A Novel ANN Ensemble and Self-calibration Model in Electronic Nose for Concentration Estimation[*]

Lei Zhang[1], Fengchun Tian[1], Lijun Dang[1] and Guorui Li[1]

[1]College of Communication Engineering　Chongqing University, Chongqing 400044, China
{leizhang & fengchuntian}@cqu.edu.cn,　danglijunshasha@163.com

Abstract - Electronics nose (e-nose), as an artificial olfaction system, has been used in environmental monitor. This paper presents a novel concentration estimation model for improving the accuracy, robustness and stability of e-nose in long-term use. In the estimation model, two models including an ANN ensemble model and a self-calibration model are studied. The ANN ensemble model is different from single ANN that it belongs to a piecewise linearly weighted prediction model but not nonlinear prediction problem. The self-calibration model is designed for correction of the threshold network in the ensemble due to that the threshold network becomes decay which is caused by sensor drift. Experimental results demonstrate that the proposed model is very effective in real time monitoring of formaldehyde.

Index Terms - Electronic nose; concentration estimation; ANN ensemble; self-calibration; threshold network

1. Introduction

Indoor air quality has attracted people's attentions in recent years. Formaldehyde is one of the most serious contaminants to humans' health, especially in a new house after fitments. New furniture, wall after paint, and new floors will release formaldehyde, benzene and toluene etc.. Therefore, how to monitor the air quality in real-time is the essential objective in our project. Electronic nose (e-nose), as an artificial olfaction system, has a wide range of applications. For example, e-nose has been used in food quality detection [1-3], discrimination of tea [4], milk [5], environment air [6-9], medical treatments and diagnosis [10, 11]. Our project will concentrate on the study of indoor contaminant prediction, especially formaldehyde using an e-nose developed in our lab. Our e-nose is based on a metal oxide semiconductor gas sensor array consisted of four gas sensors and a temperature-humidity module which can perceptron the indoor temperature and relative humidity. The cross-sensitivities and broad spectrum characteristic of gas sensor array make the odor detection possible. This paper will present a concentration estimation model using the sensor array and a novel self-calibration model based on multiple artificial neural networks. Applications of artificial neural network (ANN) in an e-nose for prediction have not been new findings. However, the robustness of ANN in e-nose for long-term application can not be good. Due to that the drift problem of semiconductor gas sensors always exists. Therefore, we present a novel self-calibration model for improving the robustness of e-nose in prediction of formaldehyde's concentration in this paper.

Accordingly, an ANN ensemble model based on a threshold network is also proposed with the self-calibration model.

2. Materials and Methods

A. Electronic Nose Data

The e-nose system has been introduced in our previous publication [12]. Our sensor array in e-nose system consists of four gas sensors from the TGS series including TGS2602, TGS2620, TGS2201A and TGS2201B. In addition, a module (SHT2230 of Sensirion in Switzerland) with two auxiliary sensors for the temperature (T) and humidity (H) are also used. A 12-bit analog-digital converter is used as interface between the Field Programmable Gate Array (FPGA) processor and the sensors. FPGA can be used for data collection, storage and processing. The e-nose system is connected to a personal computer (PC) via a Joint Test Action Group (JTAG) port which can be used to transfer data and debug programs.

For model construction, in e-nose experiments of formaldehyde, we employed 126 samples totally. Each experiment is employed in the constant temperature and humidity chamber. An input vector with 6 variables at the steady state response was selected as feature vector in each observation; for each sample, their concentration (ppm) is obtained using a spectrophotometer which can analyze the collected formaldehyde through a gas sampler.

Besides the data for model learning and test, we also collect 60 samples using another 10 e-nose instruments with completely the same type for validation of the proposed ANN ensemble model and self-calibration model.

B. ANN Ensemble Model

In this paper, the ANN ensemble model is designed with four networks, they are network_1, network_2, network_3 and the threshold network which is essentially a piecewise linear weighted model. Note that the network_1, network_2 and network_3 are used for concentration estimation through a weighting program, and threshold network is used for threshold control which can decide the specific fragment. The structure of the ensemble model has been shown by Fig.1.

The self-calibration model is based on the threshold network. The mathematical model of the ANN ensemble model is shown by Eq.(1).
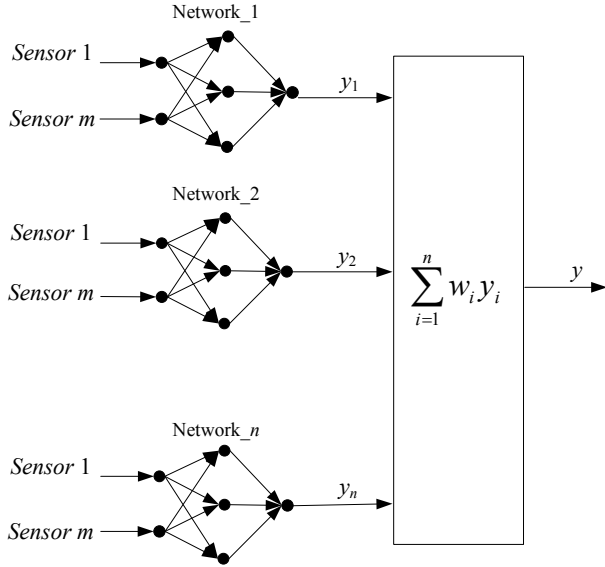
Figure 1. Structure of the ANN ensemble model.

$$y = \begin{cases} \sum_{i=1}^{3} w_1^i \cdot y_i + d_1, & AC \cdot (Th - AB) \geq a_1 \\ \sum_{i=1}^{3} w_2^i \cdot y_i + d_2, & a_2 \leq AC \cdot (Th - AB) < a_1 \\ \vdots & \vdots \\ \sum_{i=1}^{3} w_{n-1}^i \cdot y_i + d_{n-1}, & a_n \leq AC \cdot (Th - AB) < a^{n-1} \\ \sum_{i=1}^{3} w_n^i \cdot y_i + d_n, & AC \cdot (Th - AB) < a_n \end{cases} \quad (1)$$

where $y_i (i=1,2,3)$ denotes the output of the network_i; $w_j^i (i=1,2,3; j=1,\ldots,n)$ denotes the weight coefficient of the network_i in the jth fragment; $d_j (j=1,\ldots,n)$ denotes the constant value in the jth fragment, Th denotes the output of the threshold_network, AC and AB are the adjusted coefficients of the self-calibration model, $a_j (j=1,\ldots,n)$ denotes the controlled threshold value in the jth fragment, $a_1, a_2,\ldots,a_n (a_1>a_2>\ldots>a_n)$, n denotes the number of total fragments. We can find that from the ANN ensemble model we have transformed a nonlinear estimation problem of e-nose into a simplified piecewise linearly weighted estimation problem. Through long-term observation of experiments, we set three fuzzy partitions of concentrations: low concentration (0~0.3ppm), medial concentration (0.3~5ppm) and high concentration (5~20ppm). The parameters expect AC and AB in the model can be obtained through multivariate linear regression method using several experimental samples combined with the fuzzy partitions. The concentration estimation is employed through regression between concentration and concentration but not between sensor response vector and concentration.

C. Self-calibration Model

In model (1), we can find that the threshold_network plays an important role in concentration estimation. The threshold_network should be very sensitivity to formaldehyde and show a good linear characteristic, and then the ANN ensemble model should be very effectiveness. Once all the parameters except AC and AB have been fixed, they can not be adjusted. In this case, after long-term use of e-nose, the drift will occur, then the threshold_network will become weak and result in false output, there will be error in the possible fragment of model (1) and the final estimation of concentration will not be correct. Therefore, a self-calibration model for correction of the threshold_network is necessary.

Assume that the actual output of the threshold_network is Th, and the desired output is Th', then the relationship between the desired output and the actual output can be shown by

$$Th' = AC \cdot (Th - AB) \quad (2)$$

where parameters AC and AB are the self-calibration coefficients.

Due to that Th and Th' can be known, thus AC and AB are easier to be calculated. AB is first determined by fixing $AC=1$ and $Th'=0$, thus $Th-AB=0 \rightarrow AB=Th$ (e-nose is exposed to clean air). When AB is determined, the e-nose will be exposed to formaldehyde with a certain concentration C, then the desired value of Th' can be deduced by model (1). Then, AC can be calculated according to Eq.(2) by

$$AC = \frac{Th'}{Th_{new} - AB} \quad (3)$$

where Th_{new} is the predicted value of threshold_network when exposed to formaldehyde with a certain concentration C.

3. Results and Discussions

The experiments were employed in the chamber, and formaldehyde was tested using the e-nose developed in our project based on the ANN ensemble model and the self-calibration model. Totally, 60 samples were tested by using 10 e-noses with the same type. The prediction result has been shown in Fig.2 which illustrates the true and predicted concentration. We can find that from Fig.2 the prediction can follow the true values very well.

For quantification of prediction error, we have presented the bar plots of the relative error and absolute error in Fig.3. In Fig.3, the samples have been re-arranged according to the concentrations from low to high in which we can find the prediction error in different concentration level. We can find that the relative error is larger in low concentration and the absolute error is larger in high concentration. The mean relative error is 16.76% whose variance is 0.0145. The absolute error is 0.09 ppm whose variance is 0.0077. The variance shows the uniform error distribution for each sample.

Besides, we also present Table 1 in which the proportions of samples in different relative error grade. We can find that proportion 85% was obtained with the relative error less than

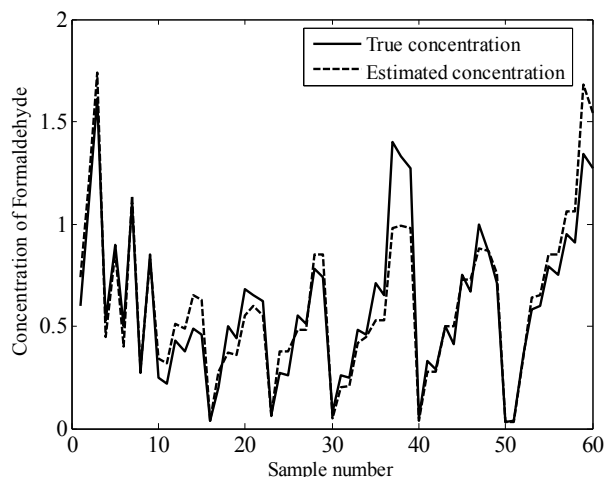30%. Therefore, we believe that the prediction accuracy is good in e-nose application.

Fig.2 Concentration predictions of e-noses based on the ANN ensemble and self-calibration model.

TABLE I. Proportions of Samples in Different Relative Error Grades

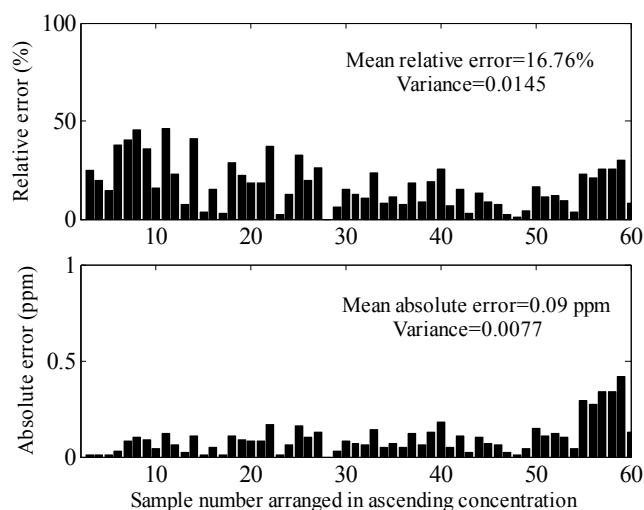| Relative error grades	Number of samples	Proportion (%)
0~10%	20	33.33
10%~20%	19	31.67
20%~30%	12	20
30%~50%	9	15

Fig.3 Relative error and absolute error of prediction.

4. Conclusions

In this paper, we propose a novel ANN ensemble model and a self-calibration model in e-nose for concentration estimation of formaldehyde. This ensemble model is different from single ANN that the proposed model belongs to a piecewise linearly prediction problem but not nonlinear prediction. Therefore, the ensemble model has a better robustness than single network. Besides, due to the sensor drift which can influence the long-term stability of e-nose, a self-calibration model is presented also for combination with the proposed ANN ensemble. Experimental results demonstrate the effectiveness of the proposed models.

5. Acknowledgment

6. References

[1] A. Berna. "Metal Oxide Sensors for Electronic Noses and Their Application to Food Analysis". Sensors 10, 2010, 3882-3910.

[2] C. Di Natale, A. Macagnano, F. Davide, A. D'Amico, R. Paolesse, T. Boschi, M. Faccio, G. Ferri. "An electronic nose for food analysis". Sensors and Actuators B: Chemical 44(1-3), 1997, 521-528.

[3] A.H. Gomez, J. Wang, G.X. Hu, A.G. Pereira. "Monitoring storage shelf life of tomato using electronic nose technique". Journal of Food Engineering 85, 2008, 625-631.

[4] H. Yu, J. Wang. "Discrimination of Long Jing Green-tea grade by electronic nose". Sensors and Actuators B: Chemical 122, 2007, 134-140.

[5] S. Ampuero, J.O. Bosset. "The electronic nose applied to dairy products: a review". Sensors and Actuators B: Chemical 94, 2003, 1-12.

[6] L. Zhang, F.C. Tian, C. Kadri, et al., "Gases Concentration Estimation using Heuristics and Bio-inspired Optimization Models for Experimental Chemical Electronic Nose". Sensors and Actuators B, Chemical 160(1), 2011, 760-770.

[7] L. Zhang, F.C. Tian, H. Nie, et al., "Classification of multiple indoor air contaminants by an electronic nose and a hybrid support vector machine". Sensors and Actuators B: Chemical 174, 2012, 114-125.

[8] S. De Vito, E. Massera, M. Piga, L. Martinotto, G. Di Francia, "On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario". Sensors and Actuators B: Chemical 129, 2008, 750-757.

[9] L. Zhang, F.C. Tian, S. Liu, et al., "Chaos based neural network optimization for Concentration Estimation of Indoor Air Contaminants". Sensors and Actuators A: physical 189, 2013, 161-167.

[10] J.W. Gardner, H.W. Shin, E.L. Hines. "An electronic nose system to diagnose illness". Sensors and Actuators B: Chemical 70(1-3), 2000, 19-24.

[11] A. D'Amico, G. Pennazza, M. Santonico, E. Martinelli, C. Roscioni, G. Galluccio, R. Paolesse, C. Di Natale. "An investigation on electronic nose diagnosis of lung cancer". Lung Cancer 68(2), 2010, 170-176.

[12] L. Zhang, F.C. Tian, C. Kadri, et al., "On-Line Sensor Calibration Transfer Among Electronic Nose Instruments for Monitoring Volatile Organic Chemicals in Indoor Air Quality". Sensors and Actuators B 160(1), 2011, 899-909.