Sensors and Actuators B xxx (2011) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Sensors and Actuators B: Chemical

journal homepage: www.elsevier.com/locate/snb

Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose

Daqi Gao*, Fangjun Liu, Ji Wang

State Key Laboratory of Bioreactor Engineering, Department of Computer Science, East China University of Science and Technology, Shanghai 200237, China

ARTICLE INFO

Article history:
Received 1 June 2011
Received in revised form 4 October 2011
Accepted 2 November 2011
Available online xxx

Keywords:
Hierarchical models
Discrimination
Quantification
Odors
Electronic nose

ABSTRACT

This paper studies hierarchical discrimination and quantification models in order to simultaneously quantify multiple kinds of odors with an improved electronic nose. Such tasks are first regard as multiple discrimination tasks and then as multiple quantification tasks, and implemented by the hierarchical models with the divide-and-conquer strategy. The discrimination models are the common classifiers, including nearest neighbor classifiers, local Euclidean distance templates, local Mahalanobis distance templates, multi-layer perceptrons (MLPs), support vector machines (SVMs) with Gaussian or polynomial kernels. Similarly, the quantification models are multivariate linear regressions, partial least squares regressions, multivariate quadratic regressions, MLPs, SVMs. We developed several types of hierarchical model and compared their capabilities for quantifying 12 kinds of volatile organic compounds with the improved electronic nose. The experimental results show that the hierarchical model composed of multiple single-output MLPs followed by multiple single-output MLPs with local decomposition, virtual balance and local generalization techniques, has advantages over the others in the aspects of time complexity, structure complexity and generalization performance.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Odors cannot be seen by eyes and felt by hands. People usually use some vague terms to describe their characteristics, such as strong, weak, stimulative, fragrant, top-quality, second-rate, normal, abnormal, sweet, foul, and abominable [1,2]. We can depict one kind of odor by means of another. For example, an odorous sample may be said to be like orange or banana, but what is the orange or the banana odor? Such questions are very difficult to be answered. Odors ever smelled can remain in our memory and be brought to our mind by imagination, but cannot be quantitatively compared by data or recode. It is not an easy thing to quantify odors only by our olfaction, namely our noses. Accordingly, sensory evaluation of odor qualities is not objective and fair enough, even if given by experts. Under the circumstances, electronic noses (ENs) emerge as the times require [3].

Electronic noses have a wide range of applications. After sensing odors with a gas sensor array and analyzing the resulting data by means of appropriate pattern recognition methods, an EN can determine products' classes, grades and freshness; distinguish genuine from sham; control production processes; readjust prescriptions; monitor environmental pollutions, etc. The application objects include perfumes [4–7], milks and teas [8–10], alcoholic beverages [11–13], fruits [14], fishes and meats [15–17], environmental air [18], water [19,20], medical treatments [21], drugs [22], warfare agents [23,24], and even bloods and bacteria [25,26]. Can we find an absolutely odorless material?

There are many kinds of odors in the natural world, and often tens, hundreds and even thousands of components in one kind. For example, there exist about 50 main aromatic components in a brewing alcoholic drink. If the aroma of the drink changes, an EN is required to judge which components change and how much they change. Currently, ENs are limited in capabilities to carry out the real-time quantitative analysis of odors [1,4–26]. Therefore, there is an urgent need to find suitable pattern recognition methods to both discriminate and quantity multiple kinds of odors, simple or complex [1].

An unfavorable case for odor quantification is that the lower the concentrations of odors, the smaller the differences between them. In other words, different sorts of low-concentration odors may be close to each other in the measure space. The relationships between strengths of multiple kinds of odors and their components may be multiple complex curved surfaces, which may intersect with each other. Therefore, the discrimination and quantification of multiple kinds of odors will bring about great difficulties and challenges to the existing pattern recognition methods, including neural

* Corresponding author at: Department of Computer Science, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, China. Tel.: +86 21 6425 3780; fax: +86 21 6425 2984.
E-mail address: gaodaqi@8163.net.cn (G. Daqi).

0925-4005/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.snb.2011.11.003

Please cite this article in press as: D. Gao, et al., Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose, Sens. Actuators B: Chem. (2011), doi:10.1016/j.snb.2011.11.003

networks [1,7,14,18,27] and support vector machines (SVMs) [28,29].

A single prediction model with multiple output units will give multiple predicted values for a specific odorous sample. Can we thus say that the sample belongs both to an odor ω_A with a concentration ρ_A and to another odor ω_B with a concentration ρ_B? It is unallowable. In other words, a single prediction model with multiple output units fails to quantify multiple kinds of odors, regardless simple or complex [1,2,30–34]. In order to recognize many kinds of odors and quantify their concentrations as well, the following approaches are available.

(A) A single multi-output (SMO) discrimination model [1] or a SMO discrimination model followed by multiple multi-output (MMO) discrimination models [30,31] is used. The two solutions actually consider the quantification task as a pure discrimination one, and a concentration point as a class. The disadvantages of these solutions include complicated structures, long learning time and serious imbalance between classes when many kinds of odors and many concentrations exist. Consequently, these two models are only suitable for a small number of classes of odors with limited concentrations [30,31].
(B) An SMO discrimination model followed by multiple single-output (MSO) quantification models, called an SMO–MSO model, or two groups of MSO quantification models in cascade, called an MSO–MSO model [32], is employed. These solutions consider the quantification task first as multiple discriminative tasks and then as multiple quantitative tasks. However, the defects that exist in case (A) will still appear when there are too many kinds of odors and concentration points.

Because the existing pattern recognition models are quite limited in their capabilities for identifying many kinds of odors and quantifying their concentrations as well [30–34], this paper devotes to studying hierarchical models and finding out the appropriate model to accomplish such tasks. The remainder of this paper is organized as follows: Section 2 illustrates an improved electronic nose and introduces its working principle. In Section 3, we propose the structure of a hierarchical model and possible component units. Furthermore, we analyze the time complexities, structure complexities and generalization performances of several hierarchical models. Section 4 presents the experimental results for quantifying 12 kinds of volatile organic compounds (VOCs). Finally, Section 5 comes to our conclusions.

2. Experimental

Fig. 1 shows an improved electronic nose [35], which consists of a test box, a personal computer (PC), six headspace vapor generators and a clean air cylinder. The test box mainly contains a thermostatic chest, an automatic lift device, a sampling needle, a miniature diaphragm vacuum pump, three two-positional two-way electromagnetic valves, a flow meter, two flow valves, a 4-channel direct current (DC) source as well as control and measure circuits. The array, which is installed within the circular chamber in the thermostatic chest, is composed of 16 TGS gas sensors, namely TGS800, 812, 813, 816, 821, 822, 823, 824, 825, 826, 830, 831, 832, 842, 880, 883T, all provided by Figaro Inc., Japan. The load resistor of each gas sensor is fixed at 10 kΩ.

The samples, liquid or solid, and headspace vapors are kept at the constant temperature of $42 \pm 0.1\,^{\circ}C$ for 30 min before measured, and the chest always kept at $55 \pm 0.1\,^{\circ}C$. In order to get good repeatability, a sample of 10 ml and its headspace vapor is measured only one time, and a glass flask of 200 ml to hold the sample and thus generate the vapor is also used only once. The responses of gas sensors are limited to the range (0.0, 10.0 V) by hardware.

The gas sensor array is calibrated by clean air before sampling. The clean air from the cylinder passes through the flow valve 2, the electromagnetic valve 3, the outlet, the interior and the inlet of the circular chamber, and the needle in sequence before exhausted into the atmosphere, at the flow rate of 600 ml/min. During the course, the electromagnetic valve 3 is on while the other two are off. Consequently, the gas sensors exactly recover to their preliminary state.

The working principle of the improved electronic nose is described as follows. While sampling, under the roles of the PC and the automatic lift device, the headspace vapor generator rises, and the sampling needle fixed under the inlet of the chamber thus contacts the headspace vapor. With the aid of the miniature diaphragm pump, the vapor in the flask is drawn into the circular chamber where the gas sensor array is mounted in at the flow rate of 600 ml/min, forced to skim across the sensitive films of gas sensors, and finally exhausted into the air at the waste gas outlet. Along with the flow of vapors, the gas sensors produce analogous sensitive responses, which are then converted into digital by the data acquisition card and stored in the PC as a data file. For a sample measured, a 16-dimensional response vector is thus gotten, called a pattern $\boldsymbol{x} \in R^{16}$ hereafter, because the maximum steady-state response of a gas sensor is regarded as a variable.

The purpose of the experiment is to discriminate twelve kinds of VOCs, ethanol, butanol, hexanol, ethyl acetate, ethyl propionate, ethyl butyrate, ethyl valerate, ethyl caproate, ethyl heptanoate, ethyl octanoate, ethyl lactate, and isoamyl acetate, and quantify their concentrations as well by measuring their headspace vapors with the improved electronic nose. These VOCs are the main fragrant components in brewing alcoholic drinks. They are diluted with distilled water into required concentrations.

3. Hierarchical models and their component units

3.1. Structure of hierarchical models

The divide-and-conquer strategy is an effective approach for the discriminative and quantitative analysis of multiple kinds of odors [30–34]. Fig. 2 illustrates a hierarchical model for implementing such tasks. Seen from the horizontal direction, a pair of hierarchical modules represents a specific kind of odor. The former module is responsible for discrimination, whose role is to separate the represented odor from the others, and the followed one is in charge of quantification, whose role is to predict the strengths and concentrations of the represented odor. If there are n kinds of odors, there are n pairs of modules, one for one. Seen from the vertical direction, there exist two parallel columns of modules. The modules in the first column are for discrimination, and those in the second column are for quantification.

3.2. Discrimination models

Because of the nonlinear distributions of odors in the measure space, the classical linear discriminant analysis (LDA) is not adopted [11,36]. The radial basis function (RBF) networks [9,11] and fuzzy inference models [1,37] are not included because they are not often employed for odor discrimination in the electronic noses. k-Nearest-Neighbor (k-NN) classifiers, Euclidean or Mahalanobis distance templates, multi-layer perceptrons (MLPs) and SVMs are able to form nonlinear decision boundaries and commonly used in the electronic noses [1,11,29–34]. They are thus chosen as the component units of discrimination modules of the hierarchical model shown in Fig. 2. In the following subsections, we will introduce the

Please cite this article in press as: D. Gao, et al., Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose, Sens. Actuators B: Chem. (2011), doi:10.1016/j.snb.2011.11.003

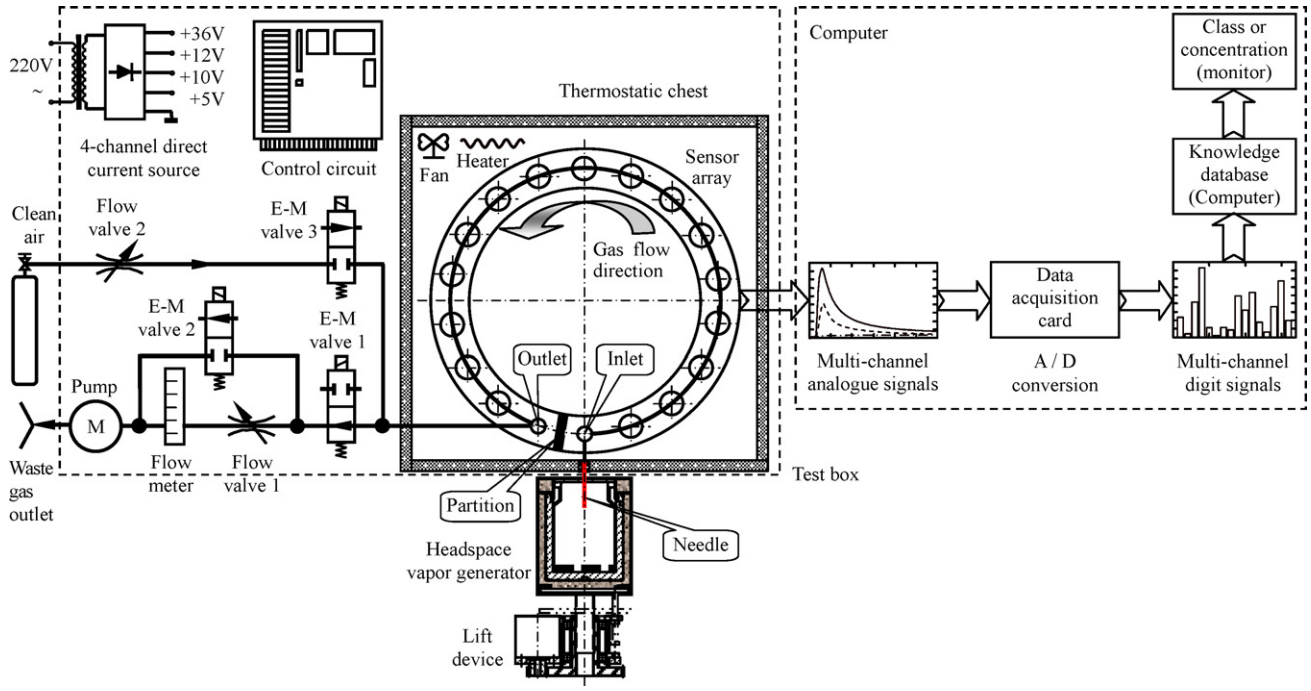


Fig. 1. Schematic diagram of an improved electronic nose.

principles and algorithms of these selected models, as well as their advantages and disadvantages.

3.2.1. Nearest neighbor classifiers

A 1-*NN* classifier with $k = 1$ needs no training process. While assigning a label to a test pattern $\boldsymbol{x} \in R^m$, however, the 1-*NN* classifier has to calculate N Euclidean distances from $\boldsymbol{x}$ to N training patterns $\boldsymbol{X} \in R^{N \times m}$ [36] and find the nearest neighbor from them. Here, $m = 16$ is the number of input dimensions or gas sensors. Naturally, all the N patterns in $\boldsymbol{X}$ must be stored as prototypes. Therefore, the time and structure complexities of 1-*NN* classifiers are high when N is large.

3.2.2. Local Euclidean distance templates

A Euclidean distance template is sometimes called a fingerprint or polar or radar plot [1,13,36]. In order to discriminate multiple kinds of odors with changing concentrations, we treat a concentration as a class and thus propose a type of local Euclidean distance (LED) template. The decision rule is to seek the minimum Euclidean distance between $\boldsymbol{x}$ and all the means $\boldsymbol{\mu}_k^{(j)}$ ($j = 1, 2, \ldots, n$). Here, n is

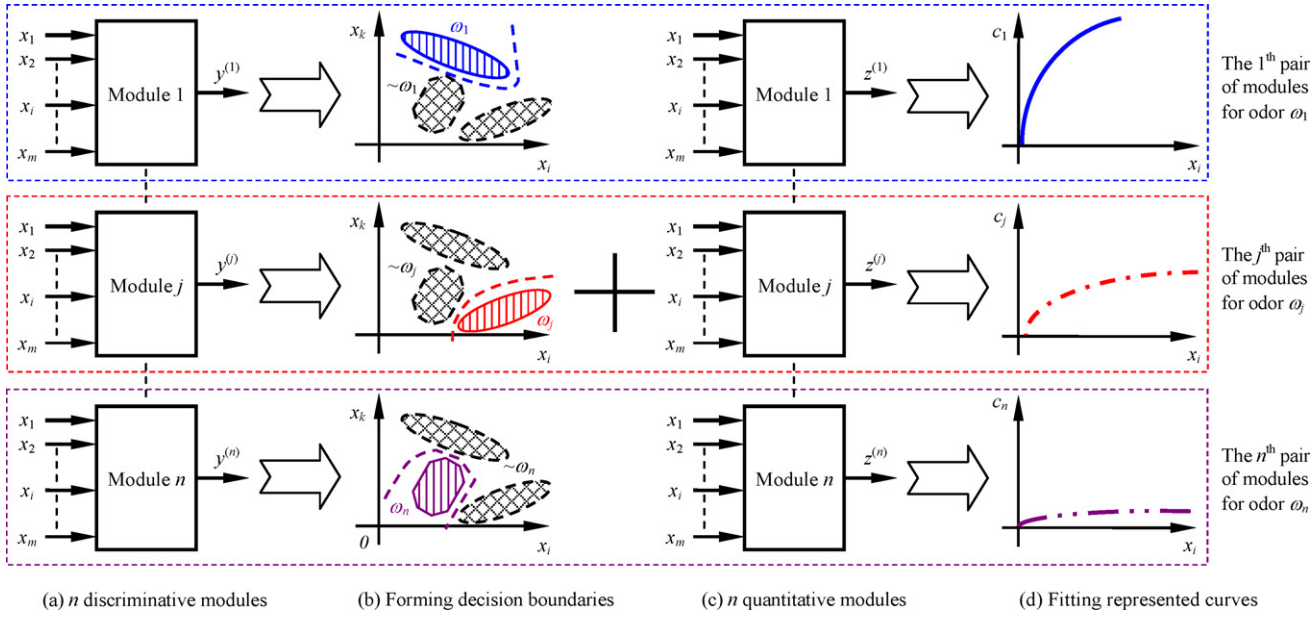


Fig. 2. Hierarchical models for discriminative and quantitative analysis of multiple kinds of odors.

Please cite this article in press as: D. Gao, et al., Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose, Sens. Actuators B: Chem. (2011), doi:10.1016/j.snb.2011.11.003

the number of classes of odors, and $\boldsymbol{\mu}_k^{(j)}$ is the mean vector of the training patterns $\boldsymbol{X}_k^{(j)}$ from the kth concentration point of odor ω_j.

3.2.3. Local Mahalanobis distance templates

A local Mahalanobis distance (LMD) template [1,33,36] is the natural generalization of an LED model. The decision rule is to seek the minimum Mahalanobis distance between $\boldsymbol{x}$ and all the centers $\boldsymbol{\mu}_k^{(j)}$. Usually, an LMD model is better in discrimination performance than an LED model. This kind of model requires to store all the mean vectors $\boldsymbol{\mu}_k^{(j)}$ and all the symmetrical covariance matrices $\boldsymbol{\Sigma}_k^{(j)}$. Here, $\boldsymbol{\Sigma}_k^{(j)}$ is determined by $\boldsymbol{X}_k^{(j)}$. The model will be no longer in force when a certain covariance matrix is singular.

3.2.4. Single-output perceptrons

Single-hidden-layer perceptrons have good capabilities to solve nonlinearly separable problems [1,27]. Each single-output perceptron with a hidden layer solves a two-class recognition problem, and its hidden and output nodes are with sigmoid activation function $f(\varphi) = 3(1 + \exp(-\varphi/3))^{-1}$ [38]. In order to improve the learning and generalization performances of perceptrons, we adopt and perfect a local one-against-all (OAA) decomposition method, a virtual balance step and a local generalization strategy [38].

Fig. 3 illustrates the formation of the jth training subset and the local generalization region for the jth perceptron. We first construct two concentric oblique ellipsoids, an initial ellipsoid Θ_{j0} and an extended ellipsoid Θ_j, the centers of which are just the mean vector $\boldsymbol{\mu}^{(j)}$ of all the training patterns $\boldsymbol{X}^{(j)}$ from ω_j. Θ_{j0} just encloses $\boldsymbol{X}^{(j)}$, whose max Mahalanobis radius $r_{\max}^{(j)} = d_M(\boldsymbol{x}, \omega_j)$ is with respect to $\boldsymbol{\mu}^{(j)}$ and the covariance matrix $\boldsymbol{\Sigma}^{(j)}$ of $\boldsymbol{X}^{(j)}$. Let $N^{(j)}$ and $N^{(\sim j0)}$ be the numbers of patterns from ω_j and all the others $\sim\omega_j$ encircled within Θ_{j0}, a suitable radius of Θ_j is $R_{\min}^{(j)} = (1 + N^{(j)}/N^{(\sim j0)})r_{\max}^{(j)}$. All the patterns inside Θ_j form the final two-class training subset $\boldsymbol{\Omega}^{(j)} = \{\boldsymbol{X}^{(j)}, \boldsymbol{X}^{(\sim j)}\}$. The final number of patterns is $N^{(j)} + N^{(\sim j)}$, and the imbalanced ratio is $\lambda_j = \max(N^{(j)}, N^{(\sim j)})/\min(N^{(j)}, N^{(\sim j)}) \geq 1$.

In order to achieve a virtual balance for the unbalanced subset $\boldsymbol{\Omega}^{(j)}$, a good solution is to add some virtual patterns to the minority side, say $\boldsymbol{X}^{(j)}$. In the specific programming realization, for the pth training pattern $\boldsymbol{x}_p$ from ω_j, we can directly multiply the weigh update increment $\Delta\boldsymbol{w}_p^{(j)}(\tau)$ by an enlargement coefficient β_j at the τth iterative epoch, that is

$$\Delta\boldsymbol{w}^{(j)}(\tau) = -\beta_j \sum_{x_p \in X^{(j)}} \frac{\partial E_j(\tau)}{\partial \boldsymbol{w}_p^{(j)}(\tau)} - \sum_{x_p \in X^{(\sim j)}} \frac{\partial E_j(\tau)}{\partial \boldsymbol{w}_p^{(j)}(\tau)} \tag{1}$$

where $E_j(\tau)$ is the root-mean-square (RMS) error of perceptron j. An appropriate value β_j is given by

$$\beta_j = \begin{cases} 1 + \dfrac{\lambda_j - 1}{3}, & \text{if} \quad 1 \leq \lambda_j \leq 4 \\ \dfrac{\lambda_j}{2}, & \text{otherwise} \end{cases} \tag{2}$$

In order to realize the local generalization, the action of every perceptron must be limited in a reasonable range through the correction of its outputs. Let the real output of perceptron j for $\boldsymbol{x}$ be $y^{(j)}$, the corrected output $\rho^{(j)}$ is

$$\rho^{(j)} = \begin{cases} y^{(j)} \exp\left(-\ln(2) \times \dfrac{r^{(j)}}{R_{\min}^{(j)}}\right), & \text{if} \quad (r^{(j)})^2 = (\boldsymbol{x} - \boldsymbol{\mu}^{(j)})^T (\Sigma^{(j)})^{-1} (\boldsymbol{x} - \boldsymbol{\mu}^{(j)}) \geq (r_{\max}^{(j)})^2 \\ y^{(j)}, & \text{otherwise} \end{cases} \tag{3}$$

The parameters that must be noted down include the class mean vectors, the class covariance matrixes, the radii of oblique ellipsoids, and all the weights and thresholds of perceptrons.

3.2.5. Support vector machines with Gaussian kernels

SVMs for discrimination decompose an n-class problem into n two-class ones and make decision according to the max rule [28]. An SVM with Gaussian kernels solves a two-class recognition problem $\{\omega_j, \sim\omega_j\}$, whose target outputs are coded as $\{+1, -1\}$. A Gaussian kernel is expressed as

$$k(\boldsymbol{x}, \boldsymbol{x}_p) = \exp\left(-\frac{||\boldsymbol{x} - \boldsymbol{x}_p||^2}{2\sigma^2}\right) = \exp(-\gamma||\boldsymbol{x} - \boldsymbol{x}_p||^2) \tag{4}$$

where $\boldsymbol{x}_p$ is called the pth support vector (SV), σ the width of kernels, and $\gamma = \sigma^{-2}/2$ the width parameter which has to be selected artificially. Hereafter, an SVM with Gaussian kernels is written as an SVM-G for short.

SVMs-G must memorize all the SVs, all the Lagrangian multipliers, and the widths of kernels.

3.2.6. Support vector machines with polynomial kernels

A polynomial kernel is expressed as

$$k(\boldsymbol{x}, \boldsymbol{x}_p) = (\langle \boldsymbol{x}, \boldsymbol{x}_p \rangle + 1)^d = (\boldsymbol{x}^T \boldsymbol{x}_p + 1)^d \tag{5}$$

where d is called the degree of polynomials, which must be also selected in advance [28].

SVMs with polynomial kernels are as a whole the same as those with Gaussian kernels except the type of kernels. Henceforth, an SVM with polynomial kernels is written as an SVM-P for short. SVMs of this kind will store all the SVs, all the Lagrangian multipliers, and the degrees of polynomials.

3.3. Quantification models

A quantification module j represents a specific kind of odor ω_j, and realizes the mapping $f: \boldsymbol{X}^{(j)} \to \boldsymbol{t}^{(j)}$, where $\boldsymbol{t}^{(j)}$ is the expected output vector. The target and predicted outputs of the regression curves often take the logarithmic concentrations [1,37]. The criteria to choose models for quantification are that (A) they are widely employed in the electronic noses and (B) they have simple structures and good capabilities to approximate nonlinear curves. The component units of quantification modules are multivariate linear regressions (MVLRs), partial least squares regressions (PLSRs), multivariate quadratic regressions (MVQRs), MLPs, SVMs, as shown in the second column of the hierarchical model in Fig. 2.

In order to predict the possible concentration points between distilled water and the lowest concentration points in each class, the patterns from distilled water are regarded as the common point of all the curves, and their expected concentrations are enforced to be 0.1 ppm. A consideration to do like this is that the TGS gas sensors used have lower sensitive limitations of 1.0 ppm and over [1].

In the following subsections, we will introduce on the principles, algorithms, as well as advantages and disadvantages of these mentioned models.

3.3.1. Multivariate linear regression

The target output of the jth MVLR model $z^{(j)} = (\boldsymbol{\beta}^{(j)})^T x + \boldsymbol{\beta}_0^{(j)}$ takes the logarithm base 10 of the expected concentrations $C_p^{(j)}$ in ppm plus 1.0, i.e., $t_p^{(j)} = \varphi(C_p^{(j)}) = 1.0 + \log_{10} C_p^{(j)}$, where $\boldsymbol{\beta}^{(j)} \in R^m$ is a coefficient vector and $\beta_0^{(j)}$ a constant. All the parameters are determined by the minimum squared-error (MSE) method [1,36,37].

Please cite this article in press as: D. Gao, et al., Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose, Sens. Actuators B: Chem. (2011), doi:10.1016/j.snb.2011.11.003

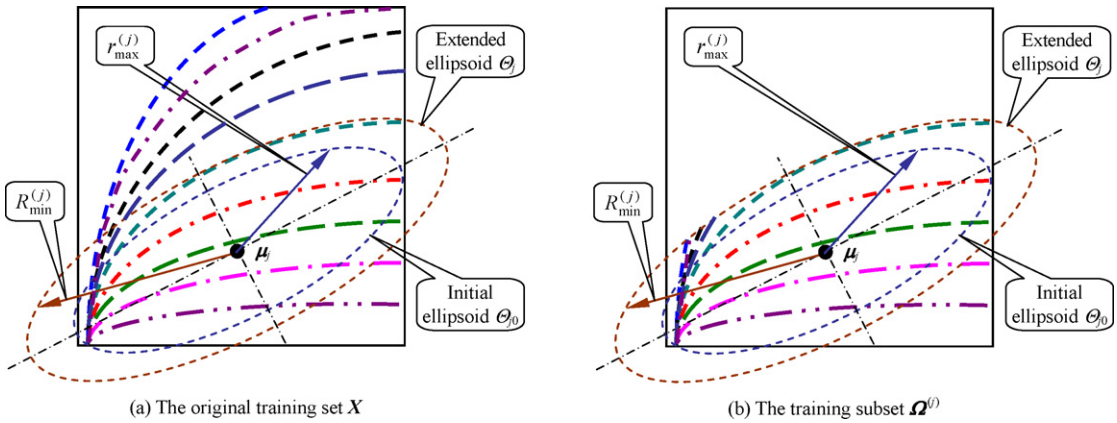


Fig. 3. Formation of the *j*th economic training subset for the *j*th perceptron classifier.

An MVLR model only needs to store $1+1/m$ equivalent patterns.

3.3.2. Partial least squares regression [39]

As the extended version of principal component regression (PCR), an PLSR model *j* first seeks the projection direction $\boldsymbol{w}_1^{(j)}$ corresponding to the maximum eigenvalue $\lambda_1^{(j)}$ of the inner product of the normalized sensor response matrix $\boldsymbol{X}_0^{(j)}$ and the normalized logarithmic concentration vector $\boldsymbol{t}_0^{(j)}$, i.e., $\boldsymbol{w}_1^{(j)} = (\boldsymbol{X}_0^{(j)})^T \boldsymbol{t}_0^{(j)} / ||(\boldsymbol{X}_0^{(j)})^T \boldsymbol{t}_0^{(j)}||$. Then, the one-variable regressions between the first principal component vector $\boldsymbol{p}_1^{(j)} = \boldsymbol{X}_0^{(j)} \boldsymbol{w}_1^{(j)}$ and $\boldsymbol{X}_0^{(j)}$ and $\boldsymbol{t}_0^{(j)}$ are done simultaneously. The regression coefficients are calculated by

$$\boldsymbol{q}_1^{(j)} = \frac{(\boldsymbol{X}_0^{(j)})^T \boldsymbol{p}_1^{(j)}}{(\boldsymbol{p}_1^{(j)})^T \boldsymbol{p}_1^{(j)}} \quad \text{and} \quad r_1^{(j)} = \frac{(\boldsymbol{t}_0^{(j)})^T \boldsymbol{p}_1^{(j)}}{(\boldsymbol{p}_1^{(j)})^T \boldsymbol{p}_1^{(j)}}.$$

Next, the one-variable regressions are applied to the residual matrix $\boldsymbol{X}_1^{(j)} = \boldsymbol{X}_0^{(j)} - \boldsymbol{p}_1^{(j)} (\boldsymbol{q}_1^{(j)})^T$ and the residual vector $\boldsymbol{t}_1^{(j)} = \boldsymbol{t}_0^{(j)} - r_1^{(j)} \boldsymbol{p}_1^{(j)}$, by which the 2nd projection direction $\boldsymbol{w}_2^{(j)}$, the 2nd principal component vector $\boldsymbol{p}_2^{(j)}$, and the 2nd regression coefficients $\boldsymbol{q}_2^{(j)}$ and $r_2^{(j)}$, are obtained accordingly.

Repeating the above steps, the predicted equations of the *j*th PLSR model for $\boldsymbol{x}$ are generated as

$$\begin{cases} p_1 = \boldsymbol{x}^T \boldsymbol{w}_1^{(j)}, \quad \ldots, \; p_k = (\boldsymbol{x}_{k-1})^T \boldsymbol{w}_{k-1}^{(j)} \\ z_1^{(j)} = r_1^{(j)} p_1 + r_2^{(j)} p_2 + \cdots + r_k^{(j)} p_k \end{cases} \tag{6}$$

where the index *k* is the computational step or the number of principal components. The computation process stops when the rate of Frobenius norms of the residual matrix $\boldsymbol{X}_k^{(j)} = \boldsymbol{X}_{k-1}^{(j)} - \boldsymbol{X}_{k-2}^{(j)}$ to the original matrix $\boldsymbol{X}_0^{(j)}$ is below 1.0%.

An PLSR model stores $1+1/m$ equivalent patterns because its linearly regressive nature.

3.3.3. Multivariate quadratic regression

The *j*th MVQR model is expressed as

$$z^{(j)} = \sum_{i=1}^{m} \boldsymbol{\alpha}_i^{(j)} (x_i^{(j)})^2 + \sum_{i=1}^{m} \boldsymbol{\beta}_i^{(k)} x_i^{(j)} + \boldsymbol{\beta}_0^{(j)}$$

where $\boldsymbol{\alpha}^{(j)}$ and $\boldsymbol{\beta}^{(j)}$ are two coefficient vectors determined by the above-mentioned MSE method. Except the coefficient vector for the quadratic term, an MVQR is completely the same as an MVLR. An MVQR model stores $2+1/m$ equivalent patterns.

Generally, the fitting capabilities of MVQRs are slightly better than those of MVLRs and PLSRs. It is worthy noting that the three mentioned models are only approximately valid for a small range of concentrations [1,36,39].

3.3.4. Single-output perceptrons

Single-hidden-layer perceptrons with sufficient number of hidden nodes have good capabilities to approximate any nonlinear curves [1,27]. The hidden and output nodes of single-output perceptrons are with sigmoid activation function $f(\varphi) = 3(1+\exp(-\varphi/3))^{-1}$ [38]. The target outputs of perceptron *j* are scaled to the real range [0.05, 2.95] by

$$t_p^{(j)} = \varphi\left(c_p^{(j)}\right) = \frac{\left(1.0 + \log_{10} c_p^{(j)}\right) \times 2.9}{5.0} + 0.05 \tag{7}$$

The learning algorithm is the classical back-propagation (BP) algorithm [27]. Let the real output of perceptron *j* be $z^{(j)}$ for $\boldsymbol{x}$, the predicted concentration is

$$c^{(j)} = 1.0 \times 10^{\left(\frac{5.0\left(z^{(j)} - 0.05\right)}{2.9} - 1.0\right)}.$$

Each perceptron for quantification only needs to store its weights and thresholds.

3.3.5. SVMs with Gaussian or polynomial kernels

SVMs with Gaussian or polynomial kernels for quantification are the same as their correspondents for discrimination in learning algorithms. Of course, all the new SVs, the Lagrangian multipliers, and the pre-selected widths of Gaussian kernels or the pre-selected degrees of polynomials, have to be stored. SVMs for quantification use the same target outputs as MVLRs do.

Now, we have discussed the component units that are suitable for the hierarchical models. To summarize, the hierarchical models may be a 1-*NN* classifier followed by *n* MVLRs; *n* LED templates followed by *n* MVQRs; *n* LMD templates followed by *n* MVLRs or PLSRs or MVQRs; *n* MLPs followed by *n* MLPs; *n* SVMs followed by *n* SVMs. In the next section, we will compare their performances through experiments.

4. Experimental results

The experiment aims at comparing the time complexities, structure complexities and generalization performances of the proposed hierarchical models for identifying and quantifying 12 kinds of VOCs. The common concentration points in the training and the test set are 1, 5, 10, 50, 100, 500, 1000, 5000 and 10,000 ppm. The

Please cite this article in press as: D. Gao, et al., Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose, Sens. Actuators B: Chem. (2011), doi:10.1016/j.snb.2011.11.003

Table 1
The concentration points arranged for 12 kinds of volatile organic compounds in the training and test set (ppm).

Dataset	Common concentration	Added concentration
The training set	1[a], 5[a], 10[b], 50, 100, 500, 1000, 5000, 10,000	
The test set	1[a], 5[a], 10[b], 50, 100, 500, 1000, 5000, 10,000	25[b], 75, 250, 750, 2500, 7500

[a] Only for ethyl caproate.
[b] Except ethanol and ethyl lactate.

specific concentration points only in the test set are 25, 75, 250, 750, 2500 and 7500 ppm. Limited by the sensitivity of sensors, the points of 1 and 5 ppm are only for ethyl caproate, and the points of 10 and 25 ppm are not for ethanol and ethyl lactate. Table 1 gives the detailed concentrations.

There are 84 concentration points, plus distilled water, 85 in total, in the training set, and 85 + 70 = 155 concentration points in the test set. For each VOC, 60 samples are made up and measured for each common concentration point. 50 patterns among them are randomly selected as a part of the training data and the remains, i.e., 10 patterns, as a part of the test data. For the concentration points only existing in the test set, 10 samples are made up and measured. As a result, the total numbers of patterns are 4250 in the training set and 1550 in the test set, respectively. Statistical terms of gas sensor responses for the training data are that the maximum standard deviation (MSD) is 0.2483 V, which occurs at ethyl acetate of 10,000 ppm measured by TGS832, and the maximum relative standard deviation (MRSD), i.e., the MSD divided by the mean, is 7.28%, which occurs at ethyl caproate of 1 ppm detected by TGS825. Fig. 4(a) and (b) shows the results of the first two principal component analyses. In the measure space, the 12 kinds of VOCs may be nonlinearly separable from each other because of the changing concentrations, and the relationships between gas sensor responses and odor concentrations are also nonlinear.

The input variables are transformed into the range (0, 1) by dividing the maximum steady-state response values of gas sensors by 10, except changed into the range [0.0, 6.0] in proportion for perceptrons or normalized for PLSRs. In the following experiments, MLPs with the BP algorithm are developed in C language, SVMs are realized with the existing software – Libsvm [40], and the others are programmed with Matlab 7.0.

Table 2 summarizes the learning and identifying results of the 6 discrimination models, namely 1-*NN* classifiers, LED templates, LMD templates, MLPs, SVMs with Gaussian kernels, and SVMs with polynomial kernels, as shown in the first column of the hierarchical model in Fig. 2.

The LED and LMD templates have the discrimination rates (DRs) of 99.93% and 100% for the training set, respectively, and their learning time is short, below 13.4 s (given by a PC with 2.6G CPU, 1.0G RAM, the same below). A 1-*NN* classifier has no training process. However, the DRs of the three models are only between 65.29% and 67.10% for the 1550 test patterns.

The structural and learning parameters of 12 single-output MLPs for discrimination are preset as follows. The number of hidden nodes is $s = 8$, the max iterative epoch $\tau_{max} = 15{,}000$, and the allowable RMS error $\varepsilon^* = 0.05$. The patterns from distilled water are adopted as the common point of all the odors. Two additional agreements are that (A) one sample is considered wrong if the real outputs from two to eleven MLPs are beyond 1.5 and (B) one sample is labeled as distilled water if all the real outputs of the 12 MLPs are larger than 1.5. According to Section 3.2.4, the sizes of training subsets are between 1514 and 2867, with an average of 2127.08, i.e., 50.05% of the original training set. The proper learning factors η_j are between 3 and 5 times of reciprocal numbers of patterns, i.e., between 0.0010 and 0.0033. Using the optimal learning subsets, virtual balance step and local generalization strategy together, the resulting DRs are 100% for the training and 1485/1550 = 95.81% for the test set.

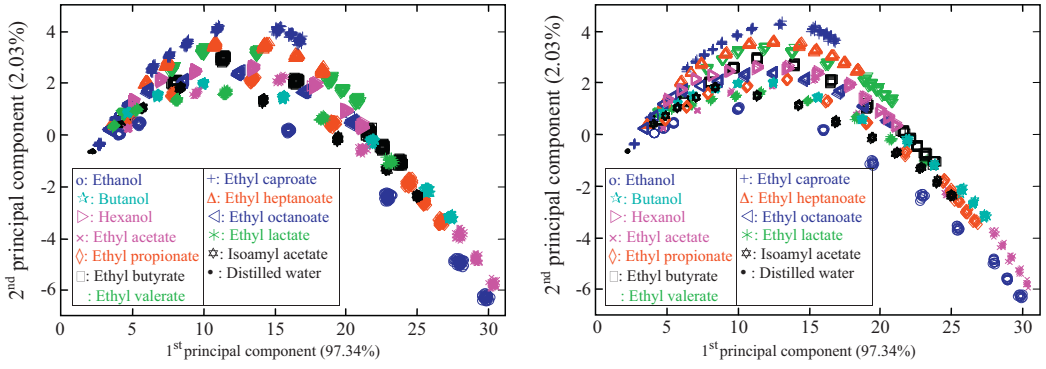
The parameters of SVMs for discrimination are preset to be $\gamma = 0.03$, $d = 3$ and $C = 1000$. The two mentioned additional agreements for the MLPs are still valid here, but the discriminative thresholds are changed from 1.5 in MLPs to 0. The SVMs with Gaussian and polynomial kernels reach the DRs of 81.74% and 75.27% for the training set, and 80.45% and 72.52% for the test set, respectively. These results show that the generalization performances of SVMs for discrimination are not as good as expected.

According to Table 2, the MSO perceptrons take the longest learning time among the discrimination models. Based on the evaluation of the predicted time, structure complexity and generalization performance, however, the MLPs are the best, and the SVMs are the second best.

Next we employ MVLRs, PLSRs, MVQRs, MLPs, SVMs with Gaussian kernels, and SVMs with polynomial kernels, to predict the concentrations of the 12 kinds of VOCs.

Table 3 gives the learning and predicted results of the 6 models for quantification. In the table, MSDs are specific to concentration points, and average relative deviations (ARDs) are specific to odor categories, while maximum absolution deviations (MADs) and maximum relative deviations (MRDs) are specific to single samples.

Because the rates of Frobenius norms of the residual matrixes to the original are below 1.0% when $k = 4$, all the 12 PLSR models are formed by the first 4 principal components. The MVLR, PLSR and MVQR models perform poorly in the learning and



(a). The full training set with 4250 samples (b). The test set with 1550 samples

Fig. 4. Principal component analysis of the training and the test set with 12 kinds of volatile organic compounds.

Please cite this article in press as: D. Gao, et al., Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose, Sens. Actuators B: Chem. (2011), doi:10.1016/j.snb.2011.11.003

Table 2
Learning and predicted results given by six types of discrimination models for 12 kinds of volatile organic compounds.

Model		1-NN	LED	LMD	MLP	SVM-G	SVM-P
The training set	Learning time (*s*)	–	10.11	13.40	2232.27	1221.45	1385.71
	No. mislabeled samples	–	3	0	0	776	1051
	Discrimination rate (%)	100	99.93	100	100	81.74	75.27
The test set	Decision time (*s*)	179.33	3.74	4.94	0.18	1.06	1.74
	No. mislabeled samples	523	538	510	65	303	426
	Discrimination rate (%)	66.26	65.29	67.10	95.81	80.45	72.52

Table 3
Learning and predicted results given by six regression models for quantifying 12 kinds of volatile organic compounds.

Regression model		MVLR	PLSR	MVQR	MLP	SVM-G	SVM-P
The training set	Learning time (s)	1.99	2.51	2.06	213.33	137.82	159.66
	Max standard deviation (ppm)	5630.55	5532.46	3580.43	806.29	736.00	2258.88
	Max absolute deviation (ppm)	25109.30	12901.80	11442.80	1985.71	2246.73	8535.32
	Average relative deviation (%)	37.49	32.61	31.56	5.57	17.63	23.22
	Max relative deviation (%)	600.55	254.88	190.71	56.15	34.97	85.35
The test set	Predicted time (s)	1.22	1.38	1.32	0.18	1.52	2.26
	Max standard deviation (ppm)	5521.27	2997.63	2535.67	568.15	624.26	1965.24
	Max absolute deviation (ppm)	16437.20	8479.10	5493.47	1714.66	2928.70	9591.60
	Average relative deviation (%)	30.09	53.59	62.64	9.58	32.65	20.68
	Max relative deviation (%)	200.38	334.45	809.78	35.30	160.83	95.92

predicted accuracies in spite of the short learning time, about 2.0 s. For example, the MSDs and MADs given by the MVLR models even reach 5521.27 ppm and 16437.20 ppm for ethyl acetate of 10,000 ppm; and the ARDs and MRDs given by the MVQR models even attain 62.64% and 809.78%, respectively, all for ethanol in the test set. Relatively speaking, PLSRs are between MVLRs and MVQRs in learning and predicted accuracies. Such results are quite identical with those demonstrated in Ref. [37].

The preset parameters of the 12 single-output perceptrons for quantification are $s = 5$, $\tau_{max} = 10{,}000$, $\eta = 0.025$, and $\varepsilon^* = 0.05$. Each

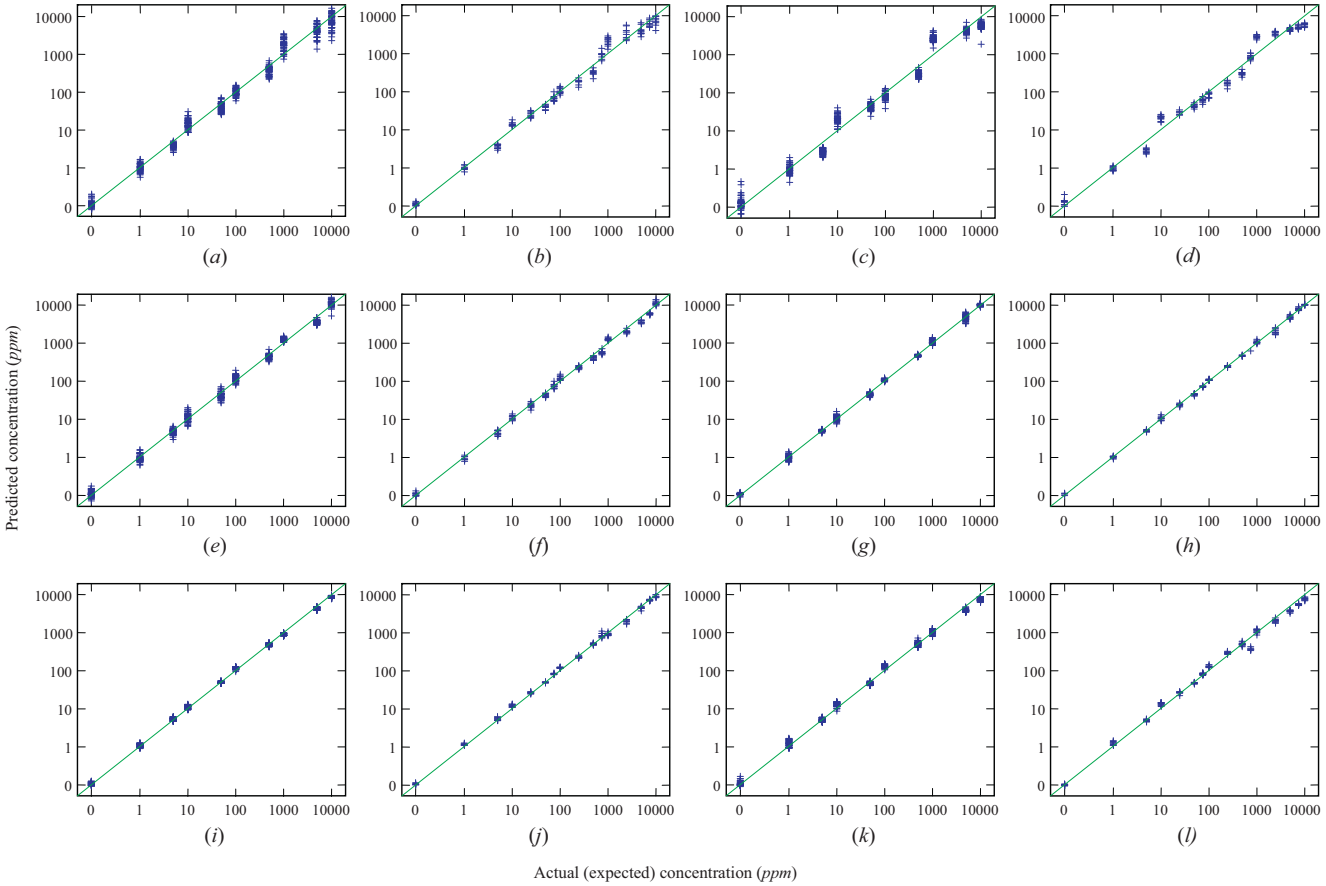


Fig. 5. Learning and predicted results of six regression models for quantifying ethyl caproate with 500 training and 160 test samples. (*a*), (*c*), (*e*), (*g*), (*i*) and (*k*) are the learning results of MVLR, PLSR, MVQR, MLP, SVM_G and SVM_P in order, and the rests are the predicted results of the corresponding models.

Please cite this article in press as: D. Gao, et al., Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose, Sens. Actuators B: Chem. (2011), doi:10.1016/j.snb.2011.11.003

Table 4
Comparison of time and structure complexities, and generalization performances of several hierarchical models.

Hierarchical model	1-*NN* + MVLR	LED + MVQR	LMD + MVLR	LMD + PLSR	LMD + MVQR	MLP + MLP	SVM-G + SVM-G	SVM-P + SVM-P
Learning time (s)	~+1.99	10.11 + 2.06	13.40 + 1.99	13.40 + 2.51	13.40 + 2.06	2232.27 + 213.33	1221.45 + 137.82	1385.71 + 159.66
Predicted time (s)	179.33 + 1.22	3.74 + 1.32	4.94 + 1.22	4.94 + 1.38	4.94 + 1.32	0.18 + 0.18	1.06 + 1.52	1.74 + 2.26
No. equivalent patterns	4250 + 12.75	85 + 24.75	722.50 + 12.75	722.50 + 12.75	722.50 + 24.75	222.75 + 68.25	4873.31 + 2856.82	5251.57 + 3941.33
Discrimination accuracy (%)[a]	66.26	65.29	67.10	67.10	67.10	95.81	80.45	72.52
Average relative deviation (%)[a]	30.09	62.64	30.09	53.59	62.64	9.5796	32.65	20.68

[a] Judged by the results for the test set.

perceptron iterates 10,000 epochs because ε^* is small. The max, min and average RMS errors of the 12 MLPs are 0.0061, 0.0307 and 0.0129 for the training subsets; and 0.0241, 0.0388 and 0.0304 for the test subsets. The ARDs are only 5.57% for the training set and 9.58% for the test set.

The 12 SVMs with Gaussian kernels for quantification are set with the parameters of $\gamma = 0.05$ and $C = 1000$, and the resulting ARDs of 17.63% are for the training and 32.65% for the test set. For the 12 SVMs with polynomial kernels, the preset parameters are $d = 3$, $C = 1000$, and the resulting ARDs are 23.22% for the training and 20.68% for the test set.

According to Table 3, based on learning speeds, the MVLR, PLSR and MVQR models are the fastest. However, judged by MSDs, MADs, ARDs and MRDs, the MLPs are the best. The SVMs do not perform as well as expected in generalization performance. An interesting phenomenon is that the 12 SVMs with Gaussian kernels perform better than those with polynomial kernels in fitting the training curves; however, the latter are a little better than the former in the aspect of average predicted accuracies.

Fig. 5 illustrates the learning and predicted results of the 6 quantification models for ethyl caproate. The results support the previous theoretical analysis and the experimental results in Table 3. Fig. 5(*l*) indicates that the SVM with polynomial kernels of $d = 3$ works quite well for predicting ethyl caproate except for the concentration point of 750 ppm which deviate its mean value a little too far.

It ought to be pointed out that the number of hidden nodes and the learning parameters for MLPs, and the widths of Gaussian kernels as well as the degrees of polynomial kernels for SVMs, are relatively optimal through repeated trials.

Table 4 compares the time complexities, structure complexities and generalization performances of several combinations of hierarchical model, including 1-*NN* classifiers followed by MVLRs, LED templates by MVQRs, LMD templates by MVLRs or PLSRs or MVQRs, MLPs by MLPs, SVMs by SVMs with Gaussian or polynomial kernels. According to the table, the combination of MSO perceptrons followed by MSO perceptrons has the longest time. However, in terms of time complexity for prediction, structure complexity and generalization performance, the hierarchical model based on MLPs is the best. The second best is the combination of SVMs.

Let us conduct a further discussion about the generalization performances of the hierarchical models mentioned above. Because the action of sigmoid activation functions is global, i.e., in the range $(-\infty, +\infty)$, and if the number of hidden nodes is sufficient and small enough, MLPs cannot only form arbitrarily complicated decision boundaries and approximate arbitrarily complicated nonlinear curves, but also avoid over-fitting [27]. Quite the reverse, Gaussian kernels are local and their widths are forced to be the same in advance for all support vectors. Furthermore, it is troublesome to pre-select suitable widths of Gaussian kernels [28], e.g., the width parameter γ chosen in the range $[2^{-10}, 2^{-9}, \ldots, 2^3, 2^4]$ in Ref. [40]. Though polynomial kernels are not certainly local, it is difficult to pre-select proper integral order of polynomials, and perhaps a non-integral order is better [28]. As a result, the generalization performances of SVMs are not so satisfactory.

5. Conclusions

This paper employs hierarchical models to implement the quantitative analysis of multiple kinds of odors with the divide-and-conquer strategy. We have developed and compared different hierarchical models such as 1-*NN* classifiers followed by MVLRs, LED templates by MVQRs, LMD templates by MVLRs or PLSRs or MVQRs, MLPs by MLPs, SVMs by SVMs.

Please cite this article in press as: D. Gao, et al., Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose, Sens. Actuators B: Chem. (2011), doi:10.1016/j.snb.2011.11.003

We measured 12 kinds of VOCs with the improved electronic nose and quantified their concentrations using the proposed hierarchical models. The experimental results for quantifying the VOCs support the following conclusions. (A) The hierarchical model composed of MSO perceptrons followed by MSO perceptrons with local decomposition, virtual balance and local generalization techniques has simple structure, fast prediction and good generalization. The more the sorts of odors and concentration points, the more superior the MLP-based hierarchical model. (B) The hierarchical models composed of SVMs are sub-optimal in comprehensive performance because of the difficulties to preselect proper widths of Gaussian kernels and appropriate degrees of polynomial kernels. (C) The hierarchical models consisting of 1-*NN* classifiers, distance templates, MVLRs, PLSRs and MVQRs perform poorly, because 1-*NN* classifiers and distance templates fail to identify new concentration points, and MVLRs, PLSRs and MVQRs are low in prediction accuracy for the large-range concentration changes.

Electronic noses will be inevitably applied to the quantitative analysis of many kinds of complex odors. The divide-and-conquer strategy and the corresponding hierarchical models will be good approaches. Future work will be devoted to these aspects.

Acknowledgments

This work is funded by the National Science Foundation of China (NSFC) under Grant Nos. 21176077 and 60675027, the High-Tech Development Program of China (863) under Grant No. 2006AA10Z315, and the Open Funding Project of the State Key Laboratory of Bioreactor Engineering.

References

[1] J.W. Gardner, P.N. Bartlett, Electronic Noses – Principles and Applications, Oxford University Press, 1999.
[2] T.C. Pearce, S.S. Schiffman, H.T. Nagle, J.W. Gardner (Eds.), Handbook of Machine Olfaction, Wiley-VCH Press, 2003.
[3] K. Persaud, G.H. Dodd, Analysis of discrimination mechanism of the mammalian olfactory system using a model nose, Nature 299 (1982) 352–355.
[4] E.Z. Panagou, N. Sahgal, N. Magan, G.J.E. Nychas, Table olives volatile fingerprints: potential of an electronic nose for quality discrimination, Sens. Actuators B: Chem. 134 (2008) 902–907.
[5] R. Stella, J.N. Barisci, G. Serra, G.G. Wallace, D.D. Rossi, Characterization of olive oil by an electronic nose based on conducting polymer sensors, Sens. Actuators B: Chem. 63 (2000) 1–9.
[6] H. Ulmer, J. Mitrovics, G. Noetzel, U. Weimar, W. Gopel, Odors and flavors identified with hybrid modular sensor systems, Sens. Actuators B: Chem. 43 (1997) 24–33.
[7] T. Nakamoto, A. Fukuda, T. Moriizumi, Perfume and flavor identification by odor-sensing system using quartz-resonator sensor array and neural-network pattern recognition, Sens. Actuators B: Chem. 10 (1993) 85–90.
[8] A. Eriksson, K.P. Waller, K.S. Sjaunja, J.E. Haugen, F. Lundby, O. Lind, Detection of mastitic milk using a gas-sensor array system (electronic nose), Int. Dairy J. 15 (2005) 1193–1201.
[9] B. Tudu, A. Jana, A. Metla, D. Ghosh, N. Bhattacharyya, R. Bandyopadhyay, Electronic nose for black tea quality evaluation by an incremental RBF network, Sens. Actuators B: Chem. 138 (2009) 90–95.
[10] Y. Huichun, W. Jun, X. Hong, L. Miao, Quality grade identification of green tea using the eigenvalues of PCA based on the E-nose signals, Sens. Actuators B: Chem. 140 (2009) 378–382.
[11] N.H. Beltran, M.A. Duarte-Mermoud, V.A. Soto Vicencio, S.A. Salah, M.A. Bustos, Chilean wine classification using volatile organic compounds data obtained with a fast GC analyzer, IEEE Trans. Instrum. Meas. 57 (11) (2008) 2421–2436.
[12] J.A.R. Sanchez, P. Chalier, D. Chevalier, M. Calderon-Santoyo, C. Ghommidh, Identification of different alcoholic beverages by electronic nose coupled to GC, Sens. Actuators B: Chem. 134 (2008) 43–48.
[13] D. Cozzolino, W. Cynkar, R. Dambergs, P. Smith, Two-dimensional correlation analysis of the effect of temperature on the fingerprint of wines analysed by mass spectrometry electronic nose, Sens. Actuators B: Chem. 145 (2010) 628–634.
[14] C.Y. Li, P. Heinemann, R. Sherry, Neural network and Bayesian network fusion models to fuse electronic nose and surface acoustic wave sensor data for apple defect detection, Sens. Actuators B: Chem. 125 (2007) 301–310.
[15] S. Panigrahi, S. Balasubramanian, H. Gu, C. Logue, M. Marchello, Neural-network-integrated electronic nose system for identification of spoiled beef, LWT – Food Sci. Technol. 39 (2006) 135–145.
[16] V.Y. Musatov, V.V. Sysoev, M. Sommer, I. Kiselev, Assessment of meat freshness with metal oxide sensor microarray electronic nose: a practical approach, Sens. Actuators B: Chem. 144 (2010) 99–103.
[17] N.E. Barbri, J. Mirhisse, R. Ionescu, N.E. Bari, X. Correig, B. Bouchikhi, E. Llobet, An electronic nose system based on a micro-machined gas sensor array to assess the freshness of sardines, Sens. Actuators B: Chem. 141 (2009) 538–543.
[18] A. Szcuzurek, P.M. Szecowka, B.W. Licznerski, Application of sensor array and neural networks for quantification of organic solvent vapors in air, Sens. Actuators B: Chem. 58 (1999) 427–432.
[19] A.C. Bastos, N. Magan, Potential of an electronic nose for the early detection and differentiation between Streptomyces in potable water, Sens. Actuators B: Chem. 116 (2006) 151–155.
[20] C. Laura, S. Selena, C. Paolo, D.R. Renato, I.G. Massimiliano, Electronic noses for the continuous monitoring of odors from a wastewater treatment plant at specific receptors: focus on training methods, Sens. Actuators B: Chem. 131 (2008) 53–62.
[21] A. D'Amico, C. Di Natale, R. Paolesse, A. Macagnano, E. Martinelli, G. Pennazza, M. Santonico, M. Bernabei, C. Roscioni, G. Galluccio, R. Bono, E. Finazzi Agrò and, S. Rullo, Olfactory systems for medical applications, Sens. Actuators B: Chem. 130 (2008) 458–465.
[22] A. Lamagna, S. Reich, D. Rodriguez, N.N. Scoccola, Performance of an e-nose in hops classification, Sens. Actuators B: Chem. 102 (2004) 278–283.
[23] T. Alizadeh, S. Zeynali, Electronic nose based on the polymer coated SAW sensors array for the warfare agent simulant classification, Sens. Actuators B: Chem. 129 (2008) 412–423.
[24] D. Matatagui, J. Martí, M.J. Fernández, J.L. Fontecha, J. Gutiérrez, I. Gràcia, C. Cané, M.C. Horrillo, Chemical warfare agents simulants detection with an optimized SAW sensor array, Sens. Actuators B: Chem. 154 (2011) 199–205.
[25] H. Knobloch, W. Schroedl, C. Turner, M. Chambers, P. Reinhold, Electronic nose responses and acute phase proteins correlate in blood using a bovine model of respiratory infection, Sens. Actuators B: Chem. 144 (2010) 81–87.
[26] H.C. Chang, L.B. Kish, M.D. King, C. Kwan, Fluctuation-enhanced sensing of bacterium odors, Sens. Actuators B: Chem. 142 (2009) 429–434.
[27] C.M. Bishop, Neural Networks for Pattern Recognition, Clarendon Press, Oxford, 1995.
[28] J.S. Taylor, N. Cristianini, Kernel Methods for Pattern Analysis, Cambridge University Press, London, 2004.
[29] C. Distante, N. Ancona, P. Siciliano, Support vector machine for olfactory singles recognition, Sens. Actuators B: Chem. 88 (2003) 30–39.
[30] J. Orts, E. Llobet, X. Vilanova, J. Brezmes, X. Correig, Selective methane detection under varying moisture conditions using static and dynamic sensor singles, Sens. Actuators B: Chem. 60 (1999) 106–117.
[31] E. Llobet, J. Brezmes, X. Vilanova, J.E. Suerias, X. Correig, Qualitative and quantitative analysis of volatile organic compounds using transient and steady-state responses of a thick-film tin oxide gas sensor array, Sens. Actuators B: Chem. 41 (1997) 13–21.
[32] G. Daqi, Y. Zeping, S. Jianli, Modular neural networks for estimating odor concentrations, in: 2008 IEEE World Cong. Comput. Intel. (WCCI'08), Hong Kong, June 1–6, 2008, pp. 3940–3947.
[33] T.G. Dietterich, P. Domingos, L. Getoor, Structured machine learning: the next 10 years, Mach. Learn. 73 (2008) 3–23.
[34] G. Horner, C. Hierold, Gas analysis by partial model building, Sens. Actuators B: Chem. 2 (1991) 173–184.
[35] G. Daqi, C. Wei, Simultaneous estimations of odor classes and concentrations using an electronic nose with function approximation model ensemble, Sens. Actuators B: Chem. 120 (2007) 584–594.
[36] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, John Wiley & Sons, Inc., New York, 2000.
[37] T. Sundic, S. Marco, A. Perera, A. Pardo, S. Hahn, N. Bârsan, U. Weimar, Fuzzy inference system for sensor array calibration: prediction of CO and CH_4 levels in variable humidity conditions, Chemometr. Intell. Lab. Syst. 64 (2) (2002) 103–122.
[38] G. Daqi, L. Chunxia, Y. Yunfan, Task decomposition and modular single-hidden-layer perceptron classifiers for multi-class learning problems, Pattern Recogn. 40 (2007) 2226–2236.
[39] P. Geladi, B. Kowalski, Partial least-squares regression: a tutorial, Anal. Chim. Acta 186 (1986), l-17.
[40] C.W. Hsu, C.J. Lin, A comparison of methods for multiclass support vector machines, IEEE Trans. Neural Network 13 (2002) 415–425 (Libsvm is available from: http://www.csie.ntu.edu.tw/~cjlin/libsvm/).

Biographies

Gao Daqi received his Ph.D degree from Zhejiang University, China, in 1996. Currently, he is a professor in East China University of Science and Technology. His research interests are electronic nose, pattern recognition and neural networks.

Liu Fangjun is currently a Ph.D student in East China University of Science and Technology. Her research interests include neural networks and machine olfactory.

Wang Ji is currently a Ph.D student in East China University of Science and Technology. His research interests include neural networks and pattern recognition.

Please cite this article in press as: D. Gao, et al., Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose, Sens. Actuators B: Chem. (2011), doi:10.1016/j.snb.2011.11.003