

A novel sensor feature extraction based on kernel entropy component analysis for discrimination of indoor air contaminants



Xiongwei Peng^a, Lei Zhang^{a,*}, Fengchun Tian^a, David Zhang^b

^a College of Communication Engineering, Chongqing University, 174 ShaZheng street, ShaPingBa District, Chongqing, 400044, China

^b Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

ARTICLE INFO

Article history:

Received 2 February 2015

Received in revised form 20 August 2015

Accepted 2 September 2015

Available online 7 September 2015

Keywords:

Electronic nose

Renyi entropy

Classification

Component analysis

Feature extraction

ABSTRACT

Component analysis techniques for feature extraction in multi-sensor system (electronic nose) have been studied in this paper. A novel nonlinear kernel based Renyi entropy component analysis method is presented to address the feature extraction problem in sensor array and improve the odor recognition performance of E-nose. Specifically, a kernel entropy component analysis (KECA) as a nonlinear dimension reduction technique based on the Renyi entropy criterion is presented in this paper. In terms of the popular support vector machine (SVM) learning technique, a joint KECA–SVM framework is proposed as a system for nonlinear feature extraction and multi-class gases recognition in E-nose community. In particular, the comparisons with PCA, KPCA and ICA based component analysis methods that select the principal components with respect to the largest eigen-values or correlation have been fully explored. Experimental results on formaldehyde, benzene, toluene, carbon monoxide, ammonia and nitrogen dioxide demonstrate that the KECA–SVM method outperforms other methods in classification performance of E-nose. The MATLAB implementation of this work is available online at <http://www.escience.cn/people/lei/index.html>

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Electronic nose (E-nose), as an artificial olfaction system, is an instrument comprised of a chemical sensor array with partial specificity and an appropriate pattern recognition algorithm [1–3]. In recent years, E-nose has an extensive range of applications such as environmental monitoring, medical diagnosis, agriculture, food and pharmaceutical industries [4–9]. However, the performance of these instruments depends heavily on the signal processing and recognition algorithms. In the past decades, we have witnessed the rapid development in pattern recognition and machine learning. In E-nose systems, feature extraction is a key step, which projects the high-dimensional data onto a well chosen low dimensional subspace while preserving the underlying structure of data and improving the discriminative capability of features. In this paper, several typical linear and nonlinear feature extraction methods based on component analysis have been studied for improving the recognition performance of E-nose.

Principal component analysis (PCA), as a well-known unsupervised dimension reduction method, transforms the original data

into the principal component space via a linear projection. PCA projects the correlated variables into another orthogonal feature space through the correlation matrix of original data, and gain a group of new feature subset with the largest variance [10,11]. Another representative linear method is independent component analysis (ICA), which is developed to solve the problem of blind source separation [12]. ICA aims at transforming the observed data into a component space with the maximum independence from each other, and it can also be used as a latent variable model.

However, the problems we face, in practical use, are usually nonlinear with very complex data structure. Based on kernel technique, the nonlinear problem can be solved in a linear way in the high dimensional feature space (i.e. Reproduced Kernel Hilbert Space, RKHS). Kernel PCA, as a non-linear extension of PCA, is one of the representative nonlinear dimension reduction methods [13]. KPCA performs the eigen-decomposition in RKHS, represented by a kernel Gram matrix. Similar to PCA, KPCA extracts principal components based on the top eigenvalues and eigenvectors of the kernel matrix. The applications of KPCA have been shown in many areas, such as de-noising, pattern recognition [14–17] etc.

In this paper, a novel feature extraction method, kernel entropy component analysis in multi-sensor system is presented for improving the classification performance of support vector machine (SVM), which is one of the most popular machine learn-

* Corresponding author. Fax: +86 23 65103544.

E-mail address: leizhang@cqu.edu.cn (L. Zhang).

ing methods and has been successfully used in various classification and regression problems [22]. The principles of KECA can be found in [18]. Different from other component analysis methods, the KECA aims at learning a basis transformation based on maximum entropy criterion but not the largest eigen-value, in the Reproduced Kernel Hilbert Space. In general, the selected principal components do not necessarily correspond to the largest eigenvalues of the kernel matrix [19]. In addition, another contribution of this paper is the proposed KECA-SVM framework, as a system with feature extraction and recognition for E-nose application.

The remainder of the paper is organized as follows: Section 2 gives a brief review of KPCA. In Section 3, KECA is introduced and the difference from KPCA is analyzed. The recognition method using support vector machine is provided in Section 4. The experiments including experimental setup and experimental data are described in Section 5. The results and comparisons are presented in Section 6. Finally, Section 7 concludes the paper.

2. Kernel PCA: a review

KPCA is a popular nonlinear dimension reduction method. The basic idea is to map the input data into a high-dimensional feature space via a nonlinear function, where PCA can be implemented. Suppose that the original input data $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathcal{R}^d \times N$, and the nonlinear mapping from the input space to a high-dimensional space \mathcal{H} (Reproducing Kernel Hilbert Space, RKHS) is defined as $\Phi(\cdot): \mathcal{R}^d \rightarrow \mathcal{H}$. Induced by Mercer kernel theorem, the kernel Gram matrix \mathbf{K} can be calculated by inner product, i.e. $K_{i,j} = \Phi^T(\mathbf{x}_i) \Phi(\mathbf{x}_j) = \kappa(\mathbf{x}_i, \mathbf{x}_j)$, where $\kappa(\cdot)$ is a kernel function. In summary, KPCA training consists of two steps: (1) compute the kernel Gram matrix $\mathbf{K} \in \mathcal{R}^N \times N$, where N is the number of training samples; (2) perform the eigen-decomposition of \mathbf{K} , where the top l eigenvectors with respect to the first l largest eigenvalues are considered to span a low dimension subspace \mathbf{P} .

3. Kernel entropy component analysis

3.1. Principle of KECA

The Renyi quadratic entropy is defined as

$$H(p) = -\log \int p^2(\mathbf{x}) d\mathbf{x} \quad (1)$$

where $p(\mathbf{x})$ is the probability density function of data sampling. Due to the monotonic nature of logarithmic function, we consider the following equation

$$V(p) = \int p^2(\mathbf{x}) d\mathbf{x} \quad (2)$$

Algorithm 1: KECA

Input: The original feature matrix is presented by $\mathbf{X} \in \mathcal{R}^{d \times M}$;

Output: The kernel ECA feature matrix \mathbf{P} ;

Procedure:

1. Compute the kernel matrix $\mathbf{K} \in \mathcal{R}^{M \times M}$ by using the kernel function.
2. Compute the Eigen-decomposition of \mathbf{K} and obtain eigen-value $\mathbf{D} = [\lambda_1, \dots, \lambda_M]$ and the eigenvector $\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_M]$.
3. Compute the Renyi entropy estimate of the each eigen-value by

$$\text{entropy}(i) = \lambda_i \cdot \|\mathbf{e}_i\|_2^2, i = 1, \dots, M$$
4. Sort the *entropy* in descending order and get the rearranged eigen-value \mathbf{D}' and eigenvector \mathbf{E}' .
5. Obtain the kernel ECA feature matrix $\mathbf{P} = \mathbf{K} \times \mathbf{E}'$.

Fig. 1. Details of the KECA algorithm.

In order to estimate $V(p)$, the Parzen window density function is used, as suggested in [19]

$$\hat{p}(\mathbf{x}) = \frac{1}{N} \sum_k k_\sigma(\mathbf{x}, \mathbf{x}_k) \quad (3)$$

where $k_\sigma(\mathbf{x}, \mathbf{x}_k)$ is the Parzen window, or kernel centered at \mathbf{x}_k and its width can be represented by the kernel parameter σ , which must be a density function itself [20,21]. Hence, there is

$$\hat{V}(p) = \frac{1}{N} \sum_k \hat{p}(\mathbf{x}_k) = \frac{1}{N} \sum_k \frac{1}{N} \sum_k k_\sigma(\mathbf{x}_k, \mathbf{x}_{k'}) = \frac{1}{N^2} \mathbf{1}^T \mathbf{K} \mathbf{1} \quad (4)$$

where element $K_{k,k'} = k_\sigma(\mathbf{x}_k, \mathbf{x}_{k'})$ and $\mathbf{1}$ is a unit vector with length N .

In addition, the Renyi entropy estimator can be expressed in terms of the eigenvalues and eigenvectors of the kernel matrix, as follows

$$\mathbf{K} = \mathbf{E} \mathbf{D} \mathbf{E}^T \quad (5)$$

where $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_N)$ is a diagonal matrix, and the columns of \mathbf{E} are the eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N$ with respect to $\lambda_1, \dots, \lambda_N$. By substituting Eq. (5) into (4), we can obtain

$$\hat{V}(p) = \frac{1}{N^2} \sum_{i=1}^N \left(\sqrt{\lambda_i} \mathbf{e}_i^T \mathbf{1} \right)^2 \quad (6)$$

we can see from Eq. (6) that each λ_i and \mathbf{e}_i have joint contribution to the entropy estimation, thus it is easy to find those eigenvalues and the eigenvectors with the most contribution to the entropy estimation. Hence, in KECA, the eigenvectors contributing the most in Eq. (6) will be selected for projection.

Note that from Eq. (4) we see that the Renyi entropy estimation obtained based on the available samples fully depends on the

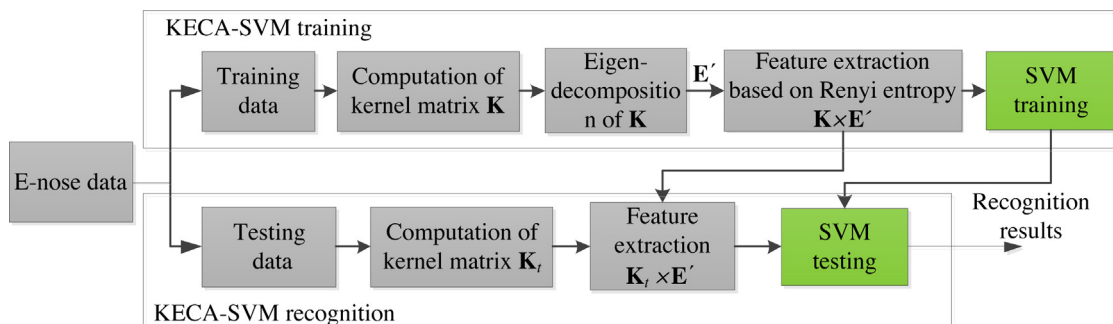


Fig. 2. Diagram of the proposed KECA-SVM framework.

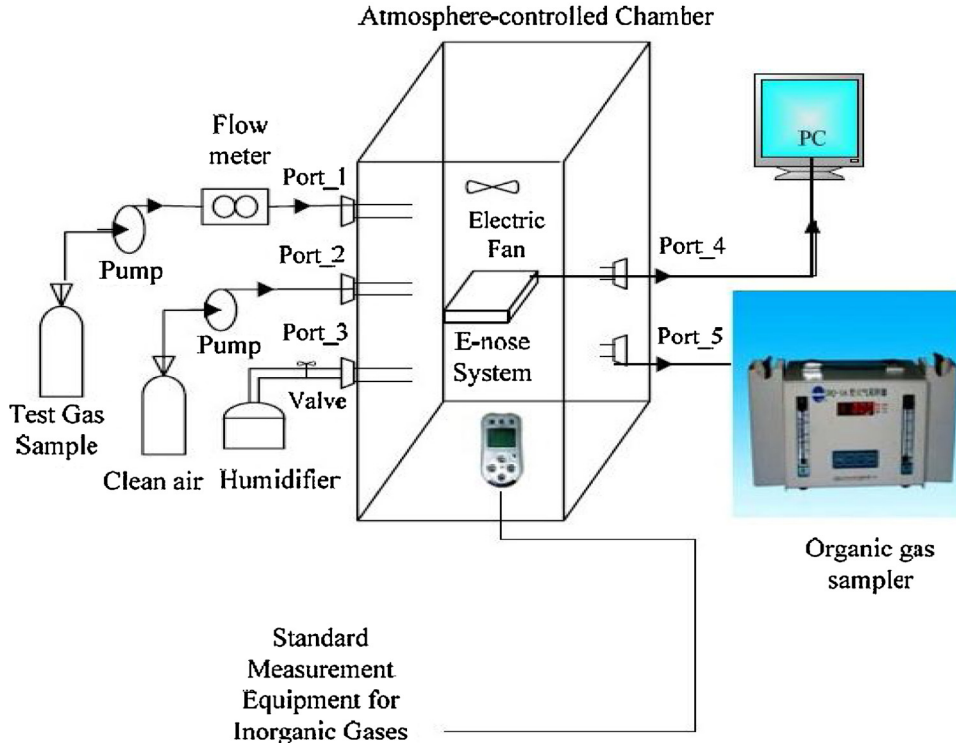


Fig. 3. The experimental platform of E-nose.

kernel matrix \mathbf{K} , which is induced by Parzen window for density estimation. If we use the first order analysis based on the Shannon entropy with Eq. (1), the kernel matrix \mathbf{K} would not be used. That is the reason why we use the Renyi entropy in second order as the information metric in this work.

3.2. The difference between KECA and KPCA

For KPCA, the eigenvectors with respect to the top eigen-values of the kernel matrix are selected for low-dimensional subspace projection. However, the e-nose data is very complex, and there is minimal difference between the gases i.e. the patterns are very similar and linear inseparable. So the variable with the largest variance in KPCA does not always represent the discriminative features in E-nose data.

Different from KPCA, KECA extracts the low-dimensional feature by not only considering the magnitude of eigen-values, but also the eigenvectors, which can find the discriminative feature from E-nose data and better reveals the cluster structure. However, KPCA only considers the ranking of eigen-values, such that the latent discriminative feature of the data may be lost. In this paper, our presented method KECA can effectively address this bottleneck problem by extracting the features with the most contribution to the Renyi entropy. The KECA is summarized as Algorithm 1 (see Fig. 1).

4. Recognition

In this paper, for recognition of six indoor air contaminants, a hybrid model KECA-SVM is developed. The model and decision function of SVM [30] are briefly introduced as follows. Given a training set of N data points $\{\mathbf{x}_i, y_i\}_{i=1}^N$, where the label $y_i \in \{-1, 1\}$, $i =$

$1, \dots, N$. SVM aims at solving the following minimization problem with inequality constraint,

$$\min_{\mathbf{w}, \xi_i} \frac{1}{2} \|\mathbf{w}\|^2 + C \cdot \sum_{i=1}^N \xi_i, \tag{7}$$

$$s.t. \xi_i \geq 0, y_i [\mathbf{w}\varphi(\mathbf{x}_i) + b] \geq 1 - \xi_i$$

where $\varphi(\cdot)$ is a linear/nonlinear mapping function, \mathbf{w} and b are the parameters of classifier hyper-plane. Generally, for optimization, the original problem Eq. (7) of SVM can be transformed into its dual formulation with equality constraint by using Lagrange multiplier method. Its dual problem of Eq. (7) can be formulated as

$$\max_{\alpha} \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} y_i y_j \alpha_i \alpha_j \varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}_j) \tag{8}$$

$$s.t. \sum_{i=1}^N \alpha_i y_i = 0, 0 \leq \alpha_i \leq C$$

The deduction of its dual formulation can be referred to [30]. The formulation of Eq. (8) is the dual problem of SVM, which is a quadratic programming (QP) problem. The α can be obtained by solving Eq. (8), then the decision function of SVM for a new sample can be described as

$$f(\mathbf{x}) = \text{sgn} \left(\sum_{i=1}^M \alpha_i y_i \kappa(\mathbf{x}_i, \mathbf{x}) + b \right) \tag{9}$$

where α_i and b are the optimal decision parameters. $\kappa(\cdot)$ is a kernel function. In this paper, the Gaussian kernel function [23] defined as $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / \sigma^2)$ is used. The flowchart of the proposed method (KECA-SVM) for feature extraction and recognition is shown in Fig. 2.

5. Experiments

5.1. Experimental setup

In previous work [24–27], the E-nose system has been presented. The E-nose consists of pattern recognition and metal oxide

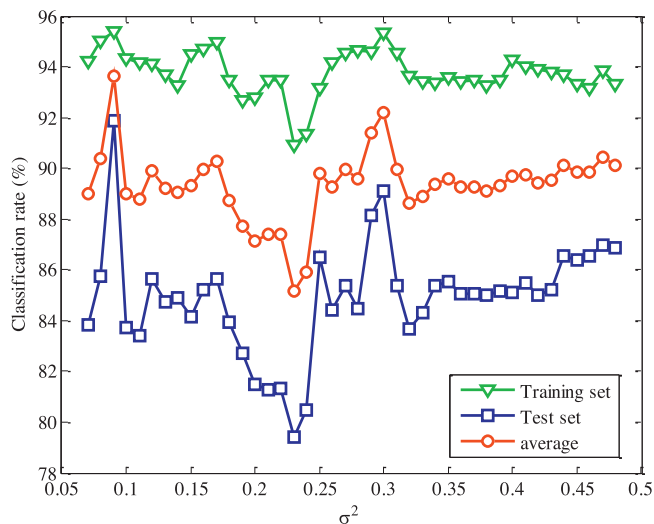


Fig. 4. Classification rates of training and test sets with different σ^2 .

semiconductor gas sensors array, in which five different types of gas sensors including four TGS sensors (i.e. TGS2602, TGS2620, TGS2201A and TGS2201B) and a module (SHT2230 of Sensirion in Switzerland) with two auxiliary sensors for temperature (T) and humidity (H) perception are considered. The sensors are installed on a designed printed circuit board (PCB) with associated electrical components, such that the sensor array can produce a group of odor patterns with six variables. The experimental platform of this work is shown in Fig. 3 [26].

5.2. Experimental data

In order to verify the effectiveness of the presented KECA–SVM framework, we investigate six common indoor air contaminants including formaldehyde (HCHO), benzene (C_6H_6), toluene (C_7H_8), carbon monoxide (CO), ammonia (NH_3) and nitrogen dioxide (NO_2). All experiments were employed by an E-nose in a chamber with temperature and humidity controlled. In experiment, a gas bag collected with target gas and nitrogen (N_2) was prepared for injection into the chamber. Note that N_2 is used to dilute the gas concentration in the gas bag, and we can obtain various samples with different concentrations by controlling the injection time. For each experiment, 12 min (2 min for baseline and 10 min for response) were consumed for sampling and extra 15 min were needed for chamber cleaning. Totally, 503 samples including 208 formaldehyde samples, 132 benzene samples, 53 toluene samples, 47 carbon monoxide samples, 39 ammonia samples and 24 nitrogen dioxide samples were obtained. Considering the ambient condition indoor, the experiments were measured at ambient temperature of 15, 25, 30 and 35 °C and humidity of 40%, 60%, and 80% RH. Note that the output of the sensors in our electronic nose system is the voltage (mV) on the sensitive resistance of sensors. Therefore, the extracted sensor feature as input that will feed into the KECA algorithm is the normalized voltage (i.e. sensor response) between 0 and 1.

5.3. Training protocol of algorithm

In the training phase of KECA–SVM, all the samples are divided into two parts: training set and testing set. Firstly, 75% of the samples are selected as training set by using the Kennard–Stone sequential (KSS) sample selection algorithm, such that the samples can cover the multidimensional space by maximizing the Euclidean distance [28]. Second, the remaining samples are used as test set.

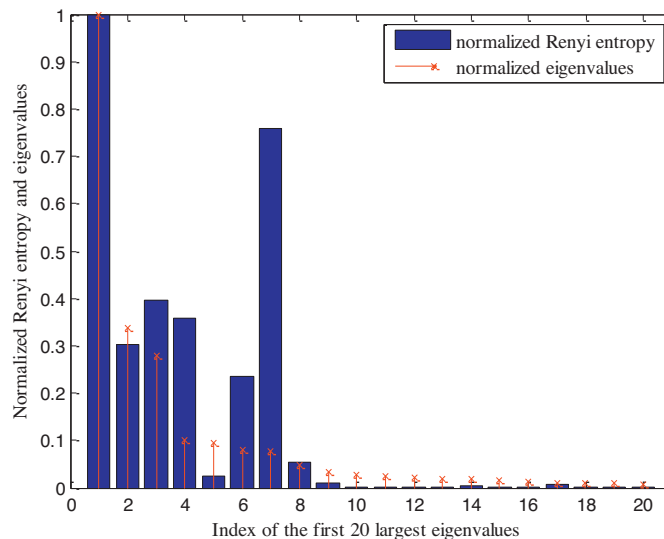


Fig. 5. Normalized Renyi entropy and eigenvalues.

Table 1
Distribution of experimental data.

Two subsets	Number of samples in the subset					
	HCHO	C_6H_6	C_7H_8	CO	NH_3	NO_2
Training set	156	99	40	35	29	18
Test set	52	33	13	12	10	6
Total	208	132	53	47	39	24

The distribution of training set and test set for each class is shown in Table 1.

6. Results and comparisons

6.1. Classification with KECA–SVM

In this section, we evaluate the effectiveness of the KECA feature extraction by calculating the average gas recognition accuracy of the E-nose. It can be seen from the Eq. (4) that, a kernel parameter σ is used in KECA for kernel matrix. A new method that uses a kernel parameter σ in the order of 10–20% of the total Euclidean distances between the feature vectors was proposed by Shi and Malik [29]. The usage of the median range is a better choice. Therefore, in this work, 10–20% of the median range is $0.07 \leq \sigma^2 \leq 0.48$. We run the KECA–SVM program over the range of kernel parameter, and the average classification rate is shown.

Fig. 4 illustrates the variation trend of classification accuracy versus σ^2 with an interval of 0.01. We can observe that when $\sigma^2 = 0.09$, the average classification rate is higher than 93%, which gives the best classification result. We can also see that for the best σ -value, both training set and test set can achieve the best recognition rate.

Table 2 shows the eigen-values and the Renyi entropy in terms of the first 20 principal components with respect to the top 20 largest eigen-values. We can see that large eigen-values do not give large Renyi entropy. The first principal component (PC1) gets the largest entropy, but the PC2 does not correspond to the second largest entropy, instead, the PC7 obtains the second largest entropy. In addition, for instance, PC10 ranks the 22nd largest Renyi entropy. For better visualization of the results, the normalized eigenvalues in descent order is described as the vertical lines in Fig. 5, in which the bars denote the normalized entropy terms $(\sqrt{\lambda_i} \alpha_i^T 1)^2$.

Table 2
Insight of KECA principal components.

PCs	Eigenvalue	Renyi entropy	Entropy terms
PC1	162.14	186.64	1
PC2	54.89	56.32	5
PC3	45.11	73.91	3
PC4	16.13	66.87	4
PC5	15.36	4.57	8
PC6	12.88	43.79	6
PC7	12.5	140.92	2
PC8	8.04	9.86	7
PC9	5.4	2.04	9
PC10	4.66	0.02	22
PC11	3.87	0.29	14
PC12	3.73	0.19	23
PC13	3.25	0.13	18
PC14	3.09	0.73	11
PC15	2.67	0.05	20
PC16	2.01	0.25	16
PC17	1.87	1.5	10
PC18	1.65	0.31	13
PC19	1.46	0.28	15
PC20	1.19	0.01	25

6.2. Comparisons between KECA and KPCA

For comparisons between KECA and KPCA, we present the results of dimension reduction and classification. For visualization, two 2D scatter sub-plots (PC-1 vs PC-2 and PC-1 vs PC-7) and two 3D scatter sub-plots are illustrated in Fig. 6 and Fig. 7, respectively. Fig. 6(a) shows a 2D scatter subplot of the projections onto the first two principal components from the KPCA, and Fig. 6(b) illustrates the projections onto the leading two principal components contributing more to the entropy extracted using KECA. From the 2D and 3D scatter plots, we can see that the KECA based clustering shows a significant improvement over that of KPCA. Particularly, the C_7H_8 and C_6H_6 patterns are more separable with other gases after KECA. This implies that the entropy-based components are better than traditional principal components based on magnitude of eigen-values for feature extraction. Numerically, we have implemented the recognition based on KPCA (PC1 vs. PC2) and KECA (PC1 vs. PC7). The average recognition accuracy of the six kinds of gases for KPCA is 42.1% and that of the KECA is 46.8%. The result denotes that the component analysis based on the largest entropies is much better than that based on largest eigenvalues.

Table 3
Comparisons of classification accuracy for different kernel functions.

Class	Classification accuracy (%)							
	KPCA				KECA			
	Linear	Poly	Cos	RBF	Linear	Poly	Cos	RBF
HCHO	82.7	92.3	93.3	90.4	86.5	92.3	88.5	92.3
C_6H_6	69.7	75.8	81.8	75.7	54.5	81.8	81.8	75.8
C_7H_8	76.9	100	76.9	100.0	92.3	100.0	76.9	100.0
CO	66.7	91.7	100.0	75.0	66.7	100.0	100.0	83.3
NH_3	50.0	50.0	80.0	100.0	60.0	60.0	80.0	100.0
NO_2	50.0	83.3	66.7	83.3	50.0	66.7	83.3	100.0
Mean	65.9	82.1	83.0	87.4	68.5	83.5	85.1	91.9

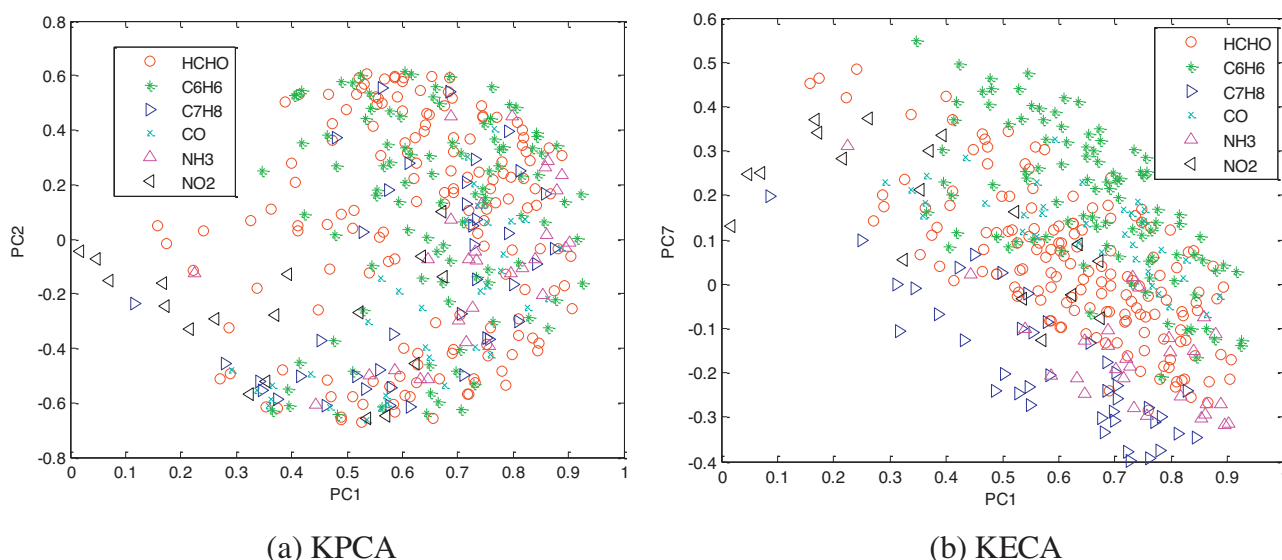
Fig. 8 shows the trend of the average classification accuracy with an increasing number of components used in SVM training and testing. It can be clearly observed that our proposed KECA-SVM outperforms the KPCA based feature extraction for classification. The best average classification accuracy is obtained when 13 principal components are used. Note that the number of selection components corresponds to the number of dimensions after feature extraction.

6.3. Comparisons of kernel functions

The choice of kernel function plays an important role in discriminative and generalization capability. In order to show the performance of different kernel functions, we compare several types of kernel functions: linear, polynomial (poly), cosine (cos) and Gaussian radial basis function (RBF) in Table 3. From the results, we can see that Gaussian RBF function shows the best performance with an average recognition accuracy of 91.9% for KECA-SVM. Cosine function achieves the second best, and linear function shows the worst performance. However, the performance of KECA is always better than KPCA with respect to any kernel function.

6.4. Comparisons with other algorithms

In order to compare KECA with other representative feature extraction algorithms, PCA and ICA are also used. The popular k nearest neighbor (KNN) and SVM classifiers are used for classification. The best results of different feature extraction methods

**Fig. 6.** The 2-D scatter plots of KECA and KPCA for feature extraction.

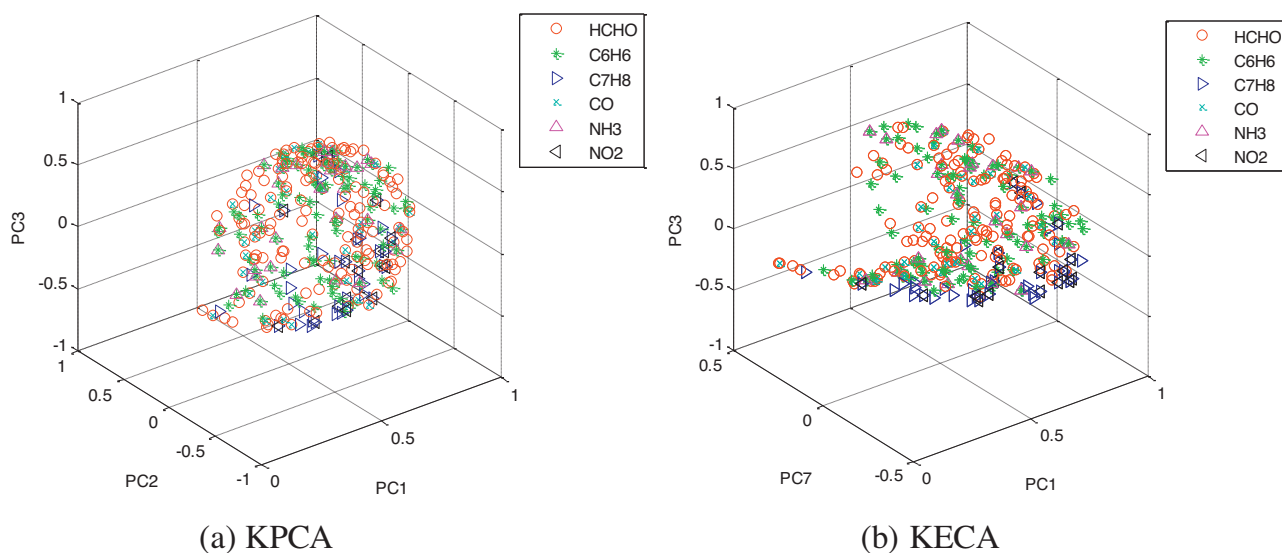


Fig. 7. The 3-D scatter plots of KECA and KPCA for feature extraction.

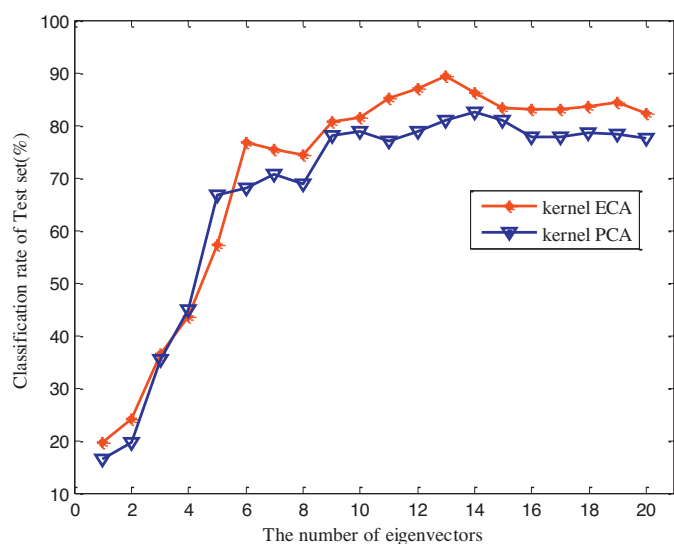


Fig. 8. Average classification rates with KPCA and KECA feature extraction.

Table 4
Recognition accuracy using different feature extraction methods.

Class	Classification accuracy (%)									
	KNN					SVM				
	Raw	ICA	PCA	KPCA	KECA	Raw	ICA	PCA	KPCA	KECA
HCHO	86.5	80.8	73.1	84.6	88.5	86.5	90.4	82.7	90.4	92.3
C ₆ H ₆	48.5	42.4	51.5	42.4	51.5	81.8	51.5	84.8	75.7	75.8
C ₇ H ₈	53.9	53.9	53.9	46.2	46.2	100	69.2	84.6	100.0	100.0
CO	58.3	41.7	58.3	58.3	58.5	58.3	75.0	58.3	75.0	83.3
NH ₃	50.0	50.0	70.0	60.0	60.0	90.0	50.0	70.0	100.0	100.0
NO ₂	50.0	33.3	33.3	50.0	50.0	50.0	50.0	50.0	83.3	100.0
Average	57.9	50.3	56.7	56.9	59.1	77.8	64.4	71.7	87.4	91.9

are shown in Table 4, from which, we can find that the presented KECA is superior to other feature extraction methods. For almost all gases, the KECA–SVM method can obtain the best recognition accuracy, except the recognition rate 75.8% of C₆H₆. For C₇H₈, NH₃ and NO₂, the recognition accuracy can reach 100% using our proposed method. Compared with PCA and ICA, KPCA and KECA are significantly better. This implies that nonlinear feature extraction

Table 5
Multi-class classification results using the proposed KECA–SVM method.

N	Class	Classified as					
		HCHO	C ₆ H ₆	C ₇ H ₈	CO	NH ₃	NO ₂
52	HCHO	48	4	0	0	0	0
33	C ₆ H ₆	5	25	0	3	0	0
13	C ₇ H ₈	0	0	13	0	0	0
12	CO	0	0	0	10	0	2
10	NH ₃	0	0	0	0	10	0
6	NO ₂	0	0	0	0	0	6

method is more competent. However, the nonlinear KPCA method gives the second best classification accuracy of 87.4%, and KECA achieves the best recognition performance of 91.9%. Specifically, Table 5 presents the classification results of test set using KECA method. The digits in diagonal line denote the number of correctly classified samples, and others mean the number of misclassified samples.

The results and comparisons in this work fully demonstrate that the nonlinear kernel method combined with entropy-based component analysis can select the principal components with better representation of odor patterns, and both structural and latent entropy information in the data can be effectively preserved for robust classification.

7. Conclusion

In this paper, a nonlinear feature extraction method in E-nose, called Kernel Entropy Component Analysis (KECA), was presented with SVM for classification of multiple indoor air contaminants by an E-nose. Different from PCA method that extracts the principal components based on the magnitude of eigen-values, KECA is based on the most entropy preserving axes, which not only chooses the best principal components with the largest entropy criterion but also reveals the latent cluster structure of the data. The performance of the proposed KECA–SVM was compared with the recognition performance based on traditional component analysis methods, and experimental results demonstrate that KECA outperforms other methods in classification of multiple indoor air contaminants by an E-nose.

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 the National Natural Science Foundation of China under Grant 61401048, the 2013 Innovative Team Construction Project of Chongqing Universities, Hong Kong Scholar Program under Grant XJ2013044, and the China Post-Doctoral Science Foundation under Grant 2014M550457.

References

- [1] J.W. Gardner, P.N. Bartlett, A brief history of electronic nose, *Sens. Actuators B* 18 (1994) 211–220.
- [2] J.E. Haugen, K. Kvaal, Electronic nose and artificial neural network, *Meat Sci.* 49 (1998) S273–S286.
- [3] L. Carmel, N. Sever, D. Lancet, D. Harel, An eNose algorithm for identifying chemicals and determining their concentration, *Sens. Actuators B* 93 (2003) 77–83.
- [4] G.C. Green, A.D.C. Chan, H.H. Dan, M. Lin, Using a metal oxide sensor (MOS)-based electronic nose for discrimination of bacteria based on individual colonies in suspension, *Sens. Actuators B* 152 (2011) 21–28.
- [5] C.H. Shih, Y.J. Lin, K.F. Lee, P.Y. Chien, P. Drake, Real-time electronic nose based pathogen detection for respiratory intensive care patients, *Sens. Actuators B* 148 (2010) 153–157.
- [6] M. Bernabei, G. Pennazza, M. Santortico, C. Corsi, C. Roscioni, R. Paolesse, C. DiNatale, A. D'Amico, A preliminary study on the possibility to diagnose urinary tract cancers by an electronic nose, *Sens. Actuators B* 131 (2008) 1–4.
- [7] H.M. Zhang, J. Wang, Detection of age and insect damage incurred by wheat with an electronic nose, *J. Stored Prod. Res.* 43 (2007) 489–495.
- [8] E.A. Baldwin, J. Bai, A. Plotto, S. Dea, Electronic noses and tongues applications for the food and pharmaceutical industries, *Sensors* 11 (2011) 4744–4766.
- [9] A.H. Gomez, J. Wang, G.X. Hu, A.G. Pereira, Monitoring storage shelf life of tomato using electronic nose technique, *J. Food Eng.* 85 (2008) 625–631.
- [10] I. Jolliffe, *Principal Component Analysis*, John Wiley & Sons, Ltd., 2005.
- [11] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, *Chemom. Intell. Lab. Syst.* 2 (1) (1987) 37–52.
- [12] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, *Neural Netw.* 13 (4) (2000) 411–430.
- [13] K.I. Kim, K. Jung, H.J. Kim, Face recognition using kernel principal component analysis, *IEEE Signal Process. Lett.* 9 (2) (2002) 40–42.
- [14] S. Mika, B. Scholkopf, A. Smola, et al., Kernel PCA and De-Noising in Feature Spaces, *Adv. Neural Inf. Process. Syst.* 11 (1999) 536–542.
- [15] R. Rosipal, M. Girolami, L.J. Trejo, et al., Kernel PCA for feature extraction and de-noising in nonlinear regression, *Neural Comput. Appl.* 10 (3) (2001) 231–243.
- [16] C. Liu, Gabor-based kernel PCA with fractional power polynomial models for face recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (5) (2004) 572–581.
- [17] M.H. Yang, N. Ahuja, D. Kriegman, Face recognition using kernel eigenfaces, *IEEE Int. Conf. Image Process.* 1 (2000) 37–40.
- [18] R. Jenssen, Kernel entropy component analysis, *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (5) (2010) 847–860.
- [19] E. Parzen, On estimation of a probability density function and mode, *Ann. Math. Stat.* 33 (3) (1962) 1065–1076.
- [20] A. Rnyi, On measures of entropy and information, *Fourth Berkeley Symp. Math. Stat. Probab.* (1961) 547–561.
- [21] B.W. Silverman, *Density Estimation for Statistics and Data Analysis*, CRC press, 1986.
- [22] T. Joachims, *Text Categorization with Support Vector Machines: Learning with Many Relevant Features*, Springer Berlin Heidelberg, 1998.
- [23] M. Girolami, Mercer kernel-based clustering in feature space, *IEEE Trans. Neural Netw.* 13 (3) (2002) 780–784.
- [24] 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 Actuators B Chem.* 160 (2011) 760–770.
- [25] 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 Actuators B Chem.* 174 (2012) 114–125.
- [26] L. Zhang, F. Tian, S. Liu, J. Guo, B. Hu, Q. Ye, L. Dang, X. Peng, C. Kadri, J. Feng, Chaos based neural network optimization for concentration estimation of indoor air contaminants by an electronic nose, *Sensors Actuators A Phys.* 189 (2013) 161–167.
- [27] L. Zhang, F. Tian, A new kernel discriminant framework for electronic nose recognition, *Anal. Chim. Acta* 816 (2014) 8–17.
- [28] 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.
- [29] J. Shi, J. Malik, Normalized cuts and image segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (8) (2000) 888–905.
- [30] C. Cortes, V. Vapnik, Support-vector networks, *Mach. Learn.* 20 (3) (1995) 273–297.

Biographies



Xiongwei Peng received his Bachelor degree in Communication Engineering in 2012 from the Chongqing University, China; he is now pursuing his MS degree in circuits and system. His research interests include artificial neural network and optimizations.



Lei Zhang received his Ph.D degree in Circuits and Systems from the College of Communication Engineering, Chongqing University, Chongqing, China, in 2013. From 2015, he is a Distinguished Research Fellow with Chongqing University. He was selected as a Hong Kong Scholar in China in 2013. He is currently also a Post-Doctoral Fellow with The Hong Kong Polytechnic University, Hong Kong. He has authored more than 30 scientific papers in electronic nose, machine olfaction, pattern recognition, and machine learning. His current research interests include machine learning, pattern recognition, artificial olfactory system, and nonlinear signal processing in Electronic Nose. Dr. Zhang was a recipient of Outstanding Doctoral Dissertation Award of Chongqing, China, in 2015, Hong Kong Scholar Award in 2014, Academy Award for Youth Innovation of Chongqing University in 2013 and the New Academic Researcher Award for Doctoral Candidates from the Ministry of Education, China, in 2012.



Fengchun Tian received PhD degree in 1997 in electrical engineering from Chongqing University. He is currently a Full-Professor with the College of Communication Engineering of Chongqing University. His research interests include Electronic nose technology, artificial olfactory systems, pattern recognition, chemical sensors, signal/image processing, wavelet, and computational intelligence. In 2006 and 2007, he was recognized as a part-time Professor of GUELPH University, Canada.



David Zhang graduated in Computer Science from Peking University. He received his MSc in 1982 and his PhD in 1985 in Computer Science from the Harbin Institute of Technology (HIT), respectively. From 1986 to 1988 he was a Postdoctoral Fellow at Tsinghua University and then an Associate Professor at the Academia Sinica, Beijing. In 1994 he received his second PhD in Electrical and Computer Engineering from the University of Waterloo, Ontario, Canada. He is a Chair Professor since 2005 at the Hong Kong Polytechnic University where he is the Founding Director of the Biometrics Research Centre (UGC/CRC) supported by the Hong Kong SAR Government in 1998. He also serves as Visiting Chair Professor in Tsinghua University, and Adjunct Professor in Peking University, Shanghai Jiao Tong University, HIT, and the University of Waterloo. He is the Founder and Editor-in-Chief, *International Journal of Image and Graphics (IJIG)*; Book Editor, Springer International Series on Biometrics (KISB); Organizer, the International Conference on Biometrics Authentication (ICBA); Associate Editor of more than ten international journals including *IEEE Transactions* and so on; and the author of more than 10 books, over 300 international journal papers and 30 patents from USA/Japan/HK/China. Professor Zhang is a Croucher Senior Research Fellow, Distinguished Speaker of the IEEE Computer Society, and a Fellow of both IEEE and IAPR.