Analytica Chimica Acta xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Analytica Chimica Acta

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

A new kernel discriminant analysis framework for electronic nose recognition

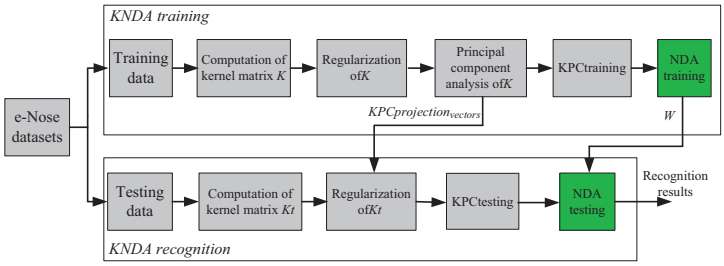
Lei Zhang [a,b,*], Feng-Chun Tian [a]

[a] College of Communication Engineering, Chongqing University, 174 ShaZheng street, ShaPingBa District, Chongqing 400044, China
[b] Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong

HIGHLIGHTS

- This paper proposes a new discriminant analysis framework for feature extraction and recognition.
- The principle of the proposed NDA is derived mathematically.
- The NDA framework is coupled with kernel PCA for classification.
- The proposed KNDA is compared with state of the art e-Nose recognition methods.
- The proposed KNDA shows the best performance in e-Nose experiments.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 25 November 2013
Accepted 23 January 2014
Available online xxx

Keywords:
Electronic nose
Discriminant analysis
Dimension reduction
Feature extraction
Multi-class recognition

ABSTRACT

Electronic nose (e-Nose) technology based on metal oxide semiconductor gas sensor array is widely studied for detection of gas components. This paper proposes a new discriminant analysis framework (NDA) for dimension reduction and e-Nose recognition. In a NDA, the between-class and the within-class Laplacian scatter matrix are designed from sample to sample, respectively, to characterize the between-class separability and the within-class compactness by seeking for discriminant matrix to simultaneously maximize the between-class Laplacian scatter and minimize the within-class Laplacian scatter. In terms of the linear separability in high dimensional kernel mapping space and the dimension reduction of principal component analysis (PCA), an effective kernel PCA plus NDA method (KNDA) is proposed for rapid detection of gas mixture components by an e-Nose. The NDA framework is derived in this paper as well as the specific implementations of the proposed KNDA method in training and recognition process. The KNDA is examined on the e-Nose datasets of six kinds of gas components, and compared with state of the art e-Nose classification methods. Experimental results demonstrate that the proposed KNDA method shows the best performance with average recognition rate and total recognition rate as 94.14% and 95.06% which leads to a promising feature extraction and multi-class recognition in e-Nose.

1. Introduction

Electronic nose (e-Nose) is an instrument comprised of a chemical sensor array with partial specificity and appropriate pattern recognition algorithm for detection of chemical analytes [1,2]. In recent years, a number of studies in e-Nose have been presented for classification applications in many fields, i.e. evaluation of tea and food quality [3–8], disease diagnosis [9–11] and environmental monitor [12–15], etc. In e-Nose, pattern recognition module with an appropriate sensor array plays an important role in usage, which decides the accuracy and robustness of e-Nose detection.

The classification methodologies have been widely studied in e-Nose applications. First, artificial neural networks (ANN) are widely

* Corresponding author at: Chongqing University, College of Computer Science, No. 174 Shazhengjie, ShaPingBa district, Chongqing, China. Tel.: +86 13629788369; fax: +86 23 65103544.
E-mail address: leizhang@cqu.edu.cn (L. Zhang).

0003-2670/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.aca.2014.01.049

used for qualitative and quantification analysis in the initial stage of e-Nose development, such as multilayer perceptron (MLP) neural network with backpropagation (BP) algorithm [4,8,15,16], RBF neural network [4,6,17], ARTMAP neural networks [18,19], etc. Then, decision tree method has also been proposed for classification [20]. With the proposal of support vector machine (SVM) with complete and theoretical proof, SVM has been widely studied in e-Nose [9,12,14,21,22] and also demonstrate that it is superior to ANN in accuracy and robustness, due to that SVM constructed as structural risk minimization is better than the only empirical risk minimization considered in ANN. The nonlinear mapping with different kernel functions (e.g. polynomial, Gaussian function, etc.) in SVM can make a linearly inseparable classification problem in original data space linearly separable in a high dimensional feature space, and realize the classification through a hyperplane. Besides, SVM is used to solve a convex quadratic programming problem and can promise global optimum which could not be achieved by ANN in a small sample of events. Although SVM seems to be the best selection in a binary classification, the algorithm complexity of SVM for a general multi-class problem may make an actual on-line application difficult in the development of our e-Nose.

Besides the classification methodologies, data preprocessing methods like feature extraction and dimension reduction methods [23–26] including principal component analysis (PCA), independent component analysis (ICA), kernel PCA (KPCA), linear discriminant analysis (LDA), singular value decomposition (SVD), etc. have also been combined with ANNs or SVMs for improving the prediction accuracy of e-Nose. Both feature extraction and dimension reduction aim to obtain useful features for classification. Dimension reduction can reduce the redundant information like denoising but may lose some useful information in original data. In addition, classification method has also ability to automatically depress the useless components in samples learning process.

The methodologies of PCA, KPCA, LDA, and the combination of KPCA and LDA have also wide application in many fields, such as time series forecasting, novelty detection, scene recognition, and face recognition as feature extraction and dimension reduction. Cao et al. employed a comparison of PCA, KPCA and ICA combined with SVM for time series forecasting, and find that KPCA has the best performance in feature extraction [27]. Xiao et al. proposed a L1 norm based KPCA algorithm for novelty detection and obtained satisfactory effect using simulation data set [28]. Hotta proposed a local feature acquisition method using KPCA for scene classification, and the performance is superior to conventional methods based on local correlation features [29]. Lu et al. [30] and Yang et al. [31] also proposed kernel direct discriminant analysis and KPCA plus LDA algorithms in face recognitions, and given a complete kernel fisher discriminant framework for feature extraction and recognition. Dixon et al. [32] presented a PLS-DA method used in gas chromatography mass spectrometry. A kernel PLS algorithm was also discussed in [33].

Inspired by these works, 2-norm between each two sample vectors x_i and x_j in between-class and within-class is considered with a similarity matrix calculated by the Gaussian function $\exp\left(-\left\|x_i - x_j\right\|^2/t\right)$. Through the construction of the between-class Laplacian scatter matrix and within-class Laplacian scatter matrix, the new discriminant analysis (NDA) framework is realized by solving an optimization problem which makes the samples between-class more separable and the samples within-class more compactable. Considering the characteristic of linearly separable in high dimensional kernel space, the Gaussian kernel function is introduced for mapping the original data space into a high dimensional space. To make the within-class Laplacian scatter matrix nonsingular in the calculation of eigenvalue problem in which a inverse operation of the within-class Laplacian scatter matrix is necessary, PCA is used to reduce the dimension of the kernel space. The contribution of this paper can be concluded as the proposed new discriminant analysis framework (KNDA) based on KPCA and its application in electronic nose for rapid detection of multiple kinds of pollutant gas components.

It is worthwhile to highlight several aspects of the proposed KNDA framework. First, in the NDA framework, each sample vector in between-class and within-class has been used for Laplacian scatter matrix, while in LDA only the centroid of between-class and within-class is used to calculate the scatter matrix in which each sample's information cannot be well represented. Second, a similarity matrix by a Gassian function is used to measure the importance of each two samples x_i and x_j with respect to their distance $\left\|x_i - x_j\right\|_2$ which is not considered in LDA. Third, the projection vector can be obtained by maximizing the between-class Laplacian scatter matrix and minimizing the within-class Laplacian scatter matrix. Fourth, the NDA is a supervised discriminant analysis framework and KNDA is the combined learning framework of unsupervised KPCA and supervised NDA for feature extraction and recognition. Fifth, the recognition in this paper is an intuitive Euclidean distance based method and promising the stability and reliability of the results.

The organization of this paper is as follows. In Section 2, we briefly review the related work such as PCA, KPCA and LDA. In Section 3, we describe the proposed approach including the proposed new discriminant analysis framework, the proposed kernel discriminant analysis learning framework and the multi-class recognition. In Section 4, electronic nose experiments and the experimental data are presented. Section 5 presents the experimental results and discussion. Conclusion is made in Section 6.

2. Related work

2.1. PCA

PCA [34] is an unsupervised method in dimension reduction by projecting correlated variables into another orthogonal feature space and thus a group of new variables with the largest variance (global variance maximization) were obtained. The PC coefficients can be obtained by calculating the eigenvectors of the covariance matrix of the original data set.

2.2. KPCA

KPCA does the PCA process in kernel space which introduces the advantage of high dimension mapping of original data using kernel trick on the basis of PCA, in which the original input vectors are mapped to a high dimensional feature space F. The mapping from the original data space to high dimensional feature space can be represented by calculating the symmetrical kernel matrix of input training pattern vectors using a Gaussian kernel function shown as

$$K(x, x_i) = \exp\left(\frac{-\left\|x - x_i\right\|^2}{\sigma^2}\right) \quad (1)$$

where x and x_i denotes the observation vectors, σ^2 denotes the width of Gaussian that is commonly called kernel parameter.

Please cite this article in press as: L. Zhang, F.-C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, Anal. Chim. Acta (2014), http://dx.doi.org/10.1016/j.aca.2014.01.049

In general, KPCA is to perform PCA algorithm in the high dimensional feature space K, and extract nonlinear feature. The size of dimension depends on the number of training vectors. The PCA process of kernel matrix K is to perform the following eigenvalue operation

$$K \times \alpha = \lambda \times \alpha \tag{2}$$

where α denotes the set of eigenvectors corresponding to d eigenvalues and λ denotes the diagonal matrix (λ_{ii} is the eigenvalue, $i = 1,\ldots,d$). The set of eigenvectors $\{\alpha|\alpha_i, i = 1,\ldots,r\}$ corresponding to the first r largest eigenvalues ordered in such a way $\lambda_1 > \lambda_2 > \cdots \lambda_r$ is the kernel principal component coefficients (projection vectors). Therefore, the kernel PC scores can be obtained by multiplying the kernel matrix K by the PC coefficients.

2.3. LDA

LDA aims to maximize the ratio of between-class variance to the within-class variance in any particular data set through a transformation vector w, and therefore promise the maximum separability. Finally, a linear decision boundary between two classes will be produced for classification.

To a binary classification (two classes), assume the dataset for the two classes to be X_1 and X_2, respectively. We write $X_1 = \{x_1^1, x_1^2, \ldots, x_1^{N_1}\}$, and $X_2 = \{x_2^1, x_2^2, \ldots, x_2^{N_2}\}$, N_1 and N_2 denote the numbers of column vectors for X_1 and X_2, $\{x_i^j, i = 1, 2; j = 1, \ldots, N_i\}$ denotes the column vector (observation sample). Then, we set the total dataset Z in R^d as $Z = \{X_1, X_2\}$.

Then, the within-class scatter matrix S_w and between-class scatter matrix S_b can be represented as

$$S_w = \sum_{i=1}^{2}\sum_{j=1}^{N_i}\left(x_i^j - \mu_i\right)\left(x_i^j - \mu_i\right)^T \tag{3}$$

$$S_b = \sum_{i=1}^{2} N_i \times \left(\mu_i - \bar{Z}\right)\left(\mu_i - \bar{Z}\right)^T \tag{4}$$

where μ_i denotes the centroid of the ith class, $\bar{Z}$ denotes the centroid of the total dataset Z, and symbol T denotes transpose.

If S_w is nonsingular, the optimal projection matrix w can be obtained by solving the following maximization problem

$$w = \text{argmax}\frac{w^T S_b w}{w^T S_w w} \tag{5}$$

The $\{w|w_i, i = 1,\ldots,m\}$ is the set of the eigenvectors of the $S_w^{-1}S_b$ corresponding to the m largest eigenvalues.

3. The proposed approach

3.1. The proposed NDA framework

Let $x_1 = [x_1^1, x_1^2, \ldots, x_1^{N_1}]$ be the training set of class 1 with N_1 samples, and $x_2 = [x_2^1, x_2^2, \ldots, x_2^{N_2}]$ be the training set of class 2 with N_2 samples. The proposed NDA framework aims to find the best projection basis W from the training sets between class 1 and class 2 to transform the training set into a low dimensional feature spaces. To our knowledge, the two classes would be more separable if the between-class scatter matrix is larger and the within-class scatter becomes more compact. The proposed NDA is a supervised dimension reduction algorithm from sample to sample, thus the similarity matrix A and B are introduced to construct the between-class Laplacian scatter and within-class Laplacian scatter. The similarity matrix A and B can be calculated as follows

$$A^{ij} = \exp\left(\frac{-\left\|x_1^i - x_2^j\right\|^2}{t}\right), \quad i = 1,\ldots,N_1; \quad j = 1,\ldots,N_2 \tag{6}$$

$$B_k^{ij} = \exp\left(\frac{-\left\|x_k^i - x_k^j\right\|^2}{t}\right), \quad i = 1,\ldots,N_k; \quad j = 1,\ldots,N_k; \quad k = 1,\ldots,c; \quad c = 2 \tag{7}$$

where t represents the width of Gaussian which is an empirical parameter. In this paper, $t = 100$.

Therefore, from the viewpoint of classification, we aim to maximize the ratio of the between-class Laplacian scatter matrix $J_1(W)$ and the within-class Laplacian scatter matrix $J_2(W)$. The specific algorithm derivation of the proposed NDA framework is shown as follows.

The between-class Laplacian scatter matrix can be represented as

$$\begin{aligned} J_1(W) &= \frac{1}{N_1 \times N_2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left\|W^T x_1^i - W^T x_2^j\right\|^2 A^{ij} \\ &= \frac{1}{N_1 \times N_2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} tr\left[\left(W^T x_1^i - W^T x_2^j\right)\left(W^T x_1^i - W^T x_2^j\right)^T\right] A^{ij} \\ &= \frac{1}{N_1 \times N_2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} tr\left[W^T\left(x_1^i - x_2^j\right)\left(x_1^i - x_2^j\right)^T W\right] A^{ij} \\ &= \frac{1}{N_1 \times N_2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} tr\left\{W^T\left[\left(x_1^i - x_2^j\right)\left(x_1^i - x_2^j\right)^T A^{ij}\right] W\right\} \\ &= tr\left\{W^T\left[\frac{1}{N_1 \times N_2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(x_1^i - x_2^j\right)\left(x_1^i - x_2^j\right)^T A^{ij}\right] W\right\} \end{aligned} \tag{8}$$

Now, we let

$$H_1 = \frac{1}{N_1 \times N_2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(x_1^i - x_2^j\right)\left(x_1^i - x_2^j\right)^T A^{ij} \tag{9}$$

Then, we get $J_1(W) = tr\left(W^T H_1 W\right)$.

Similarly, the within-class Laplacian scatter matrix can be represented as

$$\begin{aligned} J_2(W) &= \sum_{k=1}^{c}\frac{1}{N_k^2}\sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\left\|W^T x_k^i - W^T x_k^j\right\|^2 B_k^{ij} \\ &= \sum_{k=1}^{c}\frac{1}{N_k^2}\sum_{i=1}^{N_k}\sum_{j=1}^{N_k} tr\left[\left(W^T x_k^i - W^T x_k^j\right)\left(W^T x_k^i - W^T x_k^j\right)^T\right] B_k^{ij} \\ &= \sum_{k=1}^{c}\frac{1}{N_k^2}\sum_{i=1}^{N_k}\sum_{j=1}^{N_k} tr\left\{W^T\left[\left(x_k^i - x_k^j\right)\left(x_k^i - x_k^j\right)^T B_k^{ij}\right] W\right\} \\ &= tr\left\{W^T\left[\sum_{k=1}^{c}\frac{1}{N_k^2}\sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\left(x_k^i - x_k^j\right)\left(x_k^i - x_k^j\right)^T B_k^{ij}\right] W\right\} \end{aligned} \tag{10}$$

We let

$$H_2 = \sum_{k=1}^{c}\frac{1}{N_k^2}\sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\left(x_k^i - x_k^j\right)\left(x_k^i - x_k^j\right)^T B_k^{ij} \tag{11}$$

Then, we have $J_2(W) = tr\left(W^T H_2 W\right)$

Please cite this article in press as: L. Zhang, F.-C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, Anal. Chim. Acta (2014), http://dx.doi.org/10.1016/j.aca.2014.01.049

In this paper, the algorithm aims to solve a two-class problem, that is, $c = 2$. Therefore, we can rewrite the H_2 as

$$H_2 = \sum_{k=1}^{2} \frac{1}{N_k^2} \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} \left(x_k^i - x_k^j\right)\left(x_k^i - x_k^j\right)^T B_k^{ij} = \frac{1}{N_1^2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_1} \left(x_1^i - x_1^j\right)\left(x_1^i - x_1^j\right)^T B_1^{ij} + \frac{1}{N_2^2} \sum_{i=1}^{N_2} \sum_{j=1}^{N_2} \left(x_2^i - x_2^j\right)\left(x_2^i - x_2^j\right)^T B_1^{ij} \tag{12}$$

From the angle of classification, to make class 1 and class 2 more separable, we formulate the discriminative analysis model as the following optimization problem

$$\max J(W) = \max \frac{J_1(W)}{J_2(W)} = \max \frac{W^T H_1 W}{W^T H_2 W} \tag{13}$$

H_1 and H_2 have been derived in analysis, thus, we can find the projection basis W by solving the following eigenvalue problem

$$H_1 \varphi = \lambda H_2 \varphi \tag{14}$$

Then, the optimization problem (13) can be transformed into the following maximization problem

$$\max \frac{\varphi_i^T H_1 \varphi_i}{\varphi_i^T H_2 \varphi_i} \tag{15}$$

According to $H_1\varphi = \lambda H_2 \varphi$, we have

$$H_1\varphi_1 = \lambda_1 H_2 \varphi_1, \quad H_1\varphi_2 = \lambda_2 H_2 \varphi_2, \ldots, H_1\varphi_i = \lambda_i H_2 \varphi_i \tag{16}$$

Then, the maximization problem (15) can be solved as

$$\max \frac{\varphi_i^T H_1 \varphi_i}{\varphi_i^T H_2 \varphi_i} = \max \frac{\lambda_i \varphi_i^T H_2 \varphi_i}{\varphi_i^T H_2 \varphi_i} = \max \lambda_i \tag{17}$$

Let φ_1 be the eigenvector corresponding to the largest eigenvalue λ_1 ($\lambda_1 > \lambda_2 > \cdots > \lambda_d$), then the optimal projection basis W between class 1 and class 2 can be represented as $W = \varphi_1$.

3.2. The KPCA plus NDA algorithm (KNDA)

In this paper, the KPCA method is combined with the proposed NDA framework for feature extraction and recognition in e-Nose application. It is worthwhile to highlight the two reasons of KPCA in this work. First, the introduction of kernel function mapping is on the basis of the consideration that in a high dimensional kernel space, the patterns would become more separable linearly than the original data space. Second, the PCA is used for dimension reduction of the kernel data space and make the number of variables less than the number of training samples so that we can guarantee that the within-class Laplacian scatter matrix H_2 in Eq. (12) to be nonsingular in NDA framework.

The pseudocodes of the KNDA training algorithm have been described in Table 1.

The proposed NDA framework considers the two-class condition. To a multi-class (k classes, $k > 2$) problem, the NDA can also be useful by decomposing the multi-class problem into multiple two-class (binary) problems. Generally, "one-against-all (OAA)" and "one-against-one (OAO)" are often used in classification [35]. Study in [36] demonstrates that the OAO strategy would be a better choice in the case of $k \leq 10$, while this paper studies the discrimination of $k = 6$ kinds of pollutant gases. Therefore, $k \times (k-1)/2 = 15$ NDA models are designed in this paper.

3.3. Multi-class recognition

Given a test sample vector z, we first transform the sample vector z through the optimal projection vector W_i obtained using the proposed NDA method as follow

$$z' = W_i^T z \tag{18}$$

The recognition of z in a two-class (class 1 and class 2) problem can be employed in terms of the Euclidean distance 2-norm calculated as follows

$$\text{if} \|z' - \mu_i^1\|_2 > \|z' - \mu_i^2\|_2, \quad z \in \{\text{class 1}\}; \quad \text{else}, \quad z \in \{\text{class 2}\} \tag{19}$$

where symbol $\|\cdot\|_2$ denotes 2-norm, μ_i^1 and μ_i^2 denote the centroid of class 1 and class 2 in the i-model, respectively.

For a multi-class recognition, a majority voting mechanism in decision level is used based on the OAO strategy. The statics of vote number V_j for class j can be shown by

$$V_j = \sum_{i=1}^{k\times(k-1)/2} I\left(c_i = T_j\right), \quad j = 1, \ldots, k \tag{20}$$

where $I(\cdot)$ denotes the binary indicator function, c_i denotes the predicted label of the ith sub-classifier, and T_j denotes the true label of class j. The class label of class j with the largest vote number $\max V_j$ is the discriminated class of the test sample vector. The pseudocodes of the proposed KNDA recognition (testing) process have been described in Table 2.

The diagram of the proposed KNDA electronic nose recognition method has been illustrated in Fig. 1, wherein two parts are included: KNDA training and KNDA recognition. The specific implementations of KNDA training and KNDA recognition have been illustrated in Tables 1 and 2, respectively. All the algorithms in this paper are implemented in the platform of Matlab software (version 7.8) on a laptop with Intel Core i5 CPU and 2GB RAM.

4. Experiments

4.1. Sensor array based e-Nose

The sensor array based e-Nose system with Field Programmable Gate Array (FPGA) processor has been introduced in [14]. Consider the selectivity, stability, reproducibility, sensitivity and low-cost of metal oxide semiconductor (MOS) gas sensors, our sensor array in e-Nose system consists of four metal oxide semiconductor gas sensors including TGS2602, TGS2620, TGS2201A and TGS2201B. Moreover, a module with two auxiliary sensors for the temperature and humidity measurement is also used with consideration that MOS gas sensors are sensitivity to environmental temperature and humidity. Therefore, 6 variables are contained in each observation. A 12-bit analog-digital converter is used as the interface between FPGA processor and sensor array for convenient digital signal processing. The e-Nose system can be connected to a PC via a Joint Test Action Group (JTAG) port for data storage and debugging programs. The e-Nose system and the experimental platform are illustrated in Fig. 2 in which the typical response of gas sensors in the sampling process ((1) baseline, (2) transient response, (3) steady state response, (4) recover process) are also presented.

4.2. Experimental data

In this paper, six familiar chemical contaminants indoor including formaldehyde, benzene, toluene, carbon monoxide, ammonia and nitrogen dioxide are studied with an e-Nose. The experiments

Please cite this article in press as: L. Zhang, F.-C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, Anal. Chim. Acta (2014), http://dx.doi.org/10.1016/j.aca.2014.01.049

Table 1
The pseudocodes of the KNDA algorithm in the training stage.

Input: The training sets $x = [x_1, x_2, \ldots, x_m]$ of k classes, the parameter t, the kernel parameter σ of kernel space mapping, and the threshold of accumulated contribution rate of the kernel principal components
Output: Projection basis W_i, $i = 1,\ldots,k(k-1)/2$ and the centroid set μ of $k \times (k-1)/2$ models
Step 1 (Kernel function mapping of the training sets)
1.1. Computation of the symmetrical kernel matrix $K_{m \times m}$ of the training sets x using Eq. (1)
1.2. Regularization of the kernel matrix K
1.1.1. Centralization of K by using $K = K - 1/mI \times K - 1/mK \times I + 1/mI \times K \times I$
1.1.2. Normalization by using $K = K/m$
Step 2 (Kernel principal components analysis)
2.1. Eigenvalue decomposition of K by equation $K \times V = \lambda \times V$, where λ and V denote eigenvalues and eigenvectors, respectively
2.2. Sort the eigenvalues in descending order $\lambda_1 > \lambda_1 > \cdots > \lambda_m$, and the sorted new_eigenvectors
2.3. Calculate the accumulated contribution rate (ACR) which is shown by $\mathrm{ACR}_j = \sum_j^{k=1} \lambda_k / \sum_m^{i=1} \lambda \times 100, \quad j = 1, \ldots, m$
2.3. Determine the number j of kernel principal components by using $j = \arg\min \left\{ \mathrm{ACR}_j \geq \mathrm{threshold} \right\}$
2.4. The $\mathrm{KPCprojection}_{\mathrm{vectors}}$ for kernel principal components projection is determined as $\mathrm{KPCprojection}_{\mathrm{vectors}} = \left\{ \mathrm{new\_eigenvectors}_i, \quad i = 1, \ldots, j \right\}$
2.5. Calculate the kernel principal components $\mathrm{KernelPC}_{\mathrm{training}}$ of the training sets by using $\mathrm{KernelPC}_{\mathrm{training}} = K \times \mathrm{KPCprojection}_{\mathrm{vectors}}$
Step 3 (NDA framework)
For $i = 1,2,\ldots,k(k-1)/2$, repeat
3.1. Take the training vectors of the ith pair of classes from $\mathrm{KernelPC}_{\mathrm{training}}$, and calculate the between-class similarity matrix A and within-class similarity matrix B according to Eqs. (6) and (7), respectively
3.2. Calculate the between-class Laplacian scatter matrix H_1 and within-class Laplacian scatter matrix H_2 as shown in Eqs. (9) and (11), respectively
3.3. Solve the eigenvalue problem (14) and get the eigenvector φ_1 corresponding to the largest eigenvalue
3.4. Obtain the projection basis $W_i = \varphi_1$
3.5. Calculate the ith centroid pair $\mu_i = \left[\mu_i^1, \mu_i^2 \right]$ of class 1 and class 2 in the ith model
End for
Step 4 (Output the low dimensional projection basis matrix W)
Output the projection basis matrix $W = \left\{ W_i, i = 1, \ldots, k(k-1)/2 \right\}$ and the centroid set $\mu = \left\{ \mu_i, i = 1, \ldots, k(k-1)/2 \right\}$

Table 2
The pseudocodes of the KNDA recognition.

Input: The testing sets $z = [z_1, z_2, \ldots, z_n]$ of k classes, the kernel parameter σ of kernel space mapping, the $\mathrm{KPCprojection}_{\mathrm{vectors}}$, the projection basis matrix W and the centroid $\mu = \left\{ \mu_i^1, \mu_i^2 \right\}, \quad i = 1, \ldots, k(k-1)/2$, obtained in the KNDA training process
Output: The predicted labels of testing samples
Step 1 (Kernel function mapping of the testing sets)
1.1. Computation of the symmetrical kernel matrix $Kt_{n \times n}$ of the testing sets z using Eq. (1)
1.2. Regularization of the kernel matrix Kt
1.1.1. Centralization of Kt by using $Kt = Kt - 1/nI \times Kt - 1/nKt \times I1/nI \times Kt \times I$
1.1.2. Normalization by using $Kt = Kt/n$
Step 2 (Kernel principal components projection of the testing sets)
Calculate the kernel principal components $\mathrm{KernelPC}_{\mathrm{testing}}$ of testing vectors $\mathrm{KernelPC}_{\mathrm{testing}} = Kt \times \mathrm{KPCprojection}_{\mathrm{vectors}}$
Step 3 (Multi-class recognition)
For $p = 1,2,\ldots,n$, repeat
For $i = 1, \ldots, k(k-1)/2$, repeat
3.1. Low dimensional projection of NDA
$\mathrm{LowDim}_{\mathrm{projection}}^{i,p} = W_i^{\mathrm{T}} \times \mathrm{KernelPC}_{\mathrm{testing}}^{p}$
3.2. Calculate the Euclidean distance between $\mathrm{LowDim}_{\mathrm{projection}}^{i,p}$ and the centroid μ_i, and discriminate the label c_i of the pth sample in the ith classifier according to Eq. (19)
End for
3.3. Compute the vote number V_j for class j of the pth sample according to Eq. (20)
3.4. Predict the label of the pth sample as $j = \arg\max V_j$
End for
Step 4 (Output the predicted labels of testing samples)
Output the predicted labels of the n testing samples

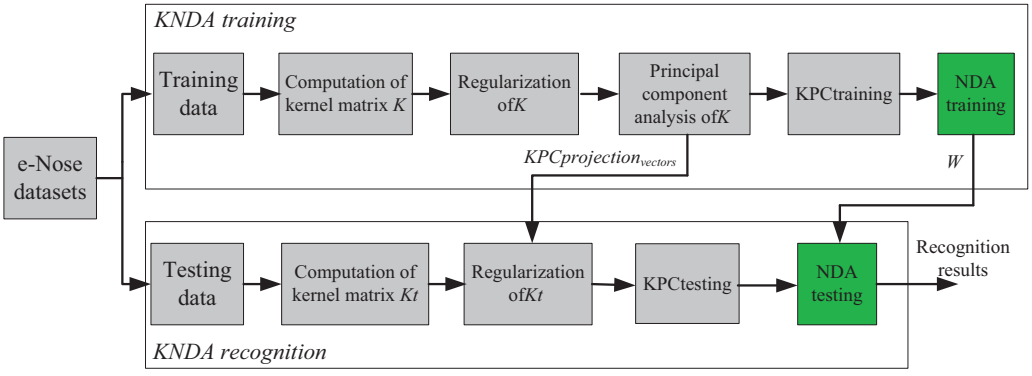


Fig. 1. Diagram of the proposed classification method.

Please cite this article in press as: L. Zhang, F.-C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, Anal. Chim. Acta (2014), http://dx.doi.org/10.1016/j.aca.2014.01.049

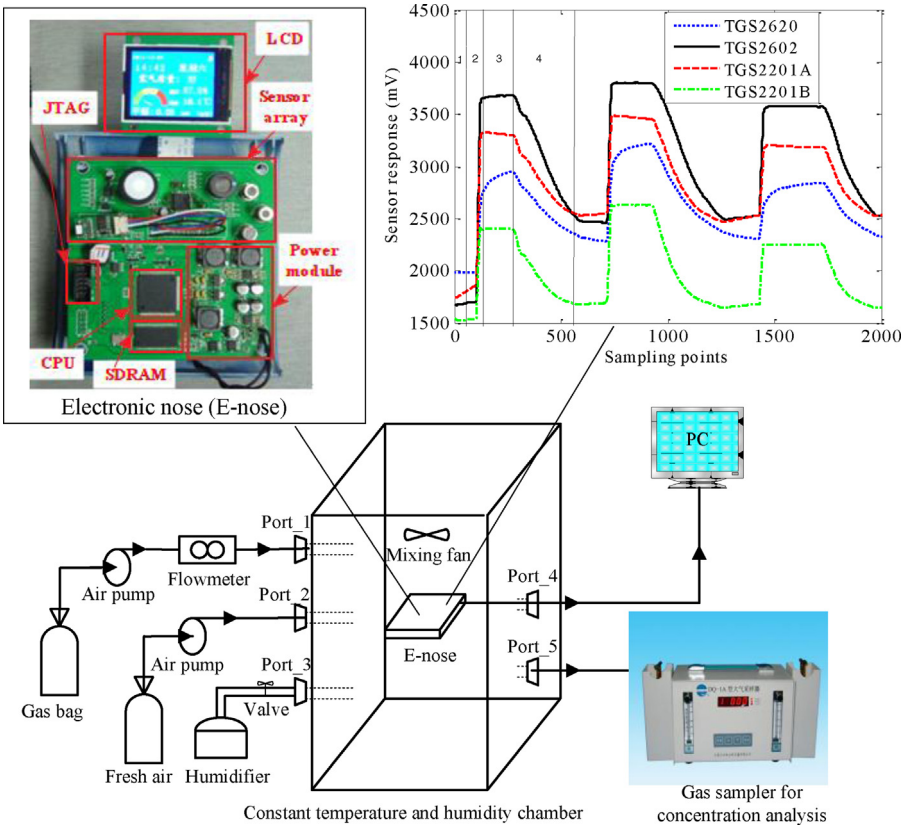


Fig. 2. Portable e-Nose, the experimental platform and the typical sensor response in this paper.

were employed in a constant temperature and humidity chamber in a condition of room temperature (15–35 °C). The experimental process for each gas is similar in which three main steps are included. First, set the target temperature and humidity and collect the sensor baseline for 2 min. Second, inject the target gas by using a flowmeter with time controlled, and collect the steady state response of sensors for 8 min. Third, clean the chamber by air exhaust for 10 min and read the data for the sample by a laptop connected with the electronic nose through a JTAG.

For more information about all the samples, we have described the experimental temperature, relative humidity, and concentration for each sample of each gas in supplementary data. The number of formaldehyde, benzene, toluene, carbon monoxide, ammonia and nitrogen dioxide samples are 188, 72, 66, 58, 60, and 38, respectively. In each sample, 6 variables (with 6 sensing units) are contained. All the experimental samples were obtained within two months by employing the e-Nose experiments continuously.

To determine the training sample index, we introduce the Kennard–Stone sequential (KSS) algorithm [37] based on Euclidean distance to select the most representative samples in the whole sample space for each gas. The selection starts by taking the pair of sample vectors (p_1, p_2) with the largest distance $d(p_1, p_2)$ among the samples for each gas. KSS follows a stepwise procedure that new selections are taken which would be the farthest from the samples already selected, until the number of training samples for each gas reaches. In this way, the most representative samples for each gas can be selected as training samples and guarantee the reliability of the learned model. The merit of KSS for training samples selection is to reduce the complexity of cross validation in performance evaluation. The remaining samples without being selected would be used for model testing. The specific number of training and testing samples after KSS selection for each gas are illustrated in Table 3.

5. Results and discussion

5.1. Contribution rate analysis

In the classification model, the kernel parameters σ^2 and the accumulated contribution rate (ACR) in KPCA are related with the actual classification performance. In experiments, six values {5, 6, 7, 8, 9, 10} of σ^2 and five values {95%, 96%, 97%, 98%, 99%} of the CR (the threshold of ACR) are selected for study and comparison, because these values have more positive effects in classifications than other values. Therefore, we do not use special optimizations to search the best parameters of KPCA. Totally, 30 kinds of combinations of (σ^2, CR) are studied in classification.

For KPCA analysis, we perform the KPCA algorithm on the total training samples (320 samples) of all gases. The size of the kernel matrix K should be 320 multiply 320. Table 4 presents the contribution rate analysis of KPCA including the number of kernel principal components (KPCs) with their ACR lower than the threshold CR. We can see that the number of KPCs (dimension) with ACR < 95%,

Table 3
Statistic of the experimental data in this paper.

Data	Formaldehyde	Benzene	Toluene	Carbon monoxide	Ammonia	Nitrogen dioxide
Training	125	48	44	38	40	25
Testing	63	24	22	20	20	13
Total	188	72	66	58	60	38

Please cite this article in press as: L. Zhang, F.-C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, Anal. Chim. Acta (2014), http://dx.doi.org/10.1016/j.aca.2014.01.049

Table 4
Contribution rate analysis of KPCA.

Threshold CR	<95%	<96%	<97%	<98%	<99%
Number of KPCs (dimension)	47	53	61	73	95
ACR	94.85%	95.91%	96.96%	97.96%	98.99%

ACR < 96%, ACR < 97%, ACR < 98%, and ACR < 99% is 47, 53, 61, 73 and 95, respectively. From the contribution rate analysis, we can find that about 99% information can be obtained by the first 95 principal components which is much lower than 320, and about 95% information is obtained by only the first 47 principal components.

5.2. Comparisons with state of the art e-Nose classification methods

The classification performance of proposed KNDA method can be shown by the average recognition rates and the total recognition rates of six kinds of chemicals. The average recognition rate denotes the average value of six recognition rates for six chemicals and it can validate the balance of multi-class recognition. The low average recognition rate demonstrates that at least one class failed in recognition. The total recognition rate represents the ratio between the number of correctly recognized samples for six chemicals and the number of total testing samples of six chemicals.

To study the kernel parameter σ^2 and the CR mentioned in KPCA, an empirical way such that six values {5, 6, 7, 8, 9, 10} of σ^2 and five values {95%, 96%, 97%, 98%, 99%} of the CR (the threshold of ACR) are selected for study and comparison of KNDA, KLDA, KPLS-DA, and KSVM. In experiments, we find that the best classification performance of KNDA is in the case of $\sigma^2 = 7$ and CR = 97% with the average recognition rate as 94.14% and the total recognition rate as 95.06%. In KPCA plus LDA (KLDA), we can find that the best performance with average recognition rate as 92.94% and total recognition rate as 94.44% is obtained in the case of $\sigma^2 = 6$ and CR = 95%. In KPLS-DA, the best performance of the average recognition rate and total recognition rate are 89.96% and 93.21% in the case of $\sigma^2 = 6$ and CR = 99%. Through the comparison of the best classification performance, we can see that the proposed KNDA has a better performance than KLDA in feature extraction for multi-class recognition. Moreover, we have implemented the KPCA plus SVM (KSVM) method for multi-class classification. In experiments, we can find that the only one case with $\sigma^2 = 10$ and CR = 98% by using KSVM has an average recognition rate 92.97% that is higher than 90%, and the corresponding total recognition rate is 95.06%. Seen from the results, SVM has an equal performance with the proposed KNDA in total recognition rate, while the proposed KNDA has a better performance from the average recognition rate. The average recognition rate demonstrates that the proposed KNDA has a better balance of recognition than SVM.

For details of the results, we present variation curves of the total recognition rate and the average recognition rate of four kernel methods (KSVM, KLDA, KPLS-DA, and KNDA) in Figs. 3 and 4, respectively, with $\sigma^2 = 5, 6, \ldots, 10$ and the ACR changes from 50% to 99% for each σ^2. The relation between the kernel parameters σ^2 and ACR of KPCA and the classification performance can be shown. From Figs. 3 and 4, we can find that the KSVM performs better in low ACR, and the KLDA, KPLS-DA and KNDA performs as good as KSVM with the increasing of ACR. When the ACR reaches 97%, the KNDA performs the best among KSVM, KLDA and KPLS-DA. In contrast, KPLS-DA shows the worst performance among the four kernel based methods. It is worth noting that the 97% is an inflection point and there is an obvious reduction of recognition rate when ACR is 98% and 99% for KLDA and KNDA. This may be explained that the most useful information of the original data is the 97% principal components, and the remaining 3% is the redundant information which is not useful for recognition. We can see that the proposed

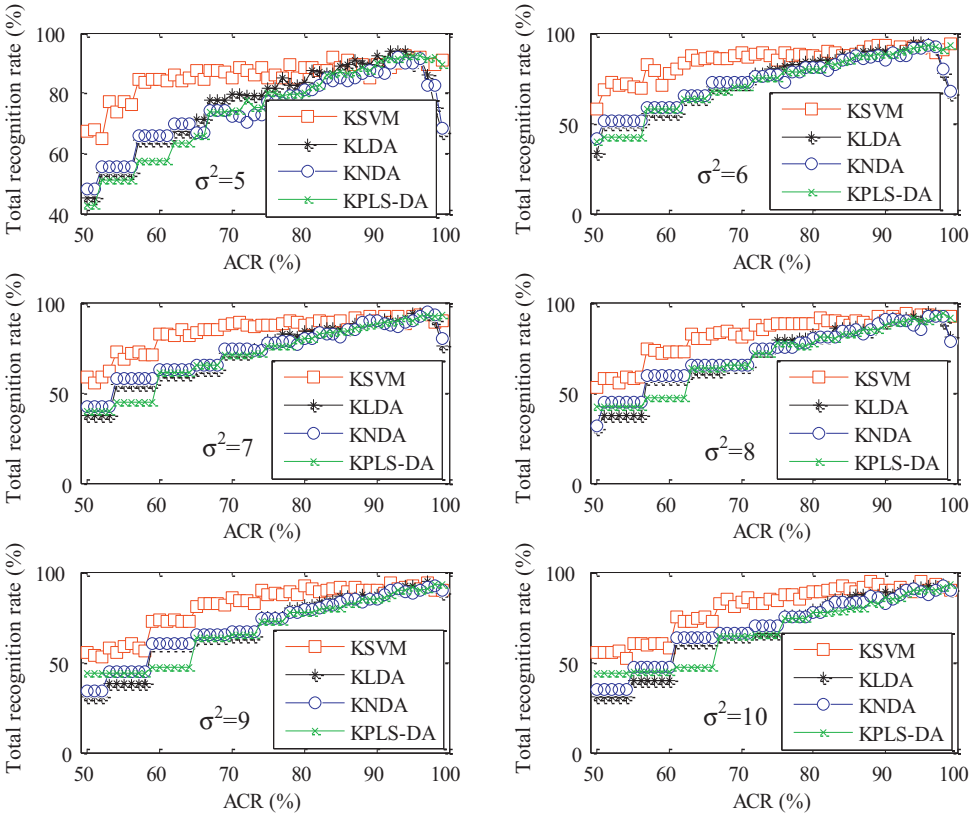


Fig. 3. Total recognition rates of all gases for four kernel based methods.

Please cite this article in press as: L. Zhang, F.-C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, Anal. Chim. Acta (2014), http://dx.doi.org/10.1016/j.aca.2014.01.049

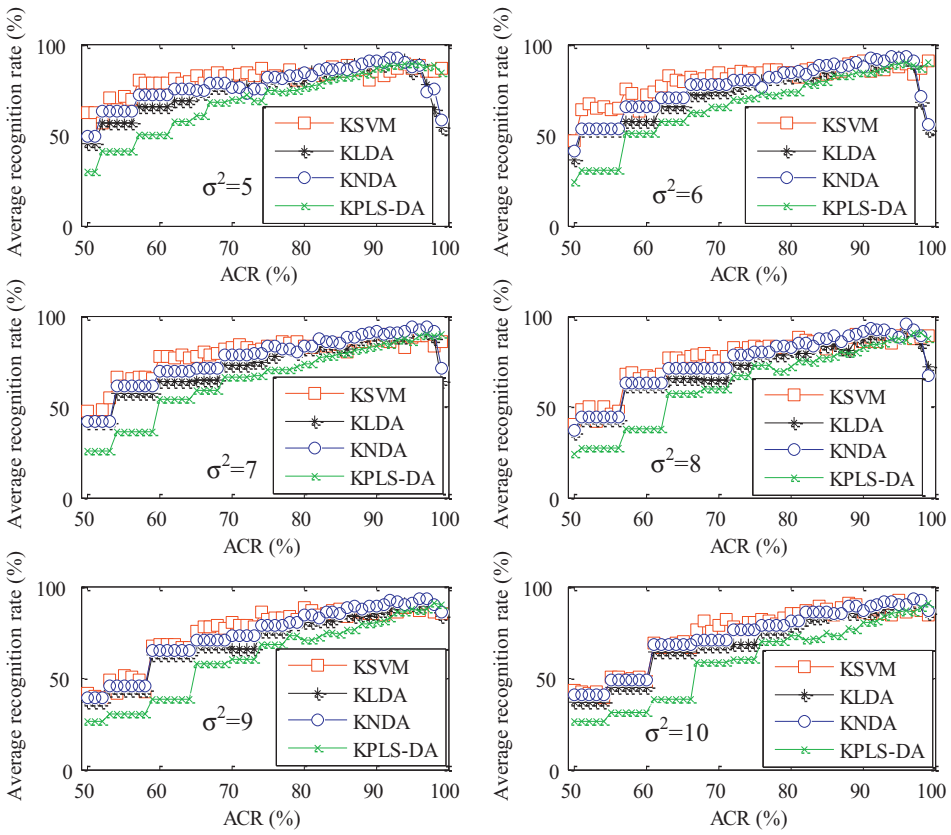


Fig. 4. Average recognition rates of all gases for four kernel based methods.

KNDA is sensitive to the noise which is similar to KLDA, while SVM is a nonlinear classification model and noise insensitive. However, considering that smaller number of principal components is also desirable on the basis of the high recognition rate, the proposed KNDA framework is superior to SVM based methods.

For completeness of comparisons, we have also studied the original SVM (SVM), PCA based SVM (PCA-SVM), original LDA (LDA), PCA based LDA (PCA-LDA), PLS-DA, kernel PLS-DA (KPLS-DA), and the proposed NDA and KNDA methods. The best classification performance of each method has been presented in Table 5 which shows the recognition rate for each gas, the average recognition rate and the total recognition rate. It is worthwhile noting that several facets should be highlighted in Table 5. First, from the comparison of LDA and the proposed NDA framework, NDA has higher recognition rates than LDA. Second, seen from the results of KSVM, the recognition rates of benzene and nitrogen dioxide are 87.50% and 76.92%, respectively. While the recognition rates of benzene and nitrogen dioxide are 100% and 84.62% which have been much improved by using KNDA. Thus, the average recognition rate of KNDA is the higher than KSVM that also demonstrates that the proposed KNDA can effectively improve the imbalance of E-nose data in SVM classification. Note that no specific method is used to improve the imbalance of experimental samples for each class in this work. Third, it is worth noting that both PLS-DA and KPLS-DA recognition methods have the same flaw of over-training as PLS, due to the unpredictable number of components in regression. However, to present the best results of each method for fair comparison, the results of PLS-DA and KPLS-DA in this paper may be over-learned.

In measurement of an e-Nose, the generalization capability is a very important factor in a system. The KNDA is actually a feature extraction method, and the recognition is based on an intuitive Euclidean distance method. Instead, SVM recognition is to solve an optimization problem. Therefore, the generalization capability of SVM in recognition depends on the optimization effect including the parameter selection in training. That is, overfitting or different SVM parameter selection, the recognition results would also be different. In general, the recognition results of KNDA should be more stable and reliable.

5.3. Computational efficiency

From the theories of these methods, the proposed NDA framework belongs to a linear discrimination, while KNDA introduces the kernel PCA in the proposed NDA framework. SVM is a nonlinear classification, and SVM aims to solve a convex quadratic programming problem. Though SVM has been widely studied for its complete theory in mathematics, SVM based classifiers have also a large computational burden which is related with the number of support vectors. For analysis of the computational efficiency of each method, the average running time including the training time and recognition time of each method for 10 times has been presented in Table 6. The consumed time in multi-class classification using SVM based classifiers (SVM, PCA-SVM, and KSVM) are generally more than 30 s, the LDA and PLS based methods (LDA, PCA-LDA, KLDA, PLS-DA, and KPLS-DA) takes less than 1 s, and the proposed NDA and KNDA methods take 1.811 and 2.221 s, respectively. Though the proposed NDA framework has a little higher computational efficiency than LDA, from the angle of synthesized consideration of the recognition accuracy and computational efficiency, the proposed KNDA is more acceptable in real application in terms of its best performance among state of the art methods.

Please cite this article in press as: L. Zhang, F.-C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, Anal. Chim. Acta (2014), http://dx.doi.org/10.1016/j.aca.2014.01.049

Table 5
Comparisons of classification accuracies with State-of-The-Art methods.

Methods	Recognition accuracy (%)							
	HCHO	C_6H_6	C_7H_8	CO	NH_3	NO_2	Average	Total
SVM	98.41	79.17	100.0	100.0	90.00	69.23	89.47	92.59
PCA-SVM	98.41	91.67	100.0	65.00	100.0	30.77	80.97	88.27
KSVM	98.41	87.50	100.0	100.0	95.00	76.92	92.97	95.06
LDA	88.89	66.67	90.91	100.0	90.00	30.77	77.87	82.72
PCA-LDA	82.54	58.33	86.36	90.00	90.00	30.77	73.00	77.16
PLS-DA	93.65	45.83	68.18	75.00	70.00	23.08	62.62	72.22
NDA	87.30	66.67	100.0	100.0	95.00	30.77	79.96	83.95
KLDA	95.24	100.0	95.45	95.00	95.00	76.92	92.94	94.44
KPLS-DA	98.41	91.67	95.45	95.00	90.00	69.23	89.96	93.21
KNDA	95.24	100.0	100.0	95.00	90.00	84.62	94.14	95.06

Table 6
Comparison of algorithms' running time (in s).

Methods	Training	Recognition	Total time
SVM	33.0	0.310	33.31
PCA-SVM	35.0	0.352	35.35
KSVM	40.0	0.620	40.62
LDA	0.144	0.011	0.155
PCA-LDA	0.194	0.011	0.205
PLS-DA	0.037	0.026	0.063
NDA	1.801	0.010	1.811
KLDA	0.425	0.092	0.517
KPLS-DA	0.280	0.091	0.371
KNDA	2.119	0.102	2.221

6. Conclusions

This paper present a rapid detection of six kinds of indoor air contaminants by a metal oxide semiconductor gas sensor array based e-Nose coupled with a KNDA method. The mathematical derivation of the proposed NDA framework has been shown. The between-class and within-class Laplacian scatter matrix from sample to sample are used in NDA. In KNDA, the KPCA contains high dimensional kernel space mapping and principal component analysis, which has two merits: first, the samples between classes become linearly separable in the high dimensional kernel space; second, the PCA is used to extract the most important information and reduce the dimension of the kernel space, and guarantee the within-class Laplacian scatter matrix nonsingular in NDA training. The specific implementations of KNDA training and recognition have been presented for readers' who are interested in the proposed classifier. Through the comparisons with LDA and PLS based methods, we find that NDA is more effective than LDA and PLS-DA. Besides, the proposed KNDA is also better than PCA-LDA, KLDA and KPLS-DA in classification. Through comparisons with state of the art SVM methods, the results demonstrate the proposed KNDA method in this paper is superior to all of them in the classification performance although KNDA has the same total recognition rate as KSVM. Because KNDA can effectively improve the imbalance of E-nose data in KSVM recognition, the average recognition rate of KNDA is higher than KSVM. From the algorithms' computational burden and time complexity, the proposed KNDA is about 20 times lower than SVM methods. Though it is a little higher than LDA methods, the running time (2 s) of KNDA is still acceptable in rapid detection with an e-Nose by virtue of its higher recognition accuracy.

Acknowledgements

We would like to express our sincere appreciation to the anonymous reviewers for their insightful comments, which have greatly improved the quality of the paper. This work was supported by Hong Kong Scholar Program (No. XJ2013044) and New Academic Researcher Award for Doctoral Candidates granted by Ministry of Education in China.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.aca.2014.01.049.

References

[1] J.W. Gardner, P.N. Bartlett, Electronic Noses: Principles and Applications, Oxford University Press, Oxford, UK, 1999.
[2] S.M. Scott, D. James, Z. Ali, Data analysis for electronic nose systems, Microchimica Acta 156 (2006) 183–207.
[3] K. Brudzewski, S. Osowski, A. Dwulit, Recognition of coffee using differential electronic nose, IEEE Transactions on Instrumentation and Measurement 61 (2012) 1803–1810.
[4] R. Dutta, E.L. Hines, J.W. Gardner, K.R. Kashwan, A. Bhuyan, Tea quality prediction using a tin oxide-based electronic nose: an artificial intelligence approach, Sensors and Actuators B: Chemical 94 (2003) 228–237.
[5] N. Bhattacharya, B. Tudu, A. Jana, D. Ghosh, R. Bandhopadhyaya, M. Bhuyan, Preemptive identification of optimum fermentation time for black tea using electronic nose, Sensors and Actuators B: Chemical 131 (2008) 110–116.
[6] B. Tudu, A. Jana, A. Metla, D. Ghosh, N. Bhattacharyya, R. Bandyopadhyay, Electronic nose for black tea quality evaluation by an incremental RBF network, Sensors and Actuators B: Chemical 138 (2009) 90–95.
[7] M. Peris, L. Escuder-Gilabert, A 21st century technique for food control: electronic noses, Analytica Chimica Acta 638 (2009) 1–15.
[8] B. Dębska, B. Guzowska-Świder, Application of artificial neural network in food classification, Analytica Chimica Acta 705 (2011) 283–291.
[9] H.L. Chen, D.Y. Liu, B. Yang, J. Liu, G. Wang, A new hybrid method based on local fisher discriminant analysis and support vector machines for hepatitis disease diagnosis, Expert Systems with Applications 38 (2011) 11796–11803.
[10] M. Bernabei, G. Pennazza, M. Santortico, C. Corsi, C. Roscioni, R. Paolesse, C. Di Natale, A. D'Amico, A preliminary study on the possibility to diagnose urinary tract cancers by an electronic nose, Sensors and Actuators B: Chemical 131 (2008) 1–4.
[11] J.W. Gardner, H.W. Shin, E.L. Hines, An electronic nose system to diagnose illness, Sensors and Actuators B: Chemical 70 (2000) 19–24.
[12] S. Güeny, A. Atasoy, Multiclass classification of *n*-butanol concentrations with *k*-nearest neighbor algorithm and support vector machine in an electronic nose, Sensors and Actuators B: Chemical 166 (167) (2012) 721–725.
[13] Q. Ameer, S.B. Adeloju, Polypyrrole-based electronic noses for environmental and industrial analysis, Sensors and Actuators B: Chemical 106 (2005) 541–552.
[14] L. Zhang, F. Tian, H. Nie, L. Dang, G. Li, Q. Ye, C. Kadri, 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.
[15] J. Getino, M.C. Horrillo, J. Gutiérrez, L. Arés, J.I. Robla, C. García, I. Sayago, Analysis of VOCs with a tin oxide sensor array, Sensors and Actuators B: Chemical 43 (1997) 200–205.
[16] L. Zhang, F. Tian, C. Kadri, G. Pei, H. Li, L. Pan, Gases concentration estimation using heuristics and bio-inspired optimization models for experimental chemical electronic nose, Sensors and Actuators B: Chemical 160 (2011) 760–770.
[17] Z. Ali, D. James, W.T. O'Hare, F.J. Rowell, S.M. Scott, Radial basis neural network for the classification of fresh edible oils using an electronic nose, Journal of Thermal Analysis and Calorimetry 71 (2003) 147–154.
[18] E. Llobet, E.L. Hines, J.W. Gardner, P.N. Bartlett, T.T. Mottram, Fuzzy ARTMAP based electronic nose data analysis, Sensors and Actuators B: Chemical 61 (1999) 183–190.

Please cite this article in press as: L. Zhang, F.-C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, Anal. Chim. Acta (2014), http://dx.doi.org/10.1016/j.aca.2014.01.049

[19] Z. Xu, X. Shi, L. Wang, J. Luo, C.J. Zhong, S. Lu, Pattern recognition for sensor array signals using Fuzzy ARTMAP, Sensors and Actuators B: Chemical 141 (2009) 458–464.
[20] J.H. Cho, P.U. Kurup, Decision tree approach for classification and dimensionality reduction of electronic nose data, Sensors and Actuators B: Chemical 160 (2011) 542–548.
[21] K. Brudzewski, S. Osowski, T. Markiewicz, J. Ulaczyk, Classification of gasoline with supplement of bio-products by means of an electronic nose and SVM neural network, Sensors and Actuators B: Chemical 113 (2006) 135–141.
[22] M. Pardo, G. Sberveglieri, Classification of electronic nose data with support vector machines, Sensors and Actuators B: Chemical 107 (2005) 730–737.
[23] B. Ehret, K. Safenreiter, F. Lorenz, J. Biermann, A new feature extraction method for odour classification, Sensors and Actuators B: Chemical 158 (2011) 75–88.
[24] Y.G. Martín, J.L.P. Pavón, B.M. Cordero, C.G. Pinto, Classification of vegetable oils by linear discriminant analysis of electronic nose data, Analytica Chimica Acta 384 (1999) 83–94.
[25] L. Nanni, A. Lumini, Orthogonal linear discriminant analysis and feature selection for micro-array data classification, Expert Systems with Applications 37 (2010) 7132–7137.
[26] S.K. Jha, R.D.S. Yadava, Denoising by singular value decomposition and its application to electronic nose data processing, IEEE Sensors Journal 11 (2011) 35–44.
[27] L.J. Cao, K.S. Chua, W.K. Chong, H.P. Lee, Q.M. Gu, A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine, Neurocomputing 55 (1–2) (2003) 321–336.
[28] Y. Xiao, H. Wang, W. Xu, J. Zhou, L1 norm based KPCA for novelty detection, Pattern Recognition 46 (1) (2013) 389–396.
[29] K. Hotta, Local co-occurrence features in subspace obtained by KPCA of local blob visual words for scene classification, Pattern Recognition 45 (10) (2012) 3687–3694.
[30] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, Face recognition using kernel direct discriminant analysis algorithms, IEEE Transactions on Neural Networks 14 (1) (2003) 117–126.
[31] J. Yang, A.F. Frangi, J.Y. Yang, D. Zhang, J. Zhong, KPCA plus LDA: a complete kernel Fisher discriminant frame work for feature extraction and recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (2) (2005) 230–244.
[32] S.J. Dixon, Y. Xu, R.G. Brereton, H.A. Soini, M.V. Novotny, E. Oberzaucher, K. Grammer, D.J. Penn, Pattern recognition of gas chromatography mass spectrometry of human volatiles in sweat to distinguish the sex of subjects and determine potential discriminatory marker peaks, Chemometrics and Intelligent Laboratory Systems 87 (2) (2007) 161–172.
[33] F. Lindgren, P. Geladi, S. Wold, The kernel algorithm for PLS, Journal of Chemometrics 7 (1) (1993) 45–59.
[34] J. Karhunen, Generalization of principal component analysis, optimization problems and neural networks, Neural Network 8 (1995) 549–562.
[35] C.W. Hsu, C.J. Lin, A comparison of methods for multiclass support vector machine, IEEE Transactions on Neural Network 13 (2002) 415–425.
[36] L.H. Chiang, M.E. Kotanchek, A.K. Kordon, Fault diagnosis based on Fisher discriminant analysis and support vector machines, Computers & Chemical Engineering 28 (2004) 1389–1401.
[37] F. Sales, M.P. Callao, F.X. Rius, Multivariate standardization for correcting the ionic strength variation on potentiometric sensor arrays, Analyst 125 (2000) 883–888.

Please cite this article in press as: L. Zhang, F.-C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, Anal. Chim. Acta (2014), http://dx.doi.org/10.1016/j.aca.2014.01.049