Electronic Nose System for Safety Monitoring at Refineries

Dr. Walaa Mohammed Hassan Khalaf
Almustansiriya University
Engineering College
Computer and Software Eng. Dep.

Abstract

In this paper an Electronic Nose (ENose) is presented which is designed for both at identifying the gas type and if it is pure or not, and at estimating the concentrations of the components of that mixture of LPG gases (Methane, Hexane, or Hydrogen) and Hydrogen Sulfide produced always inside the refineries.

Our system contains 8 sensors, 5 of them are gas sensors (of the class TGS from FIGARO USA, INC., the sensing element of two of them is catalytic (TGS-6810 and TGS-6812), the other two its sensing element is a tin dioxide (SnO_2) semiconductor (TGS-825 and TGS-2611) and the last one is an oxygen sensor (KE-50)), the remaining three sensors are auxiliary sensors for measuring a temperature and humidity (HTG–3535), and a pressure sensor (XFAM from Fujikura Ltd.).

The proposed hardware–software system uses some least squares principles for classification and regression to identify at first a new gas sample, if it is pure or mixture, and then to estimate their concentrations, respectively. In particular we adopt a training model using the least squares approach to teach the system how to discriminate among different gases. Then we apply another training model using also the least squares but this time for regression, to predict their concentrations.

The experimental results demonstrated that the proposed binary mixture of the organic vapor (Methane, Hydrogen, or Hexane) and toxic gas (Hydrogen Sulfide) classifier is effective in the identification of the tested gases.

Keywords: Electronic nose, Least square regression, Mixture of gases.

نظام الانف الالكتروني لمراقبة السلامه في مصافي النفط

الخلاصة

في هذا البحث تم تقديم انف الكتروني مصمم لتحديد نوع الغاز فيما اذا كان نقيا او لا، بالاضافة الى التنبوء بتركيز مركبات ذلك المزيج من الغازات (الميثان، الهيكسان، او الهيدروجين) و كبريتيد الهيدروجين الناتجه في داخل مصافي النفط.

النظام المستخدم يتألف من ثمانية متحسسات، خمسة منها هي متحسسات غاز(عنصر التحسس لأثنين منها هو كاتاليتك، و الاثنين الاخرين عنصر التحسس لهما هو ثنائي اوكسيد القصدير، و المتحسس الاخير هو متحسس لغاز الاوكسجين)، المتحسسات الثلاثة المتبقيه هي متحسسات مساعدة، لقياس درجة الحرارة، الرطوبه، و الضغط الجوي.

النظام المقترح في هذا البحث يستخدم مبدأ اقل المربعات للتصنيف و الارتداد، اولا للتعرف على نوع الغاز فيما اذا كان نقي ام لا، و من ثم تخمين تراكيزهما، على التوالي. بشكل خاص قمنا بتبني نموذج تدريب يستخدم مبدأ اقل المربعات لتعليم النظام كيفية التمييز بين الغازات المختلفه. بعد ذلك نطبق نموذج تدريب اخر يستخدم ايضا مبدأ اقل المربعات، لكن هذه المرة للأرتداد، من اجل التنبوء بالتراكيز.

لقد وضحت النتائج المختبريه بأن المزيج الثنائي للغازات العضوية المستخدمة (الميثان، الهيدروجين، و الهيكسان) و الغاز السام (كبريتيد الهيدروجين) قد تم تمييزها بشكل فعال.

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), LP gases, and toxic gases like hydrogen sulfide has become a serious task in many fields, especially in refineries where a great amount of these gases are produced every day and at any moment these gases are considered a dangerous and explosive compounds if its concentrations being out off control.

Electronic Nose system in general consists of an array of sensors which are specifically designed to respond to the surrounding odorant using chemical polymer or metal oxide semiconductor, etc., [1]. The output of E-nose is converted to an electronic signal so that post processing of the collected data is readily conducted via computer for applying various approaches. An important advantage of E-nose is that the synergistic effect of the sensor array can give rise to successful detection of the unknown vapors compared with a single sensor due to the sensor's unique response pattern. The use of just one sensor does not allow in general identifying 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, to identify the type of gas we use a least square model approach as a classification and concentration estimation tool. We adopt a multi-sensor scheme and useful information is gathered by combining the outputs of the different sensors.

In order to overcome the limits of high cost of monitoring systems used in refineries, we propose to use a distributed monitoring system working continuously in real-time. The main goal of this system is to have a continuous and qualitative evaluation of the air quality levels into the various monitored refinery plants. These evaluations are not as accurate as those obtained using the expensive sensors of the traditional systems, but they allow for promptly identifying eventual high risk conditions such as a gas leak.

The intelligence of an E-nose is directly reflected from both classification and estimation capabilities. Various approaches such as neural network, genetic algorithm, and fuzzy logic have been developed to improve the E-nose intelligence. For instance, the classification problem can be solved well in some cases by a number of existing classification techniques such as k-nearest neighbor (k-NN) [2], artificial neural network (ANN) [3,4], and support vector machine (SVM) [5,6].

In this paper we present the description of the system as 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.

Electronic Nose

An electronic nose combines an array of gas sensors, whose response constitutes an odor pattern [7]. To achieve high recognition rate, several sensors with different selectivity patterns are used and pattern recognition techniques must be coupled with the sensor array [1]. Our system consists of five different types of gas sensors which are from the TGS class of FIGARO USA INC as shown in Fig 2.

In particular the sensing element of two of these sensors is a tin dioxide (SnO_2) which are TGS-825 and TGS-2611, the other two sensors with a pilaster material which are TGS-6810 and TGS-6812, and the last one is a unique galvanic cell type oxygen sensor which is KE-50, Table 1 lists the types of sensors used in our system, the gases they are sensitive to and at what sensitivity. In addition three auxiliary sensors are used: they are a temperature and humidity sensor (HTG-3535 from Measurement Specialties), and a pressure sensor (XFAM from Fujikura Ltd.). Humidity and temperature changes have a strong effect on most sensors.

Table 1. Type of Sensors and Corresponding Sensitive Gas

No	Sensors	Gases	Sensitivities
1	TGS-6810	Methane and LP Gas	0 – 10000 ppm
2	TGS-6812	Hydrogen, Methane, and LP Gas	0 – 25000 ppm
3	TGS-2611	VOCs	300-10000 ppm
4	TGS-825	H_2S	5-100 ppm
5	KE-50	Oxygen	0-100%

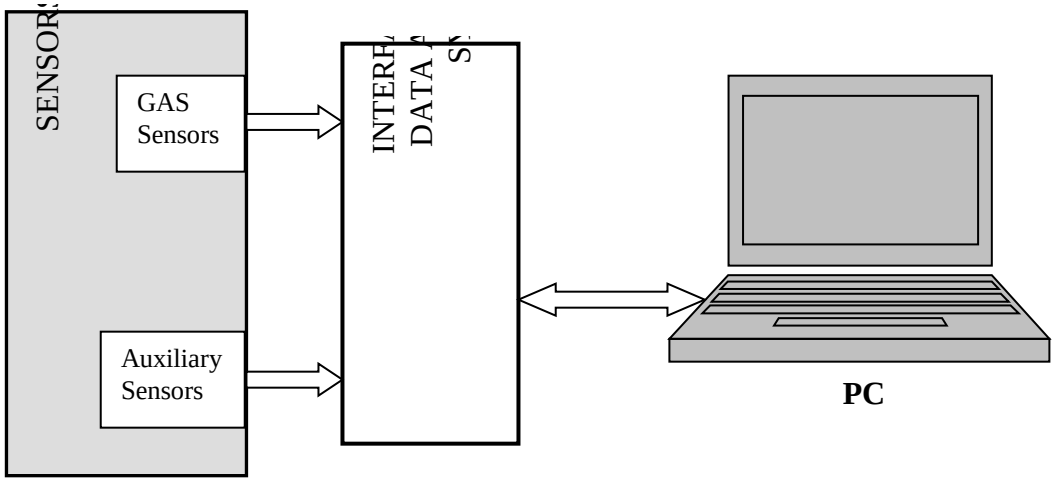

Fig. 1 Block diagram of the system

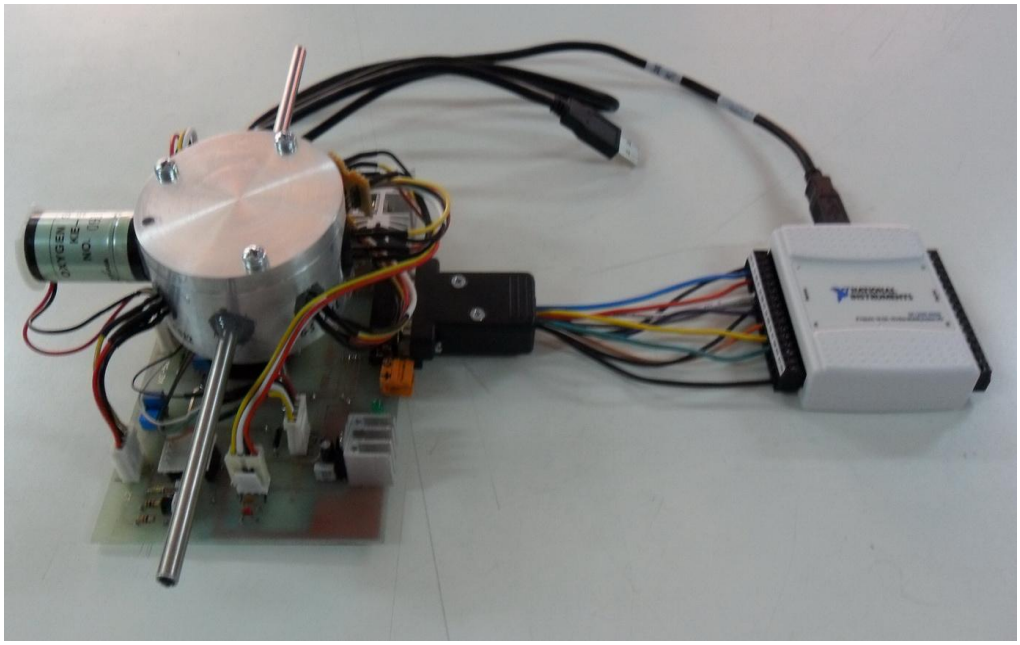

Fig. 2 Electronic Nose System

The gas sensors and auxiliary sensors are put inside a box as shown in Fig 2. All sensors are connected with a multifunction board (NI DAQ-6008), which is used in our system as an interfacing between the box that contains all the sensors and the PC. NI-DAQmx provides an interface to our LabWindows/CVI that we used inside the PC.

LabWindows package is a programming tool running on a PC. The integrated LabWindows/CVI environment features code generation tools and prototyping utilities for fast and easy C code development. It offers a unique, interactive ANSI C approach that delivers access to the full power of **C Language**. Because LabWindows/CVI is a programming environment for developing measurement applications, it includes a large set of run-time libraries for instrument control, data acquisition, and analysis.

The training for both classification and concentration estimation will be implemented inside the PC. Any machine learning system is divided into two phases (modes), training phase and testing phase; in training phase the system finds the optimal parameters for both classification and concentration estimate. In testing phase the system uses these parameters which are calculated during the training phase to analyze a gas sample of unknown type and concentration.

Least Square Model

In this section we formalize our gas type classification approach. For a certain gas index i (i = 1, 2, …, n_G), where n_G represents the number of gases under investigation, the gas class (c_i), as a function of the measures obtained by several sensors, is described as a curve in a **M** dimension space (**M** represents the number of sensors). The parametric equations are:

$$v_j = X^i_j(c_i), \qquad j = 1, 2, \ldots, \mathbf{M} \qquad \ldots(1)$$

where v_j is the response of sensor j, X^i_j is the response function to the gas i of sensor j. If we assume that a given number of samples N_i , i=1,…, n_G, of gas i. is available , we write:

$$z_{jk} = X^i_j(c_{ik}) + w_{ik}\ ,\ k=1, 2, \ldots, N_i,\quad j=1, 2, \ldots, \mathbf{M} \qquad \ldots(2)$$

where z_{jk} represents the response of sensor j for the k *th* sample of gas i, w_{ik} represents the measurement error, It is then possible to build a data fitting strategy using the least squares approach and a parametric model of the functions:

$$X^i_j(c_i) = X(c_i, \theta^i_j), \qquad \theta^i_j \in \Re^p \qquad \ldots(3)$$

where θ represents the equations parameters. From qualitative evaluation of the response curves, it appears reasonable to adopt the same model for all gases and sensors. Once the parameters have been estimated on the basis of sampling data, the parametric curve equations, for each gas class i, are:

$$x^i = \begin{bmatrix} X(c,\hat{\theta}^i_1) \\ X(c,\hat{\theta}^i_2) \\ \ldots\ldots\ldots\ldots \\ X(c,\hat{\theta}^i_{n_S}) \end{bmatrix}, \qquad c_{i,\min} \le c \le c_{i,\max} \qquad \ldots(4)$$

The construction of the curves represents the phase where our classifier is being built. It can be used, as a classification tool, as follows.

Suppose that a newly generated set of measures $z = [z_1, z_2, \ldots, z_{\mathbf{M}}]^T$ carried out on an unknown gas, at unknown concentration is available. The most natural way to obtain a classification of the new sample is to identify the gas type with associated the minimum distance between the sample and the curve that is by detecting the one which provides the minimum distance:

$$dist\ (z, x_i) = \min_{x \in x_i} \left\| z - x \right\|^2 \qquad \ldots(5)$$

Summing up, the classification-evaluation problem is divided into two steps:

1. Construction of the model (for different gases).
2. Solution of minimum distance optimization problem.

Odor Classification

By considering the data, we can find typical curves representing the response of one gas sensor j with respect to one gas type i at different concentrations.

The qualitative behaviors exhibited by the empirical results, suggest adopting a curve fitting model of the power-law type for each sensor ***j*** and for each gas ***i***, that is:

$$R_j(c_i) = \zeta^i_j(c_i, B^i_j, A^i_j) = A^i_j c_i^{B^i_j} \qquad \ldots(6)$$

where we have made explicit the meaning of the parameters θ.

Once the (2 * n_G * **M**) parameters for **A** and **B** have been obtained, we have implemented the detection phase. For any gas sample $z \in R^M$, where ***z*** is the collection of the measures

obtained from the different sensors exposed at any gas at any concentration, we have classified the gas (that is guessed the gas index ***i***) by the following formulae:

$$i_{GAS} = \arg\min_{i} \left\{ \min_{c} \sum_{j=1}^{M} (z_j - x(c, \theta_j^i))^2 \right\} \quad \ldots(7)$$

Odor Concentration Estimation

To estimate the odor consideration we have used again the least squares approach but this time for regression. We built an approximation of bilinear type of the response 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 among the two gases [8].

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

$$\min_{\alpha,\beta,\delta} \sum_{j=1}^{M} \sum_{i=1}^{N} \left(S_{ji} - \alpha_j C_{Ai} - \beta_j C_{Bi} - \delta_j C_{Ai} C_{Bi} \right)^2 \quad \ldots(8)$$

where **N** is the number of gas samples, **M** is the number of sensors (**M**=5 in our case). We index it by ***i*** for the sensor and by ***j*** for the sample, ***S*** is the sensor response in volts, $\mathbf{C}_A$ 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 [8]:

$$\min_{C_A, C_B} \sum_{j=1}^{M} \left(S_j - \alpha_j C_A - \beta_j C_B - \delta_j C_A C_B \right)^2 \quad \ldots(9)$$

Where of course the two unknowns are the concentrations $\mathbf{C}_A$ and $\mathbf{C}_B$, while the values of the parameters ($\alpha_j, \beta_j,$ and δ_j), ***j*** = 1,…, **M**, are those previously calculated.

Experiments and Results

In our experiments we used four different types of gases, Methane, Hexane, Hydrogen, and Hydrogen Sulfide, at different concentrations. Mixtures between each gas and the Hydrogen Sulfide were also considered. Thus we came out with a 7-class classification problem. The data set for these gases is made up of samples in R^8 space where each sample corresponds to the outputs of the sensors for a given couple (gas, concentration). In the first analysis, we used a least squares approach for classification that has been explained in the previous section.

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 consider this omitted subset as a testing sample to compute the error.

We used 10 gas samples for methane, 12 gas samples for Hexane, 8 gas samples for Hydrogen, and 10 gas samples for Hydrogen Sulfide as well as 8 samples for mixtures between each gas and the Hydrogen Sulfide. Each experiment was repeated twice. We got 95.63% classification correctness, making 48 leave one out 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 [5].

To this aim, we have used the least squares regression model has been mentioned previously by using equation 8. We built an approximation of the bilinear type of the response (sensor responses versus concentrations) 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 because from the experiments we knew that the relation is not linear between gas concentration and sensor response, that provide the concentration estimate as a combination of the results produced by the diverse sensors with weights coming from least square computation. They were obtained by solving equation 8 with respect to α, β, and δ.

We adopted again the least squares approach and came out with the solution of equation **9**, where of course the two unknowns are the concentrations C_A and C_B, and the values of the parameters (α_j, β_j, and δ_j), $j = 1,\ldots, M$ are those previously calculated.

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

We got **95.63%** as a classification success rate, also for the estimation values we got **0.919** and **0.942** correlation coefficients for the predicted versus real concentration of methane and H_2S, respectively.

Conclusions

In this paper, we proposed a new classification and concentration estimation method for Electronic Nose system applied in refinery zones for safety and security. Nonlinear polynomial is used to model the sensor characteristic with respect to the vapor concentration. The subsequent concentration estimation and classification are then formulated as a convex optimization problem. Moreover, the proposed method is ready for implementation as an information processing unit of an intelligent E-nose. The proposed method is applicable to single and binary mixture vapors without the need of separate classification. Experimental results confirmed the efficiency of the proposed method as shown in Table II.

Table II. Experimental Results

Original Quantity (ppm)		Estimated values (ppm)		Relative error %	
METHANE	H_2S	METHANE	H_2S	METHANE	H_2S
1000	5	987.96	5.13	17.32	37.62
1000	9	1011.18	9.21	14.45	2.84
1000	12	1023.61	11.79	14.27	6.56
1000	16	994.91	16.15	0.41	8.92
3000	5	2969.68	4.89	4.06	14.00
3000	9	2992.41	8.98	24.86	20.99
3000	12	3016.98	12.11	13.61	35.52
3000	16	3023.95	16.04	3.11	29.18
5000	5	5009.74	5.09	26.13	18.40
5000	9	5022.82	9.17	20.75	35.52
5000	12	4975.39	11.91	32.37	9.05
5000	16	4989.07	15.83	35.52	2.76
9000	5	9100.15	4.44	15.75	68.60
9000	9	9073.63	9.24	3.92	3.38
9000	12	9022.96	13.07	8.20	7.11
9000	16	9050.34	15.86	8.94	26.14

References

1. J. W. Gardner and P. Barlett, “A brief history of electronic noses,” *Sensors and Actuators B*, vol. 18, 1994, pp. 211–220.
2. K. Tang, S. Chiu, C. Pan, H. Hsieh, Y. Liang ,and S. Liu, “Development of a Portable Electronic Nose System for the Detection and Classification of Fruity Odors”, *Sensors* **2010**, *10*, pp. 9179-9193.
3. S. Sharma, “Implementation of Artificial Neural Network For Odor Identification Using E-Nose”, National Conference on Computational Instrumentation CSIO Chandigarh, INDIA, 19-20 March 2010, pp. 187-190.

4. A. Amari, N. El Barbri, E. Llobet, N. El Bari, X. Correig, B. Bouchikhi, Monitoring the freshness of Moroccan sardines with a neural-network based electronic nose, Sensors 6 (2006) 1209–1223.
5. 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.
6. M. Pardo and G. Sberveglieri. Classification of electronic nose data with support vector machines. *Sensors and Actuators B*, 107: 730–737, 2005.
7. 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.
8. W. Khalaf, C. Pace, and M. Gaudioso, "Gas Detection via Machine Learning", International Journal of Intelligent Technology Volume 3 Number 1, 2008, pp. 43-47.