Sensors and Actuators A 234 (2015) 143–149

Contents lists available at ScienceDirect

Sensors and Actuators A: Physical

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

# 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



## 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).

http://dx.doi.org/10.1016/j.sna.2015.09.009
0924-4247/

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 $\boldsymbol{X} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N] \in \Re d \times N$, and the nonlinear mapping from the input space to a high-dimensional space $\mathscr{H}$ (Reproducing Kernel Hilbert Space, RKHS) is defined as $\Phi(\cdot) : \Re d \to \mathscr{H}$. Induced by Mercer kernel theorem, the kernel Gram matrix $\mathbf{K}$ can be calculated by inner product, i.e. $K_{i,j} = \Phi^{\mathrm{T}}\left(\boldsymbol{x}_j\right)\Phi\left(\boldsymbol{x}_j\right) = \kappa\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)$, where $\kappa(\cdot)$ is a kernel function. In summary, KPCA training consists of two steps: (1) compute the kernel Gram matrix $\boldsymbol{K} \in \Re N \times N$, 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

$$H(p) = -\log \int p^2(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \tag{1}$$

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

$$V(p) = \int p^2(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \tag{2}$$

**Algorithm 1: KECA**

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

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

**Procedure:**

1. Compute the kernel matrix $\mathbf{K} \in \Re^{M\times M}$ by using the kernel function.
2. Compute the Eigen-decomposition of $\mathbf{K}$ and obtain eigen-value $\mathbf{D} = [\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_M]$ and the eigenvector $\mathbf{E} = [\mathbf{e}_1, \cdots, \mathbf{e}_M]$.
3. Compute the Renyi entropy estimate of the each eigen-value by
$$\boldsymbol{entropy}(i) = \boldsymbol{\lambda}_i \cdot \|\mathbf{e}_i\|_2^2, \boldsymbol{i} = \mathbf{1}, \cdots, \boldsymbol{M}$$
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]

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

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

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

where element $K_{k,k'} = k_\sigma(\boldsymbol{x}_k, \boldsymbol{x}_{k'})$ and $\mathbf{1}$ is an 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

$$\boldsymbol{K} = \mathbf{E}\boldsymbol{D}\boldsymbol{E}^{\mathrm{T}} \tag{5}$$

where $\boldsymbol{D} = \mathbf{diag}(\lambda_1, \cdots, \lambda_N)$ is a diagonal matrix, and the columns of $\mathbf{E}$ are the eigenvectors $\mathbf{e}1, \mathbf{e}2, \ldots, \mathbf{e}N$ with respect to $\lambda_1, \cdots, \lambda_N$. By substituting Eq. (5) into (4), we can obtain

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

we can see from Eq. (6) that each $\lambda_i$ and $\boldsymbol{e}_i$ 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

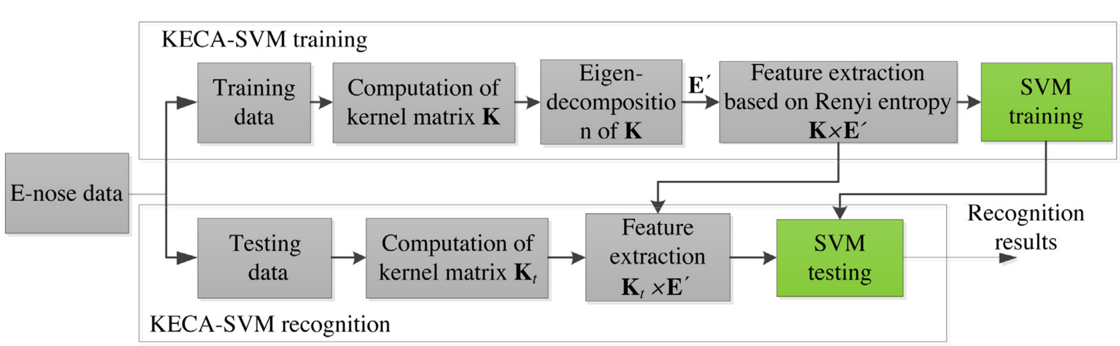


**Fig. 2.** Diagram of the proposed KECA-SVM framework.