



Anti-drift in E-nose: A subspace projection approach with drift reduction



Lei Zhang*, Yan Liu, Zhenwei He, Ji Liu, Pingling Deng, Xichuan Zhou

College of Communication Engineering, Chongqing University, Shazheng street No. 174, Shapingba district, Chongqing 400044, China

ARTICLE INFO

Article history:

Received 5 January 2017

Received in revised form 29 May 2017

Accepted 23 June 2017

Available online 26 June 2017

Keywords:

Anti-drift

Electronic nose

Subspace projection

Common subspace

Machine learning

ABSTRACT

Anti-drift is an emergent and challenging issue in sensor-related subjects. In this paper, we propose to address the time-varying drift (e.g. electronic nose drift), which is sometimes an ill-posed problem due to its uncertainty and unpredictability. Considering that drift is with different probability distribution from the regular data, a machine learning based subspace projection approach is proposed. The main idea behind is that given two data clusters with different probability distribution, we tend to find a latent projection P (i.e. a group of basis), such that the newly projected subspace of the two clusters is with similar distribution. In other words, drift is automatically removed or reduced by projecting the data onto a new common subspace. The merits are threefold: 1) the proposed subspace projection is unsupervised; without using any data label information; 2) a simple but effective domain distance is proposed to represent the mean distribution discrepancy metric; 3) the proposed anti-drift method can be easily solved by Eigen decomposition; and anti-drift is manifested with a well solved projection matrix in real application. Experiments on synthetic data and real datasets demonstrate the effectiveness and efficiency of the proposed anti-drift method in comparison to state-of-the-art methods.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Drift is a well-known issue in instrumentation and measurement. In olfaction related gas sensing instruments (e.g. electronic nose), sensor drift shows nonlinear dynamic behavior in a multi-dimensional sensor array [1]. Sensor drift effect is caused by many objective factors such as aging, poisoning, and the fluctuations of the ambient environmental variables (e.g. humidity, temperature) [2]. Temperature and humidity can also be understood as cross-sensitivity (i.e. gas sensor is also sensitive to ambient temperature and humidity, but not only odors) in E-nose. Due to the ambient variables, as a result, the instrument responds differently to a constant concentration of some contaminants at different ambient conditions. Drift effect can be understood as some uncertainty, which is difficult to intuitively estimate and compensate by using some deterministic approaches. Drift is once thought to be an ill-posed problem due to the very irregular characteristics. Therefore, motivated by these findings, we are inspired to propose an instrumental drift adaptation method from the angle of distribution consistency. That is, we attempt to improve the distribution consistency of the observed cycle data, rather than estimate the drift effect with uncertainty. Before the proposed idea, a mini-overview

of electronic nose (E-nose) is first presented as follows, so that readers who are not very familiar to E-nose community can have a quick scan.

Electronic nose, as a gas sensing device with pattern recognition unit for artificial olfaction, has witnessed a wide progress in applications, systems and algorithms during the past two decades [3–7]. A general E-nose system is shown in Fig. 1. Flammini et al. proposed a low-cost interface to high-value resistive sensors over a wide range, such that a wide detection threshold is possible [8]. Fort et al. discussed different measurement techniques for gas mixture detection by using tin oxide gas sensing device [9]. Brudzewski et al. [10] proposed a differential electronic nose for recognition of coffees. Herrero-Carrón et al. [11] proposed an active and inverse temperature modulation electronic nose for odorant classification. Hossein-Babaei and Amini proposed a single generic tin oxide gas sensor for complex odor recognition [12]. Palit et al. proposed an electronic tongue for tea taste and classification of tea taster's mark [13], by approximating human's taste. Recently, pattern recognition algorithms have been widely studied in E-nose for classification and regression of multiple kinds of analytes [14–16], such as support vector machines (SVM), neural networks (ANN), discriminant analysis (DA), learning vector quantization (LVQ), etc. Tudu et al. proposed an incremental fuzzy approach for classification of black tea quality with an E-nose [17], in which the classifier has incremental learning ability such that the newly presented patterns can be automatically included in the training set with-

* Corresponding author.

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

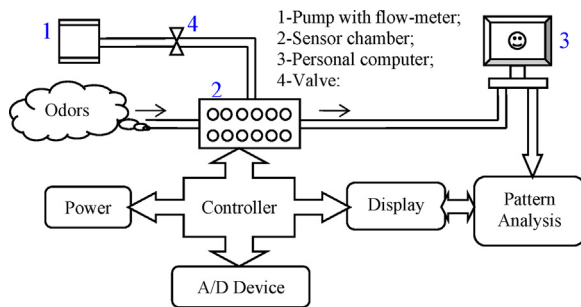


Fig. 1. General E-nose system.

out affecting the previous system. Some excellent reviews of the pattern classification and signal processing methods in machine olfaction and E-nose can be referred to as [18–20].

In this paper, we focus on the hot topic of long-term drift in E-nose. Drift is supposed to be some slow, continuous and uncertain effect, which is affecting the classification performance of an E-nose, due to the undesirable “unknown noise”. Although researchers have proposed many different methods for drift compensation from the angle of ensemble classifiers, semi-supervised learning, drift direction and drift correction, the results are still unsatisfactory since the restricts of methods [21–28]. In this paper, we propose a linear subspace alignment based drift adaptation technique, which tends to achieve drift compensation by removing the distribution discrepancy between the normal E-nose data and the drifted E-nose data. From the angle of machine learning, a basic assumption in learning theory is that the probability distribution between the training data and the testing data should be consistent. Therefore, a statistical learning algorithm can effectively capture the implied rules in the training data, which can also be adapted to the testing data. However, once the data (e.g. drift effect) violates the basic assumption of distribution consistency, the learning model may fail. This paper supposes that the drift effect seriously causes the distribution inconsistency of E-nose data. Therefore, our mission and idea is to improve the distribution consistency by subspace alignment between normal and drifted data. That is, we are not trying to calibrate the drift exactly or learn a drift-assimilation based classifier, but attempting to enhance the statistical probability distribution consistency of the normal and drifted data.

To this end, under the construction of principal component analysis, we propose a domain regularized component analysis (DRCA) method, which aims at improving the distribution consistency and achieving drift adaptation in principal component subspaces of normal (source domain) and drifted data (target domain). Inspired by transfer learning and domain adaptation [24], in this paper, we give that the normal data is from *source domain* and the drifted data is from *target domain*. The basic idea of the proposed DRCA method is illustrated vividly in Fig. 2.

The remainder of this paper is as follows. Section 2 illustrates the related work of drift compensation. Section 3 presents the proposed subspace projection based PCA synthesis approach with model formulation and optimization algorithm. The experiments and results have been discussed in Section 4. Finally, Section 5 concludes this paper.

2. Related work

2.1. Drift compensation in E-nose

Recently, many researchers have proposed different methods for handling drift issue of E-nose. Specifically, we present the existing work from classifier-level and feature level. In *classifier level*,

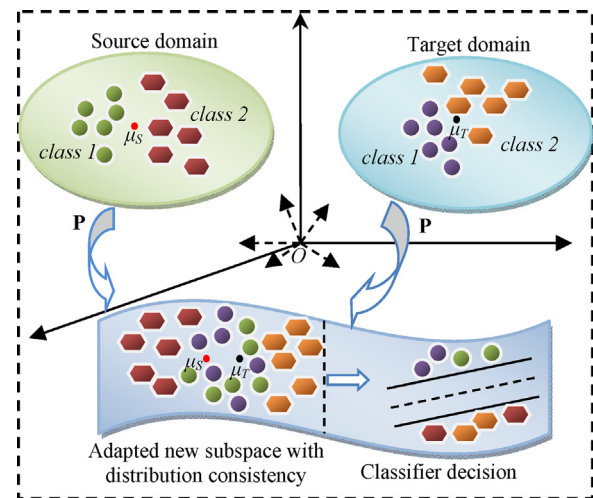


Fig. 2. Schematic diagram of the proposed DRCA method; after a subspace projection P , the source domain and target domain of different space distribution lie in a latent subspace with good distribution consistency (the centers of both domains become very close and drift is removed); in this latent subspace, the classification of two classes is successfully achieved.

Martinelli et al. proposed an artificial immune system based adaptive classifier for drift mitigation [24]. Vergara et al. proposed a classifier ensemble model for drift compensation, in which multiple SVMs with weighted ensemble are involved [22]. Also, a public long-term drift dataset was released and promotes the development in drift compensation. Liu and Tang [23] proposed a dynamic classifier ensemble for drift compensation in real application, in which the ensemble scheme is a dynamic weighted combination. Zhang et al. [24] proposed a domain adaptation extreme learning machine based transfer learning method for drift compensation, in which the classifier is learned with source and target domain simultaneously, by adaptive cross-domain learning. Liu et al. [25] also proposed a semi-supervised domain adaptation, in which manifold learning with affinity structure preservation is considered. These methods focus on the robust classifier learning, but neglect the important issue of data distribution mismatch. In *feature level*, Artursson et al. proposed a component correction based principal component analysis method, which aims at finding the drift direction and then remove the drift directly from the data [26]. The residual is recognized to be the non-drifted data. Similarly, Padilla et al. [27] proposed an orthogonal signal correction method for drift compensation, in which the components orthogonal to the signal are recognized to be drift, and then removed directly. These methods suppose that the drift is some additive noise on the data, which may violate the basic property of drift. Di Carlo et al. [28] proposed an evolutionary drift compensation, which aims at learning some transformation corresponding to the best pattern recognition accuracy. However, the learned transformation is solved by taking the recognition accuracy as optimization objective, regardless of drift essence, such that the transformation is recognition accuracy oriented and overfitted. Ziyatdinov et al. [35] proposed a common principal component analysis (CPCA) method for drift compensation, in which the drift variance can be computed. Also, the CPCA method can find common components for all gases in feature space. In this paper, the proposed DRCA method is completely different from the previous work. DRCA aims at finding a common subspace for both drifted and regular data, instead of the raw feature space. To our knowledge, it is the first time to address drift problem, by viewing drift to be some probability distribution bias in feature space between drift and regular data. In other words, we would like to find a latent common subspace where drift is not dominated.

2.2. Basic subspace projection algorithms

Subspace learning aims at learning a low-dimensional subspace under different optimization objectives. Several popular subspace methods include principal component analysis (PCA) [29], linear discriminant analysis (LDA) [30], manifold learning based locality preserving projections (LPP) [31], marginal fisher analysis (MFA) [32], and their kernelized and tensorized variants [33,34]. PCA, as an unsupervised method, aims at preserving the maximum variance of data, such that the redundant information is removed. LDA is a supervised dimension reduction method, which target at maximizing the inter-class scatter matrix and minimizing the intra-class scatter matrix, such the linear separability is maximized in the subspace. LPP is an unsupervised dimension reduction technique with manifold assumption, which preserves the affinity structure during dimension reduction process from high to low dimension. MFA is recognized to be a comprehensive version of LDA and LPP, which integrates the intra-class compactness of LDA and the graph embedding of LPP. Therefore, MFA is a supervised subspace learning method with maximum separability and affinity structure preservation.

These subspace learning methods assume that the whole distribution of the data is consistent. However, the drift issue we are facing with leads to much different distribution, so that these methods are no longer applicable. In this paper, a PCA synthesis technique, a subspace projection with transfer capability and distribution alignment, i.e. DRCA is proposed for handling the issue of distribution inconsistency, and further achieving drift adaptation in E-noses.

3. The proposed subspace projection: domain regularized component analysis (DRCA)

3.1. Notations

In this paper, the source and target domain are defined by subscript “S” and “T”, respectively. The training data of source and target domain is denoted as $\mathbf{X}_S = [\mathbf{x}_S^1, \dots, \mathbf{x}_S^{N_S}] \in \mathbb{R}^{D \times N_S}$ and $\mathbf{X}_T = [\mathbf{x}_T^1, \dots, \mathbf{x}_T^{N_T}] \in \mathbb{R}^{D \times N_T}$, respectively, where D is the number of dimensions, N_S and N_T are the number of training samples in both domains. Let $\mathbf{P} \in \mathbb{R}^{D \times d}$ represents the basis transformation that maps the original space of source and target data to some subspace with dimension of d . $\|\cdot\|_F$ and $\|\cdot\|_2$ denotes the Frobenius norm and l_2 -norm. $Tr(\cdot)$ denotes the trace operator and $(\cdot)^T$ denotes the transpose operator. Throughout this paper, matrix is written in capital bold face, vector is presented in lower bold face, and variable is in italics.

3.2. Problem formulation

As illustrated in Fig. 2, we aim to learn a basis transformation \mathbf{P} that maps the original space of source and target data to some subspace, where the feature distributions between the mapped source and target data $\mathbf{Y}_S = [\mathbf{y}_S^1, \dots, \mathbf{y}_S^{N_S}] \in \mathbb{R}^{d \times N_S}$ and $\mathbf{Y}_T = [\mathbf{y}_T^1, \dots, \mathbf{y}_T^{N_T}] \in \mathbb{R}^{d \times N_T}$ can be kept similar. Therefore, it is rational to have an idea that the mean distribution discrepancy (MDD) between \mathbf{Y}_S and \mathbf{Y}_T is minimized. Simply, the MDD concept is shown by the proposed domain distance, which is defined as the distance between the centers of the two domains. Therefore, the MDD minimization is formulated as

$$\min \|\boldsymbol{\mu}_S - \boldsymbol{\mu}_T\|_2^2 = \min \left\| \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{y}_S^i - \frac{1}{N_T} \sum_{j=1}^{N_T} \mathbf{y}_T^j \right\|_2^2 \quad (1)$$

where $\boldsymbol{\mu}_S = \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{y}_S^i$ and $\boldsymbol{\mu}_T = \frac{1}{N_T} \sum_{i=1}^{N_T} \mathbf{y}_T^i$ represents the centers in the new subspace.

According to the subspace projection, the new representation of source and target data in the lower-dimensional subspace can be formulated as

$$\mathbf{Y}_S = \mathbf{P}^T \mathbf{X}_S = \mathbf{P}^T [\mathbf{x}_S^1, \dots, \mathbf{x}_S^{N_S}] = [\mathbf{y}_S^1, \dots, \mathbf{y}_S^{N_S}] \quad (2)$$

$$\mathbf{Y}_T = \mathbf{P}^T \mathbf{X}_T = \mathbf{P}^T [\mathbf{x}_T^1, \dots, \mathbf{x}_T^{N_T}] = [\mathbf{y}_T^1, \dots, \mathbf{y}_T^{N_T}] \quad (3)$$

Therefore, we have $\mathbf{y}_S^i = \mathbf{P}^T \mathbf{x}_S^i$ and $\mathbf{y}_T^j = \mathbf{P}^T \mathbf{x}_T^j$. By substituting (2) and (3) into (1), the minimization problem (1) can be reformulated as

$$\min_{\mathbf{P}} \left\| \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{P}^T \mathbf{x}_S^i - \frac{1}{N_T} \sum_{j=1}^{N_T} \mathbf{P}^T \mathbf{x}_T^j \right\|_2^2 \quad (4)$$

For learning such a basis transformation \mathbf{P} that can minimize the mean distribution discrepancy in (4), we should also ensure that the projection does not distort the data itself, such that much more available information can be kept in the new subspace representation. Therefore, for source data it is rational to maximize the following term,

$$\max_{\mathbf{P}} Tr \left((\mathbf{P}^T \mathbf{X}_S) (\mathbf{P}^T \mathbf{X}_S)^T \right) = \max_{\mathbf{P}} Tr \left(\mathbf{P}^T \mathbf{X}_S \mathbf{X}_S^T \mathbf{P} \right) \quad (5)$$

It can be seen that by solving (5), the variance (energy) of the source data in the new subspace is maximized, such that the data cannot be distorted and the most available information in the raw data can be remained.

Similarly, for target domain data, we aim at maximizing the following term

$$\max_{\mathbf{P}} Tr \left((\mathbf{P}^T \mathbf{X}_T) (\mathbf{P}^T \mathbf{X}_T)^T \right) = \max_{\mathbf{P}} Tr \left(\mathbf{P}^T \mathbf{X}_T \mathbf{X}_T^T \mathbf{P} \right) \quad (6)$$

In actual application, there is little data in target domain by comparing to the source domain. In order to learn an effective linear subspace \mathbf{P} for drift adaptation, we propose a target domain regularized variance maximization idea and effectively avoid bias learning. Therefore, the variance maximization formulation shown in Eq. (5) and Eq. (6) can be integrated together by using a trade-off parameter as follows

$$\begin{aligned} & \max_{\mathbf{P}} Tr \left(\mathbf{P}^T \mathbf{X}_S \mathbf{X}_S^T \mathbf{P} \right) + \lambda \cdot Tr \left(\mathbf{P}^T \mathbf{X}_T \mathbf{X}_T^T \mathbf{P} \right) \\ &= \max_{\mathbf{P}} Tr \left(\mathbf{P}^T \mathbf{X}_S \mathbf{X}_S^T \mathbf{P} + \lambda \cdot \mathbf{P}^T \mathbf{X}_T \mathbf{X}_T^T \mathbf{P} \right) \\ &= \max_{\mathbf{P}} Tr \left(\mathbf{P}^T (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T) \mathbf{P} \right) \end{aligned} \quad (7)$$

where λ denotes the trade-off (regularization) parameter.

The proposed target domain regularized component analysis (DRCA) model aims at minimizing the mean distribution discrepancy (MDD) in (4) and maximizing the regularized variance in (7) of source and target data, simultaneously. Therefore, by integrating (4) and (7) together, the proposed DRSA model can be formulated as

$$\max_{\mathbf{P}} \frac{Tr \left(\mathbf{P}^T (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T) \mathbf{P} \right)}{\left\| \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{P}^T \mathbf{x}_S^i - \frac{1}{N_T} \sum_{j=1}^{N_T} \mathbf{P}^T \mathbf{x}_T^j \right\|_2^2} \quad (8)$$

Let $\mu_S = \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{x}_S^i$ and $\mu_T = \frac{1}{N_T} \sum_{i=1}^{N_T} \mathbf{x}_T^i$ be the centers of source and target domain, then the maximization problem in Eq. (8) can be finally written as

$$\begin{aligned} & \max_{\mathbf{P}} \frac{\text{Tr}(\mathbf{P}^T (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T) \mathbf{P})}{\|\mathbf{P}^T \left(\frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{x}_S^i \right) - \mathbf{P}^T \left(\frac{1}{N_T} \sum_{j=1}^{N_T} \mathbf{x}_T^j \right)\|_2^2} \\ & = \max_{\mathbf{P}} \frac{\text{Tr}(\mathbf{P}^T (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T) \mathbf{P})}{\|\mathbf{P}^T \mu_S - \mathbf{P}^T \mu_T\|_2^2} \\ & = \max_{\mathbf{P}} \frac{\text{Tr}(\mathbf{P}^T (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T) \mathbf{P})}{\text{Tr}(\mathbf{P}^T \mu_S - \mathbf{P}^T \mu_T) (\mathbf{P}^T \mu_S - \mathbf{P}^T \mu_T)^T} \\ & = \max_{\mathbf{P}} \frac{\text{Tr}(\mathbf{P}^T (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T) \mathbf{P})}{\text{Tr}(\mathbf{P}^T (\mu_S - \mu_T)(\mu_S - \mu_T)^T \mathbf{P})} \end{aligned} \quad (9)$$

$$\frac{\partial L(\mathbf{P}, \rho)}{\partial \mathbf{P}} = 0 \rightarrow \quad (12)$$

$$((\mu_S - \mu_T)(\mu_S - \mu_T)^T)^{-1} (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T) \mathbf{P} = \rho \mathbf{P}$$

From (12), we can observe that the \mathbf{P} can be obtained by solving the following Eigenvalue decomposition problem of matrix \mathbf{A} ,

$$\mathbf{A} \mathbf{P} = \rho \mathbf{P} \quad (13)$$

where $\mathbf{A} = ((\mu_S - \mu_T)(\mu_S - \mu_T)^T)^{-1} (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T)$, and ρ denotes the diagonal matrix of eigenvalues.

From (13), it is clear that \mathbf{P} denotes the eigen-vectors. Due to that the model (10) is a maximization problem, therefore, the optimal subspace \mathbf{P}^* denotes the eigenvectors with respect to the first d largest eigenvalues $[\rho_1, \dots, \rho_d]$, represented by

$$\mathbf{P}^* = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_d] \quad (14)$$

For easy implementation, the proposed DRSA algorithm is summarized in Algorithm 1.

Algorithm 1. The proposed DRCA.

Input: Heterogeneous data including source data $\mathbf{X}_S \in \mathbb{R}^{D \times N_S}$, and target data $\mathbf{X}_T \in \mathbb{R}^{D \times N_T}$, λ , d ;
Procedure:
1. Compute the centroid of source data $\mu_S = \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{x}_S^i$;
2. Compute the centroid of target data $\mu_T = \frac{1}{N_T} \sum_{i=1}^{N_T} \mathbf{x}_T^i$;
3. Compute the matrix \mathbf{A} as
$\mathbf{A} = ((\mu_S - \mu_T)(\mu_S - \mu_T)^T)^{-1} (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T)$
4. Perform the eigenvalue decomposition of \mathbf{A} by using (13);
5. Compute the optimal subspace $\mathbf{P}^* = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_d]$ by (14);
6. New subspace projection with drift adaptation by
$\mathbf{x}'_S = (\mathbf{P}^*)^T \mathbf{x}_S$
$\mathbf{x}'_T = (\mathbf{P}^*)^T \mathbf{x}_T$
Output: The basis transformation \mathbf{P} , the drift adapted (distribution aligned) data \mathbf{X}'_S and \mathbf{X}'_T .

3.3. Model optimization

In the maximization problem Eq. (9), there are many possible solutions of \mathbf{P} (i.e. non-unique solutions). To guarantee the unique property of solution, we impose an equality constraint on the optimization problem, and then Eq. (9) can be written as

$$\begin{aligned} & \max_{\mathbf{P}} \text{Tr}(\mathbf{P}^T (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T) \mathbf{P}) \\ & \text{s.t. } \text{Tr}(\mathbf{P}^T (\mu_S - \mu_T)(\mu_S - \mu_T)^T \mathbf{P}) = \eta \end{aligned} \quad (10)$$

where η is a positive constant value.

To solve (10), the Lagrange multiplier function is written as

$$\begin{aligned} L(\mathbf{P}, \rho) &= \text{Tr}(\mathbf{P}^T (\mathbf{X}_S \mathbf{X}_S^T + \lambda \cdot \mathbf{X}_T \mathbf{X}_T^T) \mathbf{P}) \\ &\quad - \rho (\text{Tr}(\mathbf{P}^T (\mu_S - \mu_T)(\mu_S - \mu_T)^T \mathbf{P}) - \eta) \end{aligned} \quad (11)$$

where ρ denotes the Lagrange multiplier coefficient.

By setting the partial derivation of $L(\mathbf{P}, \rho)$ with respect to \mathbf{P} to be 0, we have

3.4. Remarks

The proposed DRCA method is an unsupervised algorithm, in which the labels of the source and target data are not involved. In optimization of the subspace projection \mathbf{P} , DRCA holds similar computation with the popular principal component analysis (PCA) and linear discriminant analysis (LDA) method by using Eigenvalue decomposition, and the eigenvectors with respect to the first d largest eigenvalues are selected as a group of projection basis. Therefore, the computational complexity of DRCA is $O(D^3)$. Extensively, the proposed DRCA can also be further improved by considering the label information in source data, as LDA does. Therefore, a supervised or semi-supervised version can also be easily derived in the future work, if the full or partial label (class) information of the source data is available. Then, a discriminative subspace can be learned by maximizing the between-class scatter matrix and minimizing the within-class scatter matrix. In this paper, we focus on the unsupervised algorithm for distribution alignment. Additionally, the parameters such as dimensionality d and the trade-off coefficient λ are tuned in the search space $d = \{2^k, k=0,1,2,3,4,5,6,7\}$ and $\lambda = \{10^k, k=-3,-2,-1,0,1,2,3,4\}$.

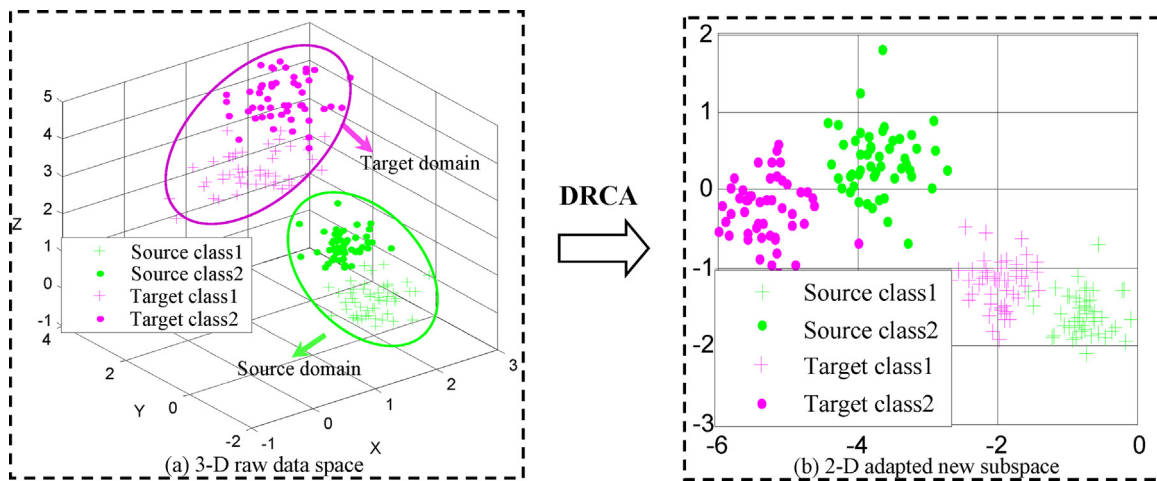


Fig. 3. Subspace adaptation of synthetic data by using the proposed DRCA method.

4. Experiments

In this section, we conduct three experiments on different datasets with knowledge bias/shift and drift for testing the effectiveness of the proposed DRCA method. Specifically, we first generate a synthetic dataset of Gaussian distribution with different parameters (e.g. mean vector and covariance matrix). Then, we test the proposed method on a public benchmark sensor drift dataset from an electronic nose. Further, we validate the proposed method on our own dataset with long-term drift by an E-nose.

4.1. Experiment on synthetic data

To validate the effectiveness of the proposed DRCA method, we randomly generate source domain data and target domain data with Gaussian distribution with respect to different mean vector and covariance matrix, respectively. For each domain, 100 data points of two classes (50 data points per class) with three dimensions are generated. In this way, the source domain and target domain are with different space distribution, such that the proposed method can be used to handle such issues. The generated synthetic data points for source and target domain are shown in Fig. 3(a) in 3D-coordinate space, from which we observe that the source data and target data are with different space distribution and the classifier of source data cannot be directly used for target data due to the distribution inconsistency.

Therefore, improving the distribution consistency between source domain and target domain is really required, such that one unified classifier can be achieved for both source data and target data. The proposed DRCA is exactly for this challenging task, that is, subspace adaptation. Visually, the adapted new subspace with improved distribution consistency by using the proposed DRCA is shown in Fig. 3(b). We observe from the 2-D coordinate subspace that the data points of the same class from source and target domain lie in nearer clusters, and one unified classifier can be easily derived for multi-classification regardless of source or target data. In other words, the distribution inconsistency between source and target domain is removed by DRCA subspace adaptation.

4.2. Experiment on benchmark sensor drift data

For verification of the proposed DRCA method, the real sensor drift benchmark dataset of three years collected by an E-nose from Vergara et al. in UCSD [22] is used in our experiment. The dataset was gathered during the period of 36 months from January 2008 to February 2011 based on a gas delivery platform. The E-nose is with

16 gas sensors, and exposed to six kinds of pure gaseous substances, such as acetone, acetaldehyde, ethanol, ethylene, ammonia, and toluene at different concentration levels. Totally, this dataset contains 13910 measurements (samples), which are divided into 10 batches of time series. The number of samples for each class with respect to each batch is summarized in Table 1. In feature extraction, 8 features were extracted for each sensor, and consequently a 128-dimensional feature vector (16×8) is formulated for each sample. We refer interested readers to as [22] for more technical details on feature extraction.

To reduce the scale variants among dimensions, the whitening processing (centralization) is conducted on the data. For visually observe the drift of the heterogeneous E-nose data, we have plotted the PCA scatter points in Fig. 4, in which the principal component projection (coefficients) is calculated on the raw clean data of batch 1, and the obtained PCA coefficients are used to compute the projected subspaces of other drifted batches. From Fig. 4, it is clear that the low-dimensional subspace distribution between batch 1 and other batches is significantly biased (inconsistent) due to drift impact over time.

Therefore, we can say that the drifted E-nose data is heterogeneous. Due to the distribution inconsistency between batch 1 (training data) and other batches (testing data), the recognition performance of the trained pattern classifier would be degraded, because it violates the basic assumption of machine learning that the training data and testing data should be with the same or similar probability distribution (i.e. independent identical distribution, i.i.d.).

To demonstrate the effectiveness of the proposed DRCA method for subspace adaptation and distribution alignment, we present the qualitative and quantitative experiments, respectively.

4.3. Qualitative result

Motivated by the PCA scatter points in Fig. 4, we provide the domain regularized principal components (PC1 vs. PC2) by using the proposed DRCA method on Fig. 4. Totally, 9 tasks with pairwise batches (batch 1 vs. batch i) ($i=2, \dots, 10$) are implemented, respectively. For each task, the projection coefficients is nominated as $\mathbf{P}_{1,i}$ ($i=2, \dots, 10$), then the projected subspaces of batch 1 and batch i based on $\mathbf{P}_{1,i}$ are plotted. As a result, 9 pairwise principal component scatter points plots highlighted by using dot rectangles are shown in Fig. 5. From the low-dimensional subspace by using the proposed DRCA, the distribution consistency between the drifted batch i and the clean batch 1 is improved.

Table 1
Benchmark Sensor Drift Dataset From UCSD.

Batch ID	Month	Acetone	Acetaldehyde	Ethanol	Ethylene	Ammonia	Toluene	Total
Batch 1	1,2	90	98	83	30	70	74	445
Batch 2	3~10	164	334	100	109	532	5	1244
Batch 3	11,12,13	365	490	216	240	275	0	1586
Batch 4	14,15	64	43	12	30	12	0	161
Batch 5	16	28	40	20	46	63	0	197
Batch 6	17,18,19,20	514	574	110	29	606	467	2300
Batch 7	21	649	662	360	744	630	568	3613
Batch 8	22,23	30	30	40	33	143	18	294
Batch 9	24,30	61	55	100	75	78	101	470
Batch 10	36	600	600	600	600	600	600	3600

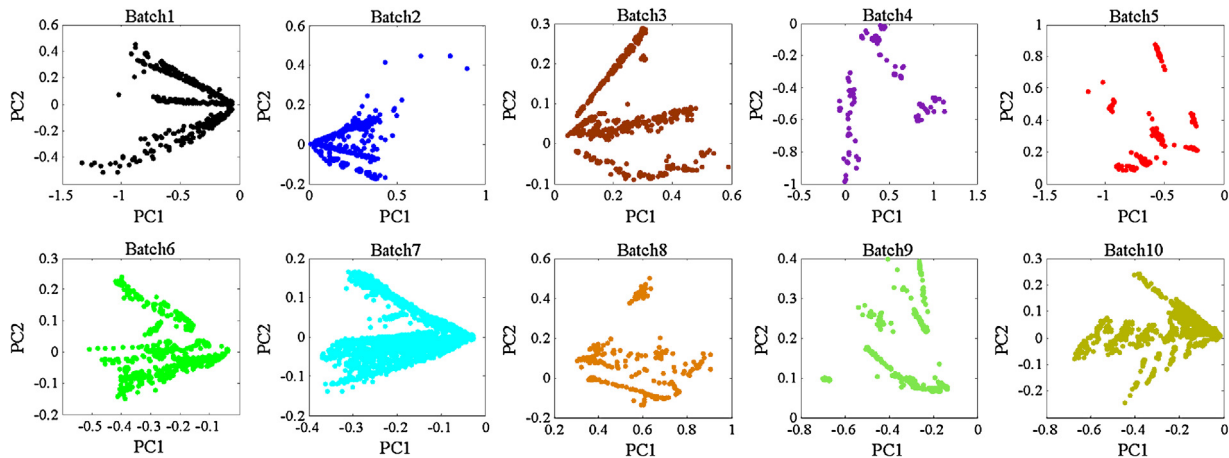


Fig. 4. Principal component subspace (PC1 vs. PC2) of the raw drifted data of 10 batches by using PCA.

It is worth noting that it may be impossible to directly calibrate the drifted sensor response due to the nonlinear dynamic behavior of sensor drift and complex sensing mechanism. Therefore, it is reasonable and necessary to handle drift compensation from the perspective of probability distribution alignment and subspace adaptation.

4.4. Quantitative result

The essential task for the proposed method is to improve the classification performance. Therefore, the classification accuracy of 6 classes on each batch is reported as evaluation metric. In comparisons, two experimental settings by following [22] are given as follows.

1) Setting 1: Take batch 1 as source domain for model training, and test on batch K , $K=2, \dots, 10$ (target domains). The classification accuracy on batch K is reported.

2) Setting 2: Take batch $K-1$ as source domain for model training, and test on batch K , $K=2, \dots, 10$ (target domains). The classification accuracy on batch K is reported.

For classification, a multi-class SVM with RBF kernel (SVM-rbf) is used as classifier. In comparisons, we compare the proposed DRCA with two baseline subspace methods such as PCA and LDA trained on source data, and eight state-of-the-art results on this benchmark dataset such as CCPCA, SVM ensemble classifier, SVM-gfk, SVM-comgfk, ML-gfk, ML-comgfk, ELM-rbf, and OSC, and two representative calibration transfer methods in E-nose, such as generalized least squares weighting (GLSW) [36] and direct standardization (DS) [37], are compared.

We conducted the experiments on *Setting 1* and *Setting 2*, respectively. The recognition results for different methods under *Experimental Setting 1* are reported in Table 2, from which we observe that the proposed DRCA achieves the best classification

performance. The average classification accuracy is 77.63%, which is 10% higher than the second best learning method, i.e. ML-comgfk. Moreover, for each batch, we also give the best parameters that the proposed method achieves the highest accuracy in Table 3.

Further, by following the experimental *Setting 2*, that is, model training on batch $K-1$ and test on batch K , and the results are reported in Table 4, and the corresponding parameters are reported in Table 5. From Table 4, we can see that the proposed DRCA performs the second best performance (74.2% in average). From the quantitative classification accuracy, the effectiveness and competitiveness of the proposed DRCA model have been clearly demonstrated. Note that a detail parameter sensitivity analysis is provided in Section 4.4.

4.5. Experiment on drift plus shift E-nose data

In this section, a more complex E-nose dataset from our lab is used. This dataset includes three parts: master data (collected 5 years ago), slave 1 data (collected now) and slave 2 data (collected now). The “complex” is represented by multiple E-nose instruments of the same type. More specifically, not only the sensor drift is implied in the data, the sensor shift caused by bad reproducibility also happens. In data acquisition, the master and two slavery E-nose systems were developed in [4]. Four TGS series sensors and two variables for ambient temperature and humidity are integrated. Therefore the dimension of features (DoF) is 6. In this dataset, six kinds of gaseous contaminants such as formaldehyde, benzene, toluene, carbon monoxide, nitrogen dioxide and ammonia are included. The detailed description of the dataset is shown in Table 6.

The PCA scatter points on the master data, slave 1 data and slave 2 data, are shown in Fig. 6, respectively, in which we can see that the points from different classes are crossed. By implementing the

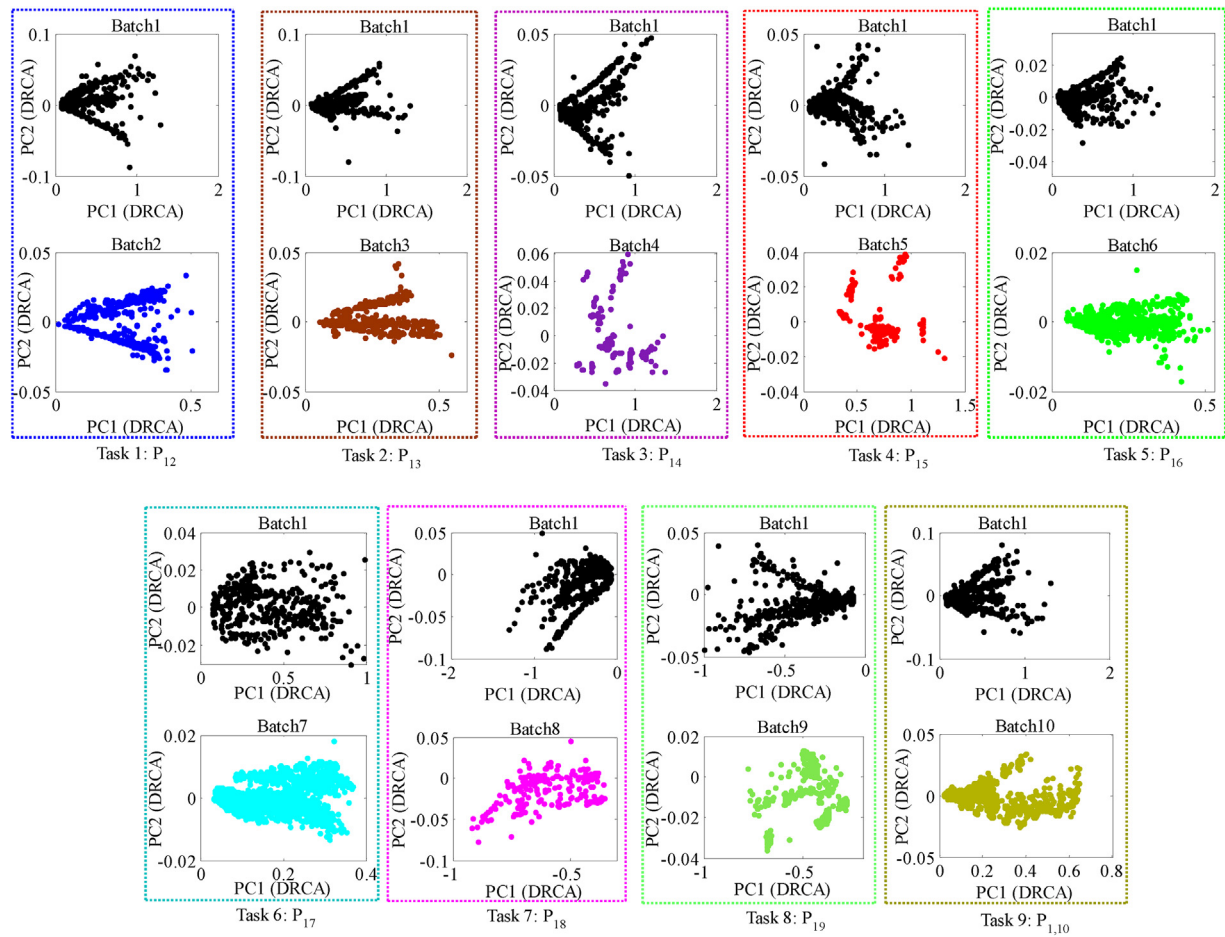


Fig. 5. Principal components projected subspace (PC1 vs. PC2) with drift adaptation by using the proposed DRCA method. Specifically, 9 tasks are performed. For each task, a new subspace projection $P_{1,i}$ ($i=2, \dots, 10$) is obtained by implementing DRCA on batch 1 and batch i ($i=2, \dots, 10$).

Table 2

Recognition Accuracy (%) Under Experimental Setting 1.

Batch ID	Batch 2	Batch 3	Batch 4	Batch 5	Batch 6	Batch 7	Batch 8	Batch 9	Batch 10	Average
PCA _{SVM}	82.40	84.80	80.12	75.13	73.57	56.16	48.64	67.45	49.14	68.60
LDA _{SVM}	47.27	57.76	50.93	62.44	41.48	37.42	68.37	52.34	31.17	49.91
CC-PCA	67.00	48.50	41.00	35.50	55.00	31.00	56.50	46.50	30.50	45.72
SVM-rbf	74.36	61.03	50.93	18.27	28.26	28.81	20.07	34.26	34.47	38.94
SVM-gfk	72.75	70.08	60.75	75.08	73.82	54.53	55.44	69.62	41.78	63.76
SVM-comgfk	74.47	70.15	59.78	75.09	73.99	54.59	55.88	70.23	41.85	64.00
ML-rbf	42.25	73.69	75.53	66.75	77.51	54.43	33.50	23.57	34.92	53.57
ML-comgfk	80.25	74.99	78.79	67.41	77.82	71.68	49.96	50.79	53.79	67.28
ELM-rbf	70.63	66.44	66.83	63.45	69.73	51.23	49.76	49.83	33.50	57.93
OSC	88.10	66.71	54.66	53.81	65.13	63.71	36.05	40.21	40.08	56.50
GLSW	78.38	69.36	80.75	74.62	69.43	44.28	48.64	67.87	46.58	64.43
DS	69.37	46.28	41.61	58.88	48.83	32.83	23.47	72.55	29.03	46.98
DRCA	89.15	92.69	87.58	95.94	86.52	60.25	62.24	72.34	52.00	77.63

Table 3

Parameters' values of the DRCA Under Experimental Setting 1.

Batch ID	Batch 2	Batch 3	Batch 4	Batch 5	Batch 6	Batch 7	Batch 8	Batch 9	Batch 10
λ	0.008	0.9	1000	0.03	0.06	8	0.01	0.001	0.08
d	30	40	39	12	12	17	8	11	10

proposed DRCA method on master \rightarrow slave1 and master \rightarrow slave2 respectively, the results are shown in Fig. 7, from which we can observe that the cross problem of different classes is eased. This qualitatively demonstrates the positive drift adaptation effect of the DRCA.

4.6. Settings

The *master* data collected 5 years ago is used as source domain data (no drift). The slave 1 and slave 2 data collected after 5 years are used as target domain data (with drift and shift). Then, we conduct the experiments with two settings as follows.

Table 4
Recognition Accuracy (%) Under Experimental **Setting 2**.

Batch ID	1 → 2	2 → 3	3 → 4	4 → 5	5 → 6	6 → 7	7 → 8	8 → 9	9 → 10	Average
PCA _{SVM}	82.40	98.87	83.23	72.59	36.70	74.98	58.16	84.04	30.61	69.06
LDA _{SVM}	47.27	46.72	70.81	85.28	48.87	75.15	77.21	62.77	30.25	60.48
SVM-rbf	74.36	87.83	90.06	56.35	42.52	83.53	91.84	62.98	22.64	68.01
SVM-gfk	72.75	74.02	77.83	63.91	70.31	77.59	78.57	86.23	15.76	68.56
SVM-comgfk	74.47	73.75	78.51	64.26	69.97	77.69	82.69	85.53	17.76	69.40
ML-rbf	42.25	58.51	75.78	29.10	53.22	69.17	55.10	37.94	12.44	48.17
ML-comgfk	80.25	98.55	84.89	89.85	75.53	91.17	61.22	95.53	39.56	79.62
ELM-rbf	70.63	40.44	64.16	64.37	72.70	80.75	88.20	67.00	22.00	63.36
GLSW	78.38	97.04	81.99	73.60	36.57	74.48	60.54	81.91	26.31	67.87
DS	69.37	53.59	67.08	37.56	36.30	26.57	49.66	42.55	25.78	45.38
DRCA	89.15	98.11	95.03	69.54	50.87	78.94	65.99	84.04	36.31	74.22

Table 5
Parameters' values of the DRCA Under Experimental **Setting 2**.

Batch ID	1 → 2	2 → 3	3 → 4	4 → 5	5 → 6	6 → 7	7 → 8	8 → 9	9 → 10
λ	0.008	1	0.1	0.001	0.01	10000	0.1	0.09	0.1
d	30	84	61	8	110	109	15	88	55

Table 6
Data Description of The Complex E-nose Data.

E-nose System	DoF	Formaldehyde	Benzene	Toluene	Carbon monoxide	Nitrogen dioxide	Ammonia	Total
Master (no drift)	6	126	72	66	58	38	60	420
Slave 1 (drift + shift)	6	108	108	106	98	107	81	608
Slave 2 (drift + shift)	6	108	87	94	95	108	84	576

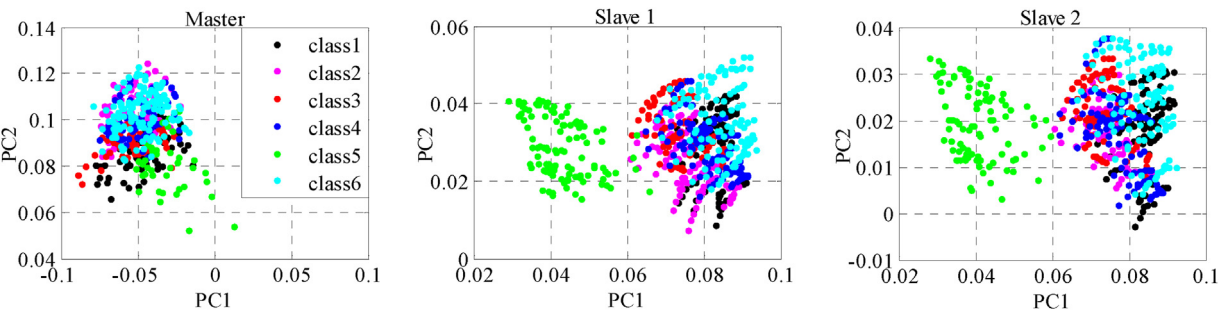


Fig. 6. The PCA scatter points of the master, slave 1 and slave 2 data, respectively.

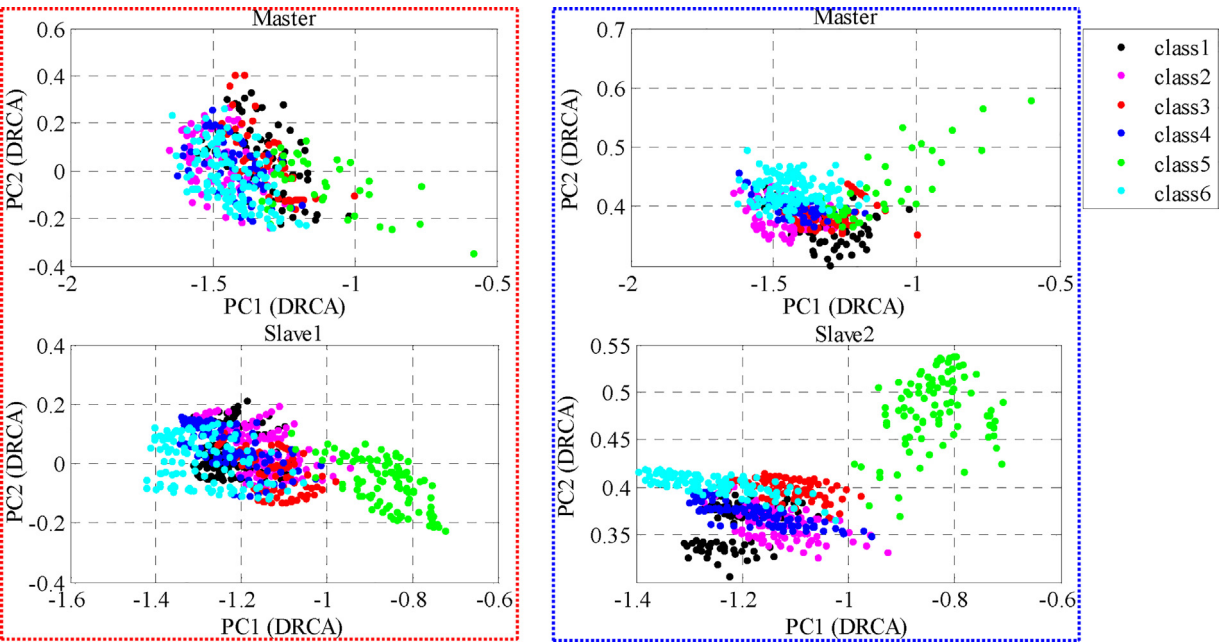


Fig. 7. The scatter points by using DRCA from master → slave1 (red dot) and master → slave2 (blue dot). Note that the analytes from class 1 to class 6 are toluene, benzene, ammonia, carbon monoxide, nitrogen dioxide, and formaldehyde, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

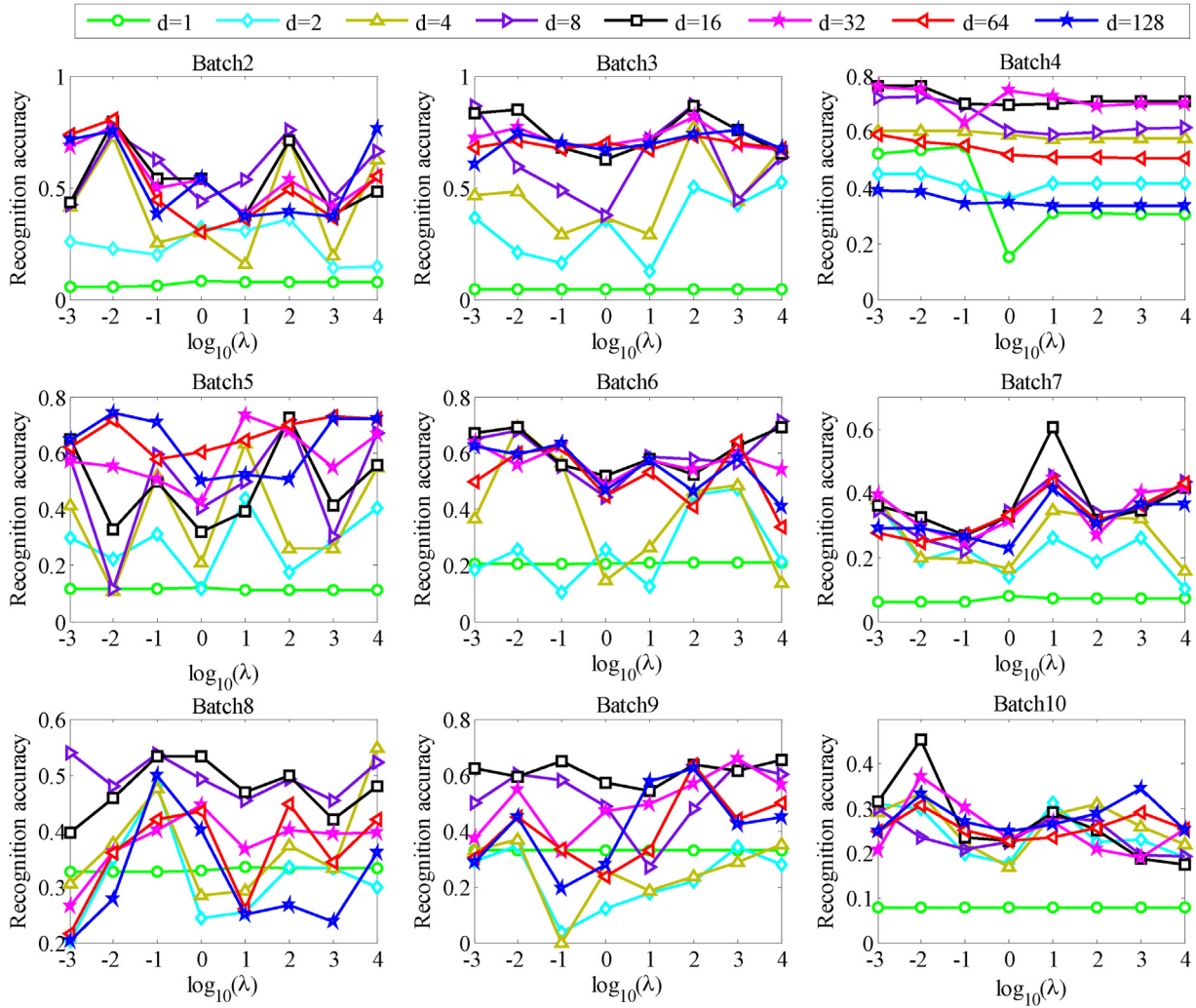


Fig. 8. The performance curves under different dimension d and λ .

Table 7
Recognition Accuracy (%) With Sensor Shift Calibration (Setting 1).

Task	SVM	PCA	LDA	GLSW	DS	DRCA
master \rightarrow slave1	45.89	46.22	42.11	41.45	40.30	57.07
master \rightarrow slave2	31.08	41.84	41.32	48.09	39.76	52.95

Setting 1 (Drift): Due to that there is also sensor discreteness between slaves and master, the sensor calibration between slaves and master is made by using linear regression according to [38]. Therefore, only drift exists between the source and target data.

Setting 2 (Drift + Shift): The sensor calibration step is omitted, which implies that both the sensor drift and shift exist between the source and target data.

4.7. Results

In this section, the average recognition accuracy of the six kinds of gases is reported. For each setting, two tasks including master \rightarrow slave1 and master \rightarrow slave2 are conducted. By following setting 1 (with shift calibration), the experimental results of the proposed DRCA method are presented in Table 7, in which the SVM, PCA, LDA and two representative calibration transfer methods in E-nose, such as generalized least squares weighting (GLSW) [35] and direct standardization (DS) [36] are compared. From the results, we can observe that the proposed method achieves the best

Table 8
Recognition Accuracy (%) Without Sensor Shift Calibration (Setting 2).

Task	SVM	PCA	LDA	GLSW	DS	DRCA
master \rightarrow slave1	51.97	51.97	51.97	47.53	40.46	58.55
master \rightarrow slave2	60.59	60.59	56.77	59.38	40.63	61.63

recognition accuracies on two slaves, with 10% improvements. This demonstrates that the proposed DRCA can effectively address the long-term drift problem (5 years).

Further, to verify the proposed method for drift plus shift adaptation, the experimental results by following setting 2 are presented in Table 8. We can observe that the proposed DRCA still outperforms other methods. Additionally, by comparing the results between Table 7 and 8, we get that the proposed method can well issue the sensor shift. Therefore, we confirm that the proposed method is effective in handling heterogeneous E-nose data caused by both drift and shift.

4.8. Parameter sensitivity analysis

In the proposed DRCA model, there are two parameters: the dimension d and regularization coefficient λ . To observe the performance variations in tuning the two parameters, we tune from the parameter set $d=\{2^k, k=0,1,2,3,4,5,6,7\}$ and $\lambda=\{10^k, k=-3,-2,-1,0,1,2,3,4\}$. We select the benchmark sensor drift dataset for

experiments. When tune the parameter λ , the value of d is fixed, and vice versa. The tuning results are shown in Fig. 8, which can help us quickly determine the range of the optimal parameters.

4.9. Discussion

The proposed DRCA method, from the viewpoint of machine learning, is a cross-domain subspace learning technique. The so-called “cross-domain” denotes the cross-system from master to slave in E-nose topic. The advantage of DRCA is that a common subspace can be learned for source domain (master) and target domain (slave), such that the drift information implied in target domain is eliminated in this common subspace. That is, in this common subspace, the master and slave share similar class information, which greatly helps the classification without drift blur. There is also a limitation that the proposed DRCA is essentially a linear technique, by transformation via a projection \mathbf{P} , which may not seriously guarantee the “drift-less” property. Therefore, our future improvement may be the nonlinear subspace learning by considering the kernel technique. Additionally, the proposed DRCA can be easily implemented in on-line manner. Specifically, the focus of DRCA is the learning of \mathbf{P} based on source (master) and target (slave) dataset. In on-line application, the target dataset may be time-varying, therefore, the time interval may be defined by users, for updating the projection \mathbf{P} .

5. Conclusion

In this paper, we propose a domain regularized component analysis (DRCA) model for heterogeneous E-nose application. Drift is often publicly recognized as an ill-posed problem due to the uncertainty and unpredictability of its degeneration behavior. Therefore, machine learning may be a powerful solution for drift adaptation. The proposed DRCA method is motivated by transfer learning, and the difference of probability distribution between source data and target data incurs the failure of machine learning in data mining. For learning a common subspace for heterogeneous data, an intuitive idea by regularizing the target domain and minimizing the mean distribution discrepancy (MDD) based domain distance is proposed. Motivated by PCA and the proposed idea, a DRCA model is formulated. The optimization algorithm of the proposed DRCA is also induced. Experiments on synthetic dataset, benchmark sensor drift dataset and sensor drift plus shift dataset demonstrate that the proposed DRCA method outperforms several state-of-the arts methods.

Acknowledgments

The authors would like to thank the anonymous reviewers and the editor for their valuable suggestions. This work was supported in part by Fundamental Research Funds for the Central Universities (No. 106112017CDJQJ168819), National Natural Science Foundation of China (Grant 61401048, 61471073), and in part by the research fund for Central Universities and young Scientist Foundation of Chongqing (cstc2013kjrc-qncr0080). The authors would also like to thank Dr. A. Vergara and his team from UCSD for their provided long-term artificial olfactory data.

References

- [1] L. Zhang, F. Tian, S. Liu, L. Dang, X. Peng, X. Yin, Chaotic time series prediction of e-nose sensor drift in embedded phase space, *Sens. Actuators, B* 182 (2013) 71–79.
- [2] F. Hossein-Babaei, V. Ghafarinia, Compensation for the drift-like terms caused by environmental fluctuations in the response of chemoresistive gas sensors, *Sens. Actuators, B* 143 (2010) 641–648.
- [3] N. Bhattacharyya, R. Bandyopadhyay, M. Bhuyan, B. Tudu, D. Ghosh, A. Jana, Electronic nose for black tea classification and correlation of measurements with tea taster marks, *IEEE Trans. Instrum. Meas.* 57 (7) (2008) 1313–1321.
- [4] L. Zhang, F. Tian, Performance study of multilayer perceptrons in a low-Cost electronic nose, *IEEE Trans. Instrum. Meas.* 63 (7) (2014) 1670–1679.
- [5] R. Gosangi, R. Gutierrez-Osuna, Active temperature programming for metal-Oxide chemoresistors, *IEEE Sens. J.* 10 (6) (2010) 1075–1082.
- [6] L. Zhang, F.C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, *Anal. Chim. Acta* 816 (2014) 8–17.
- [7] X. Yin, L. Zhang, F. Tian, D. Zhang, Temperature modulated gas sensing E-nose system for low-cost and fast detection, *IEEE Sens. J.* 16 (2) (2016) 464–474.
- [8] A. Flammini, D. Marioli, A. Taroni, A low-Cost interface to high-Value resistive sensors varying over a wide range, *IEEE Trans. Instrum. Meas.* 53 (4) (2004) 1052–1056.
- [9] A. Fort, N. Machetti, S. Rocchi, M.B.S. Santos, L. Tondi, N. Olivieri, V. Vignoli, G. Sberveglieri, Tin oxide gas sensing: comparison among different measurement techniques for gas mixture classification, *IEEE Trans. Instrum. Meas.* 52 (3) (2003) 921–926.
- [10] K. Brudzewski, S. Osowski, A. Dwulit, Recognition of coffee using differential electronic nose, *IEEE Trans. Instrum. Meas.* 61 (6) (2012) 1803–1810.
- [11] F. Herrero-Carrón, D.J. Yáñez, F. de, Borja Rodríguez, P. Varona, An active, inverse temperature modulation strategy for single sensor odorant classification, *Sens. Actuators, B* 206 (2015) 555–563.
- [12] F. Hossein-Babaei, A. Amini, Recognition of complex odors with a single generic tin oxide gas sensor, *Sens. Actuators, B* 194 (2014) 156–163.
- [13] M. Palit, B. Tudu, P.K. Dutta, A. Dutta, A. Jana, J. Kumar Roy, N. Bhattacharyya, R. Bandyopadhyay, A. Chatterjee, Classification of black tea taste and correlation with tea taster's mark using voltammetric electronic tongue, *IEEE Trans. Instrum. Meas.* 59 (8) (2010) 2230–2239.
- [14] I. Rodríguez-Lujan, J. Fonollosa, A. Vergara, M. Homer, R. Huerta, On the calibration of sensor arrays for pattern recognition using the minimal number of experiments, *Chemom. Intell. Lab. Syst.* 130 (2014) 123–134.
- [15] S.K. Jha, K. Hayashi, R.D.S. Yadava, Neural, fuzzy and neuro-fuzzy approach for concentration estimation of volatile organic compounds by surface acoustic wave sensor array, *Measurement* 55 (2014) 186–195.
- [16] S.J. Dixon, R.G. Brereton, Comparison of performance of five common classifiers represented as boundary methods: euclidean Distance to Centroids, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Learning Vector Quantization and Support Vector Machines, as dependent on data structure, *Chemom. Intell. Lab. Syst.* 95 (2009) 1–17.
- [17] B. Tudu, A. Metla, B. Das, N. Bhattacharyya, A. Jana, D. Ghosh, R. Bandyopadhyay, Towards versatile electronic nose pattern classifier for black tea quality evaluation: an incremental fuzzy approach, *IEEE Trans. Instrum. Meas.* 58 (9) (2009) 3069–3078.
- [18] S. Marco, A. Gutiérrez-Gálvez, Signal and data processing for machine olfaction and chemical sensing: a review, *IEEE Sens. J.* 12 (11) (2012) 3189–3214.
- [19] F. Röck, N. Barsan, U. Weimar, Electronic nose: current status and future trends, *Chem. Rev.* 108 (2008) 705–725.
- [20] S.M. Scott, D. James, Z. Ali, Data analysis for electronic nose systems, *Microchim. Acta* 156 (2007) 183–207.
- [21] E. Martinelli, G. Magna, S. De Vito, R. Di Fuccio, G. Di Francia, A. Vergara, C. Di Natale, An adaptive classification model based on the Artificial Immune System for chemical sensor drift mitigation, *Sens. Actuators, B: Chem.* 177 (2013) 1017–1026.
- [22] A. Vergara, S. Vembu, T. Ayhan, M.A. Ryan, M.L. Homer, R. Huerta, Chemical gas sensor drift compensation using classifier ensembles, *Sens. Actuators, B: Chem.* 167 (2012) 320–329.
- [23] H. Liu, Z. Tang, Metal oxide gas sensor drift compensation using a dynamic classifier ensemble based on fitting, *Sensors* 13 (2013) 9160–9173.
- [24] L. Zhang, D. Zhang, Domain adaptation extreme learning machines for drift compensation in E-nose systems, *IEEE Trans. Instrum. Meas.* 64 (7) (2015) 1790–1801.
- [25] Q. Liu, M. Ye, S.S. Ge, X. Du, Drift compensation for electronic nose by semi-Supervised domain adaption, *IEEE Sens. J.* 14 (3) (2014) 657–665.
- [26] T. Artursson, T. Eklov, I. Lundstrom, P. Martensson, M. Sjöström, M. Holmberg, Drift correction for gas sensors using multivariate methods, *J. Chemom.* 14 (5–6) (2000) 711–723.
- [27] M. Padilla, A. Perera, I. Montoliu, A. Chaudry, K. Persaud, S. Marco, Drift compensation of gas sensor array data by Orthogonal Signal Correction, *Chemom. Intell. Lab. Syst.* 100 (2010) 28–35.
- [28] S. Di Carlo, M. Falasconi, E. Sanchez, A. Scionti, G. Squillero, A. Tonda, Increasing pattern recognition accuracy for chemical sensing by evolutionary based drift compensation, *Pattern Recognit. Lett.* 32 (2011) 1594–1603.
- [29] M. Turk, A. Pentland, Eigenfaces for recognition, *J. Cogn. Neurosci.* 3 (1) (1991) 71–86.
- [30] P.N. Belhumeur, J. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, *IEEE Trans. Pattern Anal. Mach. Intell.* 19 (7) (1997) 711–720.
- [31] X. He, S. Yan, Y. Hu, P. Niyogi, H.J. Zhang, Face recognition using laplacianfaces, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (3) (2005) 328–340.
- [32] S. Yan, D. Xu, B. Zhang, H. Zhang, Q. Yang, S. Lin, Graph embedding and extensions: a general framework for dimensionality reduction, *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (1) (2007) 40–51.

- [33] S. Yan, J. Liu, X. Tang, T. Huang, A parameter-free framework for general supervised subspace learning, *IEEE Trans. Inf. Forensics. Secur.* 2 (1) (2007) 69–76.
- [34] S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, H. Zhang, Multilinear discriminant analysis for face recognition, *IEEE Trans. Image Process.* 16 (1) (2007) 212–220.
- [35] A. Ziyatdinov, S. Marco, A. Chaudry, K. Persaud, P. Caminal, A. Perera, Drift compensation of gas sensor array data by common principal component analysis, *Sens. Actuators B: Chem.* 146 (2) (2010) 460–465.
- [36] L. Fernandez, S. Guney, A. Gutiérrez-Gálvez, S. Marco, Calibration transfer in temperature modulated gas sensor arrays, *Sens. Actuators B: Chem.* 231 (2016) 276–284.
- [37] J. Fonollosa, L. Fernández, A. Gutiérrez, R. Huerta, S. Marco, Calibration transfer and drift counteraction in chemical sensor arrays using Direct Standardization, *Sens. Actuators B: Chem.* 236 (2016) 1044–1053.
- [38] L. Zhang, F. Tian, C. Kadri, B. Xiao, H. Li, L. Pan, H. Zhou, On-line sensor calibration transfer among electronic nose instruments for monitoring volatile organic chemicals in indoor air quality, *Sens. Actua. B* 160 (1) (2011) 899–909.

Biographies

Lei Zhang received his Ph.D degree in Circuits and Systems from the College of Communication Engineering, Chongqing University, Chongqing, China, in 2013. He is currently a Professor/Distinguished Research Fellow with Chongqing University. He was selected as a Hong Kong Scholar in China in 2013, and worked as a Post-Doctoral Fellow with The Hong Kong Polytechnic University, Hong Kong, from 2013 to 2015. He has authored more than 50 scientific papers in top journals, including the *IEEE Transactions on Instrumentation and Measurement*, the *IEEE Transactions on Image Processing*, the *IEEE Transactions on Multimedia*, the *IEEE Sensors Journal*, *Information Fusion*, *Sensors & Actuators B: Chemical*, and *Analytica Chimica Acta*. His current research interests include machine olfaction, machine learning, pattern recognition, computer vision, and multi-sensor system. 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.

Liu Yan received her Bachelor degree in Information Engineering in 2014 from Chengdu Polytechnic University, China. Since September 2014, she is currently pursuing a MS degree in Chongqing University. Her research interests include electronic nose and intelligent algorithm.

Zhenwei He received his Bachelor degree in Information Engineering in 2014 from Tianjin University, China. From July 2014 to June 2016, he worked in Chongqing Cable Network Inc. Since September 2016, he is currently studying for his Master degree in Chongqing University. His research interests include deep learning and computer vision.

Ji Liu received his Bachelor degree in Lamps and Lighting in 2016 from Hubei Engineering University, China. Since September 2016, he is currently studying for his Master degree in Chongqing University. His current research interests include machine learning and computer vision.

Pingling Deng received her Bachelor degree in Information Science and Engineering in 2015 from Lanzhou University, China. Since September 2015, she is currently pursuing a Master degree in Chongqing University. Her research interests include machine learning and electronic nose.

Xichuan Zhou received the B.S. and Ph.D degrees from Zhejiang University, China, in 2005 and 2010 respectively. Since the end of 2010, he is an Associate Professor and the Assistant Dean of the College of Communication Engineering, Chongqing University (CQU), China. He received the Young and Middle Age Backbone Faculty Award Program of Chongqing province in 2016. Dr. Zhou's research interests include high performance computing, circuit and system design and machine learning. He has published over 20 papers in the area of electronic engineering and information science on peer-reviewed journals including *IEEE Transactions on Circuits and Systems*, *IEEE Transactions on Biomedical Engineering* and *IEEE Transactions on Electron Devices*.