A Combinative Function Approximation Model and Its Applications to Electronic Noses

Gao Daqi, Tong Zhen, Li Yongli
Department of Computer Science
State Key Laboratory of Bioreactor Engineering
East China University of Science and Technology
Shanghai 200237, PR China

Abstract: ***This paper focuses on combinative and modular approximation models to simultaneously estimate odor classes and strengths. We first decompose a many-to-many approximation task into multiple many-to-one tasks, and then realize them using multiple many-to-one approximation models. A single model is regarded as an expert, and a panel or ensemble is made up of multiple such experts. Each expert is either a multivariate logarithmic regression model, or a multilayer perceptron (MLP), or a support vector machine (SVM). A panel is on behalf of a kind of odor. The most similar panel gives the class label and strength of an odor. The experiment for estimating 4 kinds of fragrant materials shows that the proposed model is effective.***

1. INTRODUCTION

Experts can recognize over 4,000 odors, but they are often poor to estimate odor strengths [1-2]. Compared with that, capabilities of electronic noses are not worth mentioning. Up to now, electronic noses are only able to recognize limited kinds of odors, or estimate concentrations of no many single-component odors, but few reported for simultaneously determining multiple odor classes and concentrations in the literature [1-8], which is far from man's expectation. It is reasonable to make electronic noses undertake such tasks. Some of the main obstacles are that the sensitivity and selectivity of gas sensors are not high enough, and the classification and approximation capabilities of analysis methods are not ideal. This paper focuses on the latter.

In the early stage of electronic nose development, some data processing methods in common use were multivariate regression [3], partial least square (PLS) [1-2], cluster analysis (CA) [4], principle component analysis (PCA) [4-5], discriminant function analysis (DFA) [1-2], template matching [1-2], etc. Today, artificial neural networks (ANNs) [1-2, 6-8], support vector machine (SVM) [9], independent component analysis [10], and the others, are often used.

From a long-term point of view, electronic noses will carry out such tasks as simultaneously estimating odor classes and strengthens. The problems of this kind can be really treated as many-to-many function approximation ones. Unfortunately, up to now there are few reports about successfully solving high dimensional and multi-output approximation problems in the literature [11]. In principle, a many-to-many multilayer perceptron (MLP) is able to approximate any complicated function with any accuracy under condition that it has sufficient hidden nodes and long learning time. So are many-to-many polynomial functions and SVMs. In fact, neither of them can well finish such tasks alone. One of main difficulties is that a continuous function always provides an output for an arbitrary input.

The tasks of simultaneously estimating odor classes and concentrations are really multiple function approximation ones. In order to implement such tasks, two following methods can be employed. (*A*). Transform a many-to-many approximation problem into multiple classification problems. The shortcoming to do like that is that too many classifiers are needed. It is especially impractical when a great quantity of strengths in a kind of odor needs to be estimated. (*B*). Take a type of many-to-many or multiple many-to-one function approximation models to simultaneously fit the strength curves of multiple odors one after another. However, the two methods are not ideal in effect. The main drawback of the former is that a multi-output approximation model is often of complicated structure, low approximation accuracy, and even long learning time. And the main defect of the latter is that an approximation model may give self-contradictory results for a response of the sensor array, because a single approximation model doesn't know to say "No". For example, what can we do if one single-output MLP says that an odor sample is 100 *ppm* (parts per million) of methanol, and another says it is 1000 *ppm* of ethanol?

The concept of classifier combination or ensemble was presented in 1990s [12-15], because the capability of a single kind of classifiers is limited after all. For solving the large-scale learning problems, task decompositions and modular classifiers [16-17] are proposed. If each module is regarded as an expect, we can determine the class label of a sample according to the majority vote, sum, product, max, min, and the other combinative rules [12-15]. With the help of combinative idea, this paper decomposes a complicated many-to-many approximation task into multiple simple many-to-one tasks, and correspondingly, proposes a kind of combinative and modular approximation models for simultaneously estimating many kinds of odor classes and concentrations.

This paper is organized as follows. In Section II, we briefly introduce an experimental electronic nose. The task decomposition method, the structures of each modular approximation components as well as the combination strategies, are formulated in Section III. In Section IV, the presented method is applied to simultaneously estimating classes and strengths for 4 kinds of fragrant materials, namely as ethanol, ethyl acetate, ethyl caproate and ethyl lactate, 21 kinds of concentrations altogether. Finally, Section V comes to our conclusions.

2. EXPERIMENTAL

Figure 1 is the schematic diagram of a practical electronic nose, which is mainly made up of a test box, a thermostatic cup and a personal computer (PC). The test box consists of a

0-7803-9048-2/05/$20.00 ©2005 IEEE

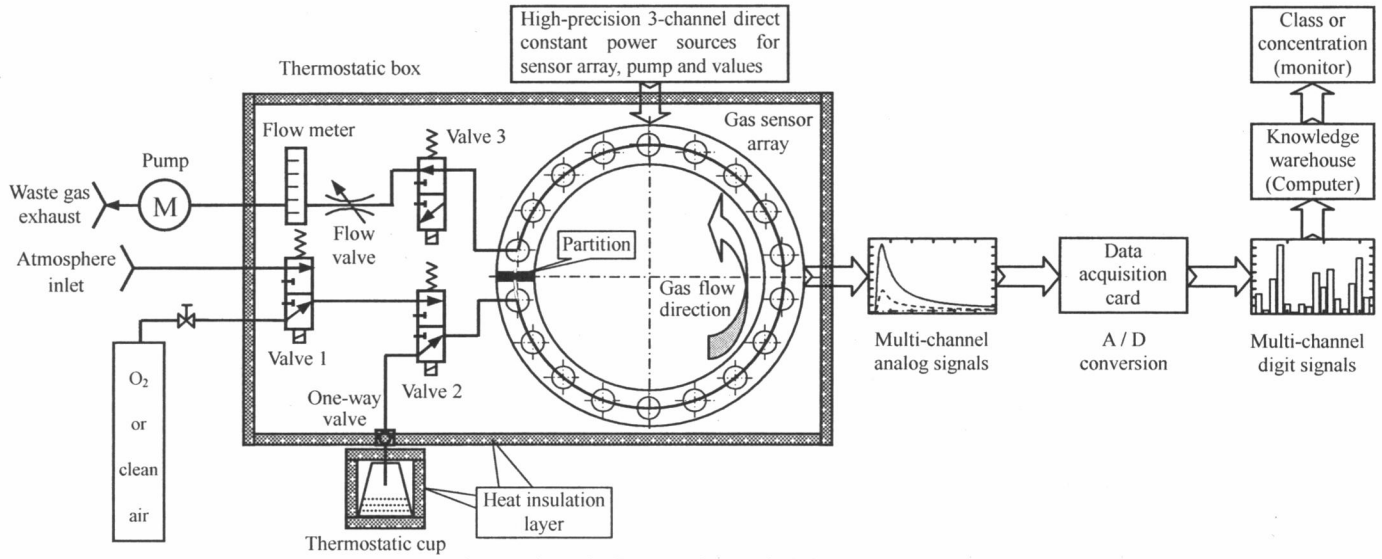


Fig. 1. Schematic diagram of a practical electronic nose

gas sensor array, 3-channel high-precision direct constant power source (5±0.01V for the sensor heaters, 10± 0.001V for the sensor working circuits, 12±0.1V for the others), gas sampling and flow control units, waste gas exhauster, thermostatic unit, etc. The sensor array is composed of 16 TGS sensors, namely TGS800, 812, 813, 816, 821, 822, 823, 824, 825, 826, 830, 831, 832, 842, 880, 883T, all provided by Figaro Engineering Inc., Japan. The array is set up within a 200-*ml* circular chamber. A 16-bit high-precision data acquisition card is inserted in the expansion slot of the PC motherboard [18].

In order to ensure the same measurement condition, the following steps are adopted.

(A) Purge the sensor array by O_2 or dry air in the speed of 60 *mm*/*sec* for 40 *sec* before sampling.

(B) Pour 30-*ml* liquid sample in a 250 *ml* flask and keep it within the thermostatic cup for 30 to 60 *min*. The cup is wrapped with heat insulation material and heated by a semiconductor heater. The thermostatic accuracy is 55± 0.1°*C*.

(C) Keep the sensor array, the inlet pipes and three electro-magnetic values in a thermostatic box. The test box is wrapped with heat insulation materials and heated by a resistance heater. The thermostatic accuracy is 55±0.1°*C*.

A PC controls all the operating procedure. What an operator needs to do is only to load or unload an awaiting sample according to the PC's order. Fig. 2 is the flow chart of a sampling procedure. The maximum value from the response curve of a sensor is taken as a feature component, namely x_{pi} in Fig. 2. Because the array is made up of 16 sensors, the input dimensionality is 16. Such a 16-dimensional pattern is then transformed into an odor's class, strength, concentration or the other results in agreement with human perceptions by means of appropriate classification and nonlinear mapping methods.

3. Combinative Function Approximation Models

3.1 *Combinative and modular approximation models*

In order to simultaneously estimate the odor classes and concentrations, it is quite ideal to employ the divide-and-conquer and combinative strategies. A panel or expert ensemble is on behalf of a kind of odor and made up of several experts, and naturally the number of panels is equal to the number of kinds of odors. The most consistent panel ought to be taken to determine the class label and concentration of a sample. For the purpose of clarity, let us suppose that there are three panels and three experts in each ensemble. For sample $\boldsymbol{x}$, the predicted results of the three panels are shown in *Table* I. The relative standard deviation (RSD) of *Panel* 1 is only 4.08%, the smallest among three panels, or the three members in the first ensemble have the most identical views, so one can make such a decision that $\boldsymbol{x}$ belongs to the odor represented by *Panel* 1 and its concentration is 1000 *ppm*.

As stated above, a single many-to-many approximation model is unable to effectively cope with multi-output approximation problems. Therefore, we have to resort to multiple single-output approximation models. Does a many-

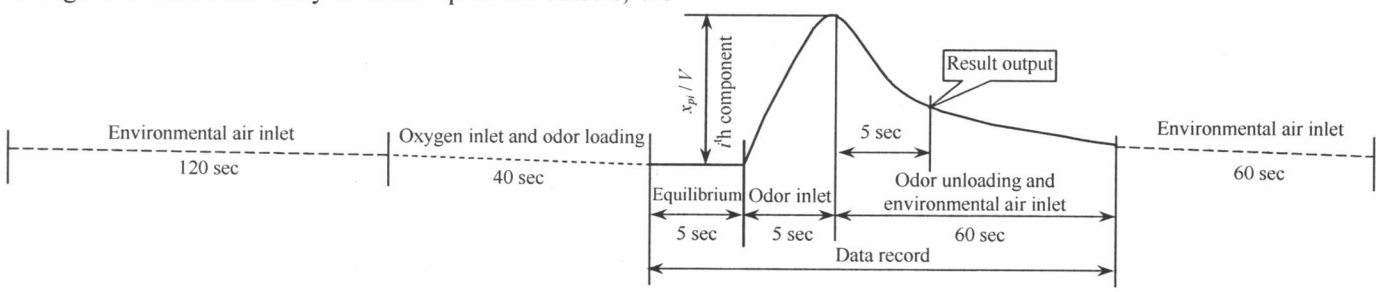


Fig. 2. Flow chart of a sampling procedure

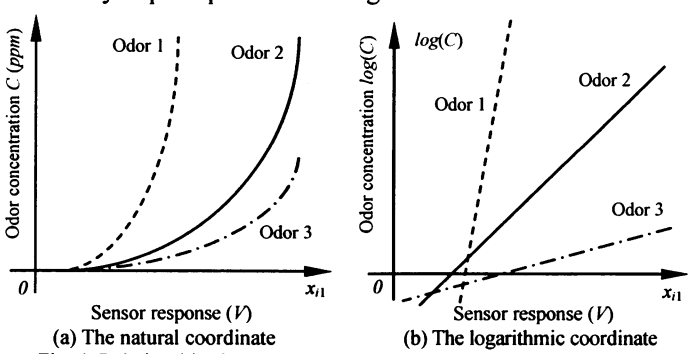
Table I. Predicted concentrations of sample x by three 3-member panels (*ppm*)

Predicted result	Panel 1			Panel 2			Panel 3		
	Expert 1	Expert 2	Expert 3	Expert 1	Expert 2	Expert 3	Expert 1	Expert 2	Expert 3
Predicted conc.	1000	1050	950	1500	2500	500	500	2000	500
Average conc.	1000			1500			1000		
RSD (%)	4.08			54.43			70.71		

to-one approximation model trained with the dataset from odor j have a real output for sample x from odor q? Yes! And the reverse is true. For instance, three many-to-one MLPs, called MLP1, MLP2 and MLP3, are used to predict the class and concentration of x. Maybe MLP1 takes x for a 1000-*ppm* ethanol, MLP2 for a 750-*ppm* methanol, and MLP3 for a 1500-*ppm* acetone. At the moment we are unable to determine x's label and concentration. The embarrassing situation shows that the class and concentration of x can be not certainly decided by only a single type of many-to-one approximation models, which looks just like the lack of selectivity of a single sensor. So we turn to multiple many-to-one approximation models for help.

Our basic ideas are as follows. A panel consists of several many-to-one approximation models, and stands for a specific kind of odor. And n ensembles are behalf of n kinds of odors, as shown in Fig. 3. In a panel, all members, or all many-to-one approximation models, are used to simultaneously imitate the total behaviors of the sensor array for a specified kind of odors. If x is indeed from odor j, the outputs of all the approximation models in ensemble j will simultaneously be close in on a certain concentration point; otherwise their outputs will be quite divergent. In this way, we can determine the class label and concentration of a sample using multiple expert panels, not by only a single type of approximation models, according to a certain kind of combinative rule. In other words, the expert panel with the highest similarity determines the class label of an undecided sample, and the average output of all experts in the panel decides its concentration or strength.

The relationship between concentrations of an odor with only a kind of changing component and responses of the sensor array is a hyper-dimensional curve, and the points on the curve are in the name of odor strengths. The change of responses of the sensor array means the change of odor strengths. A lot of experiments show that the relationships are close to logarithmic [1-2], as shown in Fig. 4, thus multivariate logarithmic regressions (MVLR) models can be used to fit them. Because the changes of logarithmic curves are not fiercely undulate [1-8], it is suitable to use single-hidden-layer perceptrons with sigmoid activation functions



(a) The natural coordinate (b) The logarithmic coordinate

Fig. 4. Relationships between sensor responses and odor concentrations

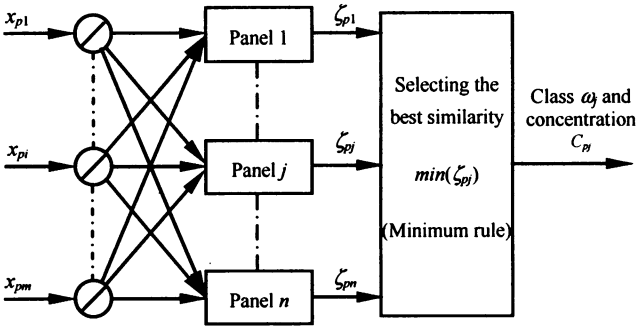


Fig. 3. A combinative and modular FAM

and SVMs with radial basis function (RBF) kernels to execute such function approximation tasks. Therefore, this paper selects MVLRs, MLPs with sigmoid activation functions, and SVMs with RBF kernels to play the roles of experts, as shown in Fig. 5.

Let $x_p=(x_{p1},x_{p2},\ldots,x_{pi},\ldots,x_{pm})^T\in R^m$ be the preprocessed response of a sensor array for the pth sample, which is taken as the real input of an approximation model, and x_{pi} is the response of the ith sensor, or the ith input component of the approximation model. The real output of panel j for x_p is the similar degree ζ_p of a known odor j, given according to the sum and average combinative rules, as shown in Fig. 5.

Suppose there are r experts in a panel j, all experts in the panel fit the relationship between the response x_p of a sensor array and the real concentrations C_{pj} of odor j. In the learning stages, the target outputs $d_{pj}=f(C_{pj})$ of all experts in the panel j are the same, namely the logarithmic concentration values of odor j. At the moment, only the samples from odor j take part in learning. In the test stages, for a certain sample x_p, if ζ_{pj} is the smallest among the similar degrees given by all panels, one can decide that x_p belongs to odor j and its concentration is the function $C_{pj}=f^{-1}(\bar{y}_{pj})$ of the average output of all experts in panel j.

Because an approximation model always gives an output for any input, no matter how absurd it may be, this paper presents the following combinative and modular approximation models to simultaneously estimate odor classes and concentrations.

(1).Decompose a multi-input multi-output approximation task $f: X\rightarrow D$ into multiple multi-input single-output ones $f: X^{(j)}\rightarrow d_j$. Here $X\in R^{N\times m}$, $X^{(j)}\in R^{Nj\times m}$, $D\in R^{N\times n}$ and $d_j\in R^{Nj\times 1}$, N is the total number of learning samples. Each training subset is only composed of N_j samples from odor j, and learned by the corresponding many-to-one approximation model.

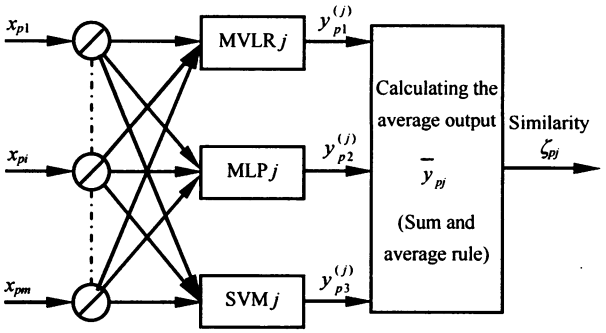


Fig. 5. Block diagram of approximation model panel j

(2).Use an MVLR j, an MLP j, and an SVM j, all with the m-to-1 structures, to form a panel j to simultaneously fit the m-to-1 curve f: $X^{(j)} \rightarrow d_j$. In panel j, the target outputs of all the compositions are the same, the preprocessed concentration values in a certain interval.

(3).Predetermine the concentration of a sample by all experts in panel j according to the average combinative rule:

$$\bar{y}_{pj} = \frac{1}{r}\sum_{k=1}^{r} y_{pk}^{(j)} \tag{1}$$

Here, $y_{pk}^{(j)}$ is the real output of the kth expert in panel j, k=1,2,...,r. The similarity to odor j given by panel j is:

$$\zeta_{pj} = \frac{1}{\bar{y}_{pj}}\sqrt{\frac{1}{r}\sum_{k=1}^{r}\left(y_{pk}^{(j)} - \bar{y}_{pj}\right)^2} \times 100\% \tag{2}$$

In fact, ζ_{pj} is similar to the RSD and $\bar{y}_{pj}$ to the mean value.

(5).Determine the final class label ω_j and concentration C_{pj} of an undecided odor x_p through comparing all the panels with the help of the minimum combinative rule:

$$x_p \in \omega_j \quad and \quad C_{pj} = f^{-1}\left(\bar{y}_{pj}\right)$$
$$if \quad \zeta_{pj} = \min_{1\le k\le n}\left(\zeta_{pk}\right) \quad j = 1,2,\ldots,n \tag{3}$$

Here f^{-1} shows a certain inverse transformation between the real concentration C_{pj} and the predicted value $\bar{y}_{pj}$.

3.2 *Multivariate logarithmic regressions*

For a kind of odor j, an MVLR is really to realize a many-to-one mapping f:$X^{(j)} \rightarrow d_j$, $X^{(j)} \in R^{Nj\times m}$, and $d_j \in R^{Nj\times 1}$, namely

$$\log\left(y_1^{(j)}\right) = \alpha^T X^{(j)} + \alpha_0 + \varepsilon_1 = \log\left(d_j\right) + \varepsilon_1 \tag{4}$$

in order to minimize the sum-of-squared error ε_1. Here, $\alpha \in R^m$ is a coefficient vector, α_0 a constant, and $y_1^{(j)}$ the output of MVLR j. Obviously, there are m+1 undecided values.

3.3 *Single-hidden-layer perceptrons*

Figure 6 is the detailed structure of an m-s-1 MLP, which also works for a many-to-one mapping f:$X^{(j)} \rightarrow d_j$, the same as an MVLR does. Here, s is the number of hidden nodes. All MLPs use the batch-learning back-propagation algorithm to adjust their weights and biases. The learning process of an MLP expert j is to minimize the sum-of-squared error E_j between d_j and the real output $y_2^{(j)}$, or

$$E_j = \left\|d_j - y_2^{(j)}\right\|_2^2 = \frac{1}{2}\sum_{p=1}^{N_j}\left(d_{pj} - y_{p2}^{(j)}\right)^2 = min \tag{5}$$

We know that variable scales and activation functions have great influences on the generalization and convergence of MLPs [21]. In our experiment, the activation functions in the hidden and the output layers are $f(\varphi)$=3(1+exp(-φ/3))$^{-1}$, all the input variables are scaled in proportion to the range of [0.0, 6.0], and the target outputs to the range of [0.0, 3.0] [21].

It is not feasible to directly change the concentration values into the range of [0.0, 3.0], because that will further lead to lessen the expected distinction between low-concentration samples. Therefore, we suppose the concentration of distilled water is 1.0×10^{-7}, namely 0.1 *ppm*, and apply the common logarithmic transformation to the real concentrations. Let the real concentration be C_{pj} (*ppm*) for sample x_p, the target of MLP j is given by:

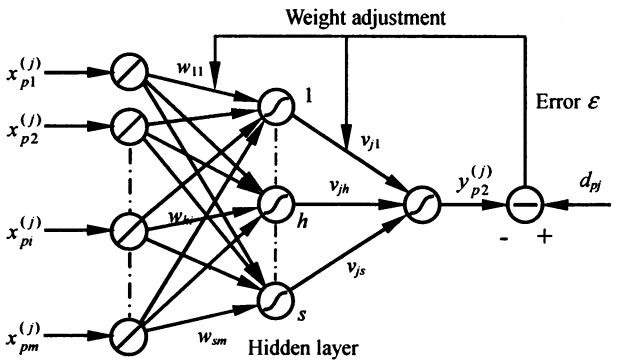


Fig. 6. An m-s-1 perceptron expert j

$$d_{pj} = \left(1 + \log\left(C_{pj}\right)\right) * 3.0/5.0 \tag{6}$$

In consideration of the cut-off and saturated states of sigmoid activation functions, the range of target outputs should be further cut down, namely d_{pj}=0.1 corresponds to the distilled water (C_{pj}=0.1 *ppm*) and d_{pj}=2.9 to C_{pj}=10,000 *ppm*. Therefore, (6) is revised to be:

$$d_{pj} = (1 + \log\left(C_{pj}\right)) * 2.8/5.0 + 0.1 \tag{7}$$

In addition, we must take the following two situations into accounts:

(A) If the predicted concentration of sample x_p by MLP j is over 10,000 *ppm*, x_p is considered not to belong to class ω_j. Similarly, an MLP steps down from decision if its real output is over 2.95.

(B) If the lower limit of sensitivity of a sensor for an odor is C_θ, the corresponding input component is set to be 0.

3.4 *Support vector machines*

We use the RBF kernel in each SVM expert:

$$k\left(x, x_p\right) = \exp\left(-\gamma\left\|x - x_p\right\|^2\right) \tag{8}$$

Here, γ is a width parameter determined artificially.

The real output of the jth SVM expert is

$$y_3^{(j)} = f\left(x, X_{\text{sup}}^{(j)}, \beta^{(j)}\right) = \sum_{s=1}^{N_j}\beta_s^{(j)} k\left(x, x_s^{(j)}\right) + b^{(j)} \tag{9}$$

Here, $X_{\text{sup}}^{(j)}$ is the support vector matrix, $\beta_s^{(j)}$ a constant of the sth kernel related to the support vector $x_s^{(j)}$, and $b^{(j)}$ is a constant, too. All the constants are determined through learning.

The preprocessing means of dataset for training an SVM expert is just the same as done in the corresponding MLP expert, omitted here. The learning process of an SVM expert is also evaluated by the sum-of-squared error E_j.

4. Experimental Results

Four kinds of fragrant materials are ethanol, ethyl acetate, ethyl caproate and ethyl lactate, and each of them is diluted with distilled water into 4-6 kinds of concentrations, as shown in Table II.

The measurement method is just the same as stated in *Section* 2. Fifty samples for each concentration are collected, and 40 of them are used as a part of the training set, and the others as a part of the test set. Therefore there are 40×21=840 samples in the training set, and 10×21=210 patterns in the test set. Fig. 7 shows the difference of responses between

Table II. Concentrations of fragrant materials in water (*ppm*)

Ethanol	Ethyl acetate	Ethyl caproate	Ethyl lactate
Distilled water	Distilled water	Distilled water	Distilled water
		1	
100	10	5	100
2500	100	10	2500
5000	1000	100	5000
7500		1000	7500
10000	10000	10000	10000

distilled water and 1.0 *ppm* ethyl caproate in liquid. Fig. 8 is the PCA result for all the original data set. The reproducibility of data is relatively good. Obviously the fitted curves are different from each other. The linear regression method fails to solve the problems effectively, and the multivariate polynomial regression method will give some unreasonable results.

There are 4 panels and 3 experts in each panel. The number of input dimensions is 16, because there are 16 sensors in the array and 4 kinds of odors.

The multivariate logarithmic regression experts are realized with MATLAB6.2. Their learning time is within 1 *sec* (PIV860 CPU 256 RAM, the same below). Fig. 9 is the predicted result for ethyl caproate with 70 testing samples, and at the moment, the root-mean-squared (RMS) error of the MVLR expert is 0.152.

For all the MLP experts, let the number of hidden nodes be 5, the learning factor η=0.02, and the momentum parameter α=0.075. All the 16-5-1 MLPs learn 15,000 epochs and take 17.39, 14.65, 20.51, 17.58 sec, respectively. Fig. 10 shows the predicted results of the 16-5-1 MLP expert for ethyl caproate, and the predicted RMS error is 0.031. The approximation and predicted performance of the MLP expert is satisfactory.

Let the width parameter γ=1/2, all the learning time of the 4 SVM experts is within 30 *sec*. The number of support vectors in each expert is 165, 124, 215, and 161 in sequence. Fig. 11 gives the learning and predicted result for ethyl caproate with all 350 samples, and the RMS error is 0.010, which is quite satisfactory.

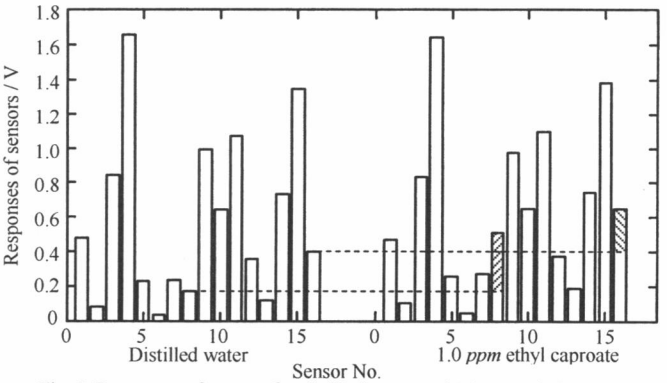


Fig. 7. Responses of sensors for distilled water and 1.0 *ppm* ethyl caproate

Relatively speaking, the approximation and predicted accuracy of an SVM is the highest among the 3 experts, shown in Fig. 9 to Fig.11. From the angle of memory requirement for test, an SVM expert has not any advantage, which needs to memory too many support vectors, compared with an MVLR or MLP expert. It has been made clear in *Section* 3.1 that we can not depend upon any kind of single-output approximation model to determine the label and concentration of an odor sample. Therefore, the predicted results of the above 3 experts shown in Fig. 9 to Fig.11 only mean to be a little closer to the memorized concentration points than those of the other experts do. In other words, the predicated results of a single kind of many-to-one approximation models are not so convincing.

The simultaneous class and strength estimation correct rate of the proposed combinative approximation models is 208/210=99.05% for the test set. Only two 1-*ppm* samples from ethyl caproate are mistaken for distilled water. In fact, any single-kind of many-to-many approximation models is

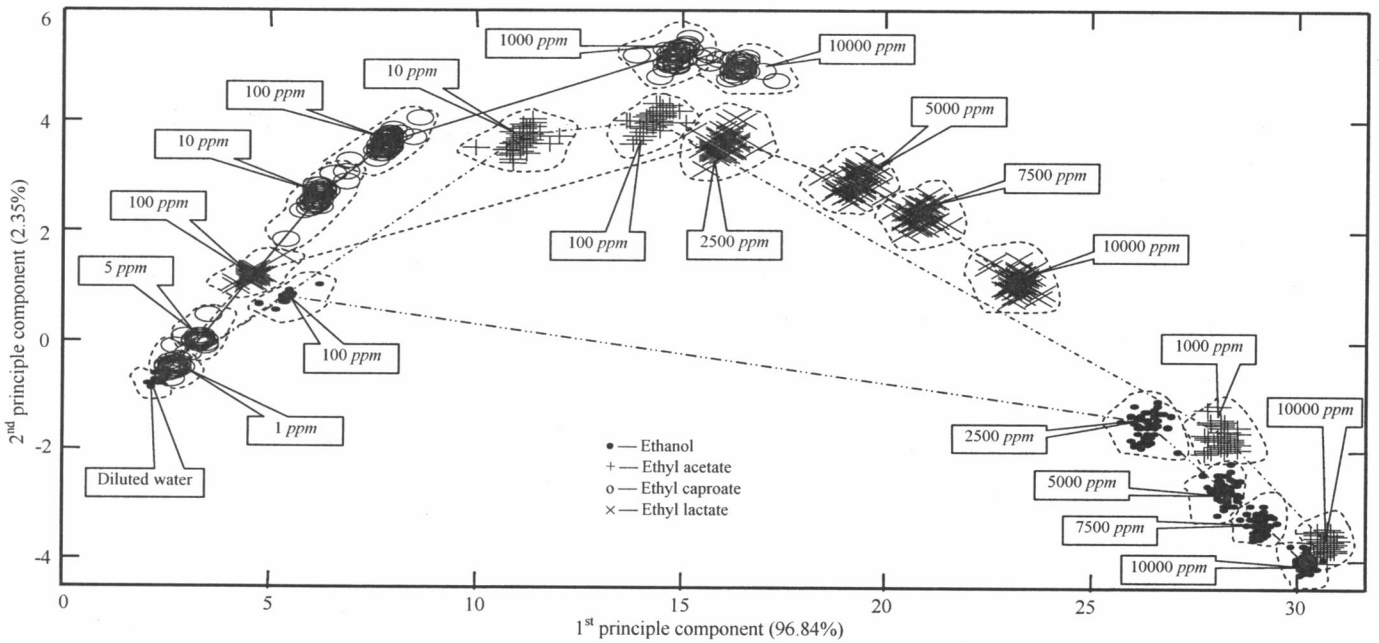


Fig. 8. Principle component analysis for the original data set

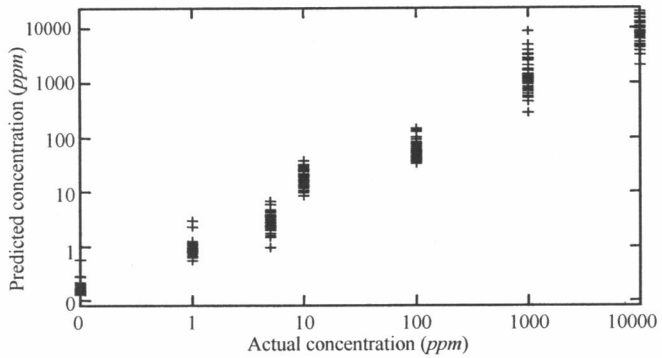


Fig. 9. Learning and predicted results for ethyl caproate by the multivariate logarithmic regression expert

not able to obtain high approximation accuracy. For example, a 16-8-4 MLP cannot converge to the required accuracy even after 50,000 iterating epochs. The predicted correct rates of the 2 many-to-many function approxiamtion models, namely a 16-to-4 MVLR and an MLP, are only 47.31%, 62.38%, resprectively.

5. Conclusions

This paper decomposes a many-to-many approximation task into multiple many-to-one approximation tasks, and takes multiple many-to-one experts to simultaneously implement such approximation tasks. We employ MVLRs, MLPs and SVMs as the basic components of panels, and use the average and minimum combinative rules to determine the final labels and strengths of odors. An experiment for estimating 4 kinds of fragrant materials in 21 grades of concentrations shows that the proposed combinative approximation model is quite effective for simultaneously determining the labels and strengths of multiple odors, so improves the recognition performances of electronic noses. With the combinative approximation models, an electronic nose can well execute the simultaneous estimation tasks of many kinds of odor classes and concentrations on condition that a kind of odor is either single-component or only a kind of changeable component.

Acknowledgment

This work is supported by the National Science Foundation of China (NSFC) under Grant No. 60275017, 60373073, the Key Science and Technology Development Foundation of Shanghai, China, under Grant No. 04DZ05010, and the open project program of the State Key Laboratory of Bioreactor Engineering.

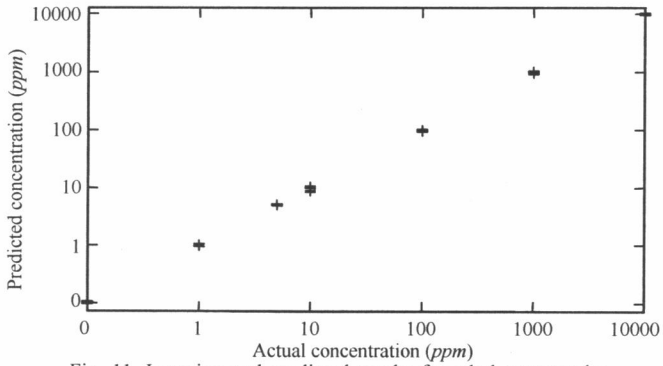


Fig. 11. Learning and predicted results for ethyl caproate by the support vector machine expert

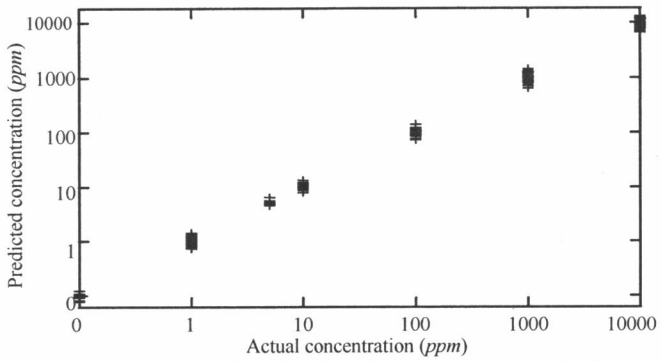


Fig. 10. Learning and predicted results for ethyl caproate by the 16-5-1 perceptron expert

References

[1] T.C. Pearce, S.S. Schiffman, H.T. Nagle, and J.W. Gardner (Eds), Hand of machine olfaction, Wiley-VCH Press, 2003.

[2] J.W. Gardner, P.N. Bartlett, Electronic noses: principles and applications, Oxford University Press, 1999.

[3] M.S. Nayak, R. Dwivedi, S.K. Srivastava, Application of iteration technique in association with multiple regression method for identification of mixtures of gases using an integrated gas-sensor array, *Sens. Actuators B: Chem.* 1994, 21: 11-16.

[4] J.W. Gardner, Detection of vapours and odours from a multisensors array using pattern recognition, part I. Principal component and cluster analysis, *Sens. Actuators B: Chem.* 1991, 4: 109-115.

[5] S. Capone, M. Epifani, F. Quaranta, et al, Monitoring of rancidity of milk by means of an electronic nose and a dynamic PCA analysis, *Sens. Actuators B: Chem.* 2001, 78: 174-179.

[6] G. Huyberechts, P. Szecowka, J. Roggen, et al, Simultaneous quantification of carbon monoxide and methane in humid air using a sensor array and an artificial neural network, *Sens. Actuators B: Chem.* 1997, 45: 123-130.

[7] A.K. Srivastava, Detection of volatile organic compounds (VOCs) using SnO_2 gas-sensor array and artificial neural network, *Sens. Actuators B: Chem.* 2003, 96(1-2): 24-37.

[8] Z. Haoxian, M.O. Balaban, J.C. Principe, Improving pattern recognition of electronic nose data with time-delay neural networks, *Sens. Actuators B: Chem.* 2003, 96(1-2): 385-389.

[9] C. Distante, N. Ancona, P. Siciliano, Support vector machines for olfactory signals recognition, *Sens. Actuators B: Chem.* 2003, 88: 30-39.

[10] C.D. Natale, E. Martinelli, A.D. Amico, Counteraction of environmental disturbances of electronic nose data by independent component analysis, *Sens. Actuators B: Chem.* 2002, 82: 158-165.

[11] C.M. Bishop, Neural networks for pattern recognition, *Clarendon Press, Oxford*, 1998.

[12] L.K. Hansen, P. Salamon, Neural network ensembles, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(10): 993-1001.

[13] J. Kittier, M. Hatef, R.P.W. Duin, et al. On combining classifiers, *IEEE Transactions on PAMI*, 1998, 20(3): 226-239.

[14] D. Windridge, J. Kittler, A morphologically optimal strategy for classifier combination: multiple expert fusions as a tomographic process, *IEEE Transactions on PAMI*, 2003, 25(3): 343-353.

[15] C.A. Shipp, L.I. Kuncheva, Relationships between combination methods and measures of diversity in combining classifiers, *Information Fusion*, 2002, 3: 135-148.

[16] G. Uda, M. Amel, Modular neural network classifiers: a comparative study, *Journal of Intelligent and Robotic Systems*, 1998, 21: 117-129.

[17] R. Anand, K.G. Mehrotra, C.K. Mohan, et al, Efficient classification for multiclass problems using modular neural networks, *IEEE Transactions on Neural Networks*, 1995, 6(1): 117-124.

[18] G. Daqi, W. Shuyan, J. Yan, An electronic nose and modular radial basis function network classifiers for recognizing multiple fragrant materials, *Sens. & Actuators B*, 2004, 97(2-3): 391-401.

[19] R.O. Duda, P.E. Hart, D.G. Stork, Pattern classification (2nd ED), John wiley & Sons Inc., 2001.

[20] G. Thimm, P. Moerland, E. Fiesler, The interchangeability of learning rate and gain in backpropagation neural networks. *Neural Computation*, 1996, 8(2): 451-460.

[21] G. Daqi, Y. Genxing, Influences of variable scales and activation functions on the performances of multilayer feedforward neural networks, *Pattern Recognition*. 2003, 36 (4): 869-878.