, are those previously calculated.

In Table II we show the real concentrations versus our estimated concentrations as well as the percentage relative error between them for each gas type.

We got 96.61% as a classification success rate, while for the estimation values we got 0.979 and 0.964 correlation coefficients for the predicted versus real concentration of methanol and acetone, respectively. The comparison of our results with other similar results presented in the literature (e.g. [11]) appears extremely encouraging.

ACKNOWLEDGMENT

The authors are grateful to Annabella Astorino for some useful discussions about classification. Also, we wish to thank Prof. Giuseppe Cocorullo from Dipartimento di Elettronica, Informatica e Sistemistica, Università della Calabria-Italy, for supporting this project.

REFERENCES

[1] P. Bartlett and J. Gardner. Odour sensors for an electronic nose, in: Sensors and sensory systems for an electronic nose. *NATO ASI Series E: Applied Science*, 212: 197–216, 1992.

[2] C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and knowledge Discovery,* 2 (2): 121-167, 1998.

[3] C. C. Chang and C. J. Lin. Libsvm: A library for support machines. [Online]. *Available :http://www.csie.ntu.edu.tw/ cjlin/libsvm,* 2001.

[4] N. Cristianini and J. Shawe-Taylor. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press, 2000.

[5] C. Distante, N. Ancona, and P. Siciliano. Support vector machines for olfactory signals recognition. *Sensors and Actuators B*, 88 (1): 30-39, 2003.

[6] M. Gaudioso, W. Khalaf, and C. Pace. On the use of the SVM approach in analyzing an electronic nose. *Proceedings of the seventh International Conference on Hybrid Intelligent Systems*, Kaiserslautern-Germany, pages 42-46, 17-19 September 2007.

[7] R. Gutierrez-Osuna. Pattern analysis for machine olfaction: A review. *IEEE Sensors Journal*, 2 (3): 189–202, June 2002.

[8] K. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf. An introduction to kernel-based learning algorithms. *IEEE Trans. on Neural Networks* 12 (2): 181–201, March 2001.

[9] M. Pardo and G. Sberveglieri. Classification of electronic nose data with support vector machines. *Sensors and Actuators B,* 107: 730–737, 2005.

[10] T. C. Pearce, S. S. Schiffman, H. T. Nagle, and J. W. Gardner. *Handbook of machine olfaction: Electronic nose technology*. WILEY–VCH, 2003.

[11] M. Penza, G. Cassano, and F. Tortorella. Identification and quantification of individual volatile organic compounds in a binary mixture by saw multisensor array and pattern recognition analysis. *Measurement Science and Technology*, 13: 846–858, 2002.

[12] V. N. Vapnik. *Statistical learning theory*. John Wiley and Sons, 1998.

[13] X. Wang, H. Zhang, and C. Zhang. Signals recognition of electronic nose based on support vector machines. *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics,* Guangzhou, pages 3394–3398, 18–21 August 2005.

Walaa Khalaf received the B.Sc. degree in electrical engineering in 1996 from AlMustansirya University, Baghdad-Iraq and, in 2001, the M.Sc. degree in electronic and communication engineering from the same university. Thesis title was "Image compression using wavelet transform". He is a Ph.D. student in the University of Calabria, Italy, since 2004.

He worked in AlMustansirya University/College of Engineering as a researcher from 2002 till 2004. His research interest in Electronic Nose, classification and pattern recognition using support vector machine.

Calogero Pace received the degree in electronic engineering in 1990 from the University of Palermo and, in 1994, the doctoral degree in electronic engineering from the same university.

In 1996 he joined the University of Messina as assistant professor and, in 2002 he moved to the University of Calabria where he is currently associate professor of electronics. He is currently involved in research projects, on the design of low noise electronic instrumentation, on the design and characterization of electronic gas sensors and on the study of nanocrystal memory devices. He is the Coordinator of the Italian Ministry of Foreign Affairs international project "RHESSA" on the radiation hardness of electronic devices and systems for space applications. He is co-author of more than 30 scientific and technical papers published in international refereed journals.

Manlio Gaudioso graduated in Electrical Engineering at Università di Napoli in 1973. He is currently professor of Operations Research at the Università della Calabria.

His interests include numerical optimization for nonsmooth problems and applications of mathematical programming to logistics problems.