G Model
SNB-14774; No. of Pages 6
ARTICLE IN PRESS
Sensors and Actuators B xxx (2012) xxx–xxx
Contents lists available at SciVerse ScienceDirect
Sensors and Actuators B: Chemical
journal homepage: www.elsevier.com/locate/snb

Improving gas identification accuracy of a temperature-modulated gas sensor using an ensemble of classifiers

Amir Amini [a,*], Mohammad Ali Bagheri [b], Gholam Ali Montazer [b]

[a] Electronic Materials Laboratory, Electrical Engineering Department, K.N. Toosi University of Technology, Tehran 16315-1355, Iran
[b] Department of Information Technology, School of Engineering, Tarbiat Modares University, Tehran, Iran

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Ensemble classification
Gas recognition
Information fusion
Metal oxide gas sensor
Operating temperature modulation
Gas diagnosis

ABSTRACT

Data processing methods commonly used in conjunction with the array- and quasi-array-based gas identification systems generally include a dimension reduction followed by categorization using a classifier. Here, we have applied an ensemble of classifiers, directly to the high dimensional feature vectors and fused their verdicts by majority voting. The quasi-array investigated is a metal oxide sensor temperature-modulated with different rectangular heating voltage pulses. The experimental database was developed by recording the temporal responses obtained at different conditions to methanol, ethanol and 1-butanol vapors. Features related to each response were extracted by wavelet transform. The classification rates achieved with traditional methods were compared to that obtained using an ensemble of classifiers. The classification rate was improved by majority voting among the classifiers, each trained on different feature subsets, for the classification verdict.

1. Introduction

The problem of high dimensional data in e-noses, also referred to as "the curse of dimensionality" in statistical pattern recognition, significantly increase the complexity of the classification algorithm, time and memory requirements. Many of the features of the recorded patterns are irrelevant or redundant due to the cross-selectivity of the responses of the array components or the outputs of the virtual components of the virtual array utilized [1]. A simple strategy to reduce the number of features is to select a subset of the available features, feature subset selection (FSS). The goal of FSS is to find an optimal subset of features that maximizes prediction or classification accuracy. An exhaustive search of all possible subsets of features will guarantee that the optimal subset is found. However, this is computationally impractical even for a moderate number of features. The performance of different sensors and feature selection methods have been studied by various researchers in the electronic nose community [2–5], but the potential improvement in classification through feature fusion by ensemble-based approach [6–8] has not been a topic in research. While the feature selection seeks to find an optimal subset of features, the goal of classifier ensembles is to combine the outputs of diverse classifiers to achieve optimal accuracy. This approach generally belongs to the multiple classifier systems which are explained in detail in the following section.

The responses of a chemoresistor temperature-modulated with a heating voltage waveform contain significant amount of information related to the nature of the prevailing analyte in the background atmosphere [9–12]. Different voltage waveforms, such as staircases, pulse trains, sinusoidals, and step functions have been applied to the microheater of these sensors resulting in different success levels in analyte recognition by utilization of a single classifier [13–17]. Traditionally, a dimension reduction is carried out on the features of the recorded response patterns, and the classifier is trained based on the training database comprising all the low dimension features.

Here, we demonstrate the significance of using an ensemble of classifiers, in the categorization of gases using an operating temperature modulated gas sensor for the first time. A generic tin oxide-based gas sensor is temperature modulated with six different rectangular heating voltage pulses. The responses obtained are introduced to an ensemble of six classifiers, each trained on different feature subsets, for classification. The obtained classification rate is compared with those achieved with traditional processing methods utilizing a single classifier trained on all the recorded feature database.

2. Multiple classifier systems

Combining multiple classifiers to achieve higher accuracy is one of the foremost research areas in machine learning. It is known under various names, such as multiple classifier systems, classifier

* Corresponding author. Tel.: +98 21 8406 2328; fax: +98 21 8876 8289.
E-mail addresses: amir.amini@ee.kntu.ac.ir, amir.amini.elec@gmail.com, amir.amini1@yahoo.com (A. Amini), a.bagheri@modares.ac.ir (M.A. Bagheri), montazer@modares.ac.ir (G.A. Montazer).

0925-4005/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.snb.2012.10.140

Please cite this article in press as: A. Amini, et al., Improving gas identification accuracy of a temperature-modulated gas sensor using an ensemble of classifiers, Sens. Actuators B: Chem. (2012), http://dx.doi.org/10.1016/j.snb.2012.10.140

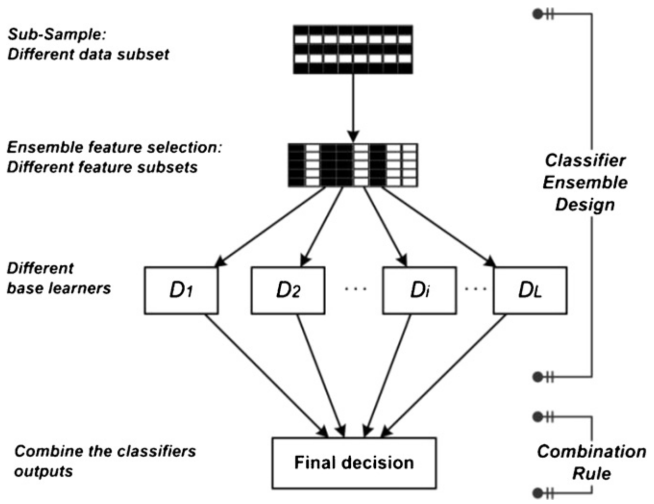

Fig. 1. The structure of a multiple classifier system.

ensemble, committee of classifiers, and classifier fusion. Multiple classifier systems can generate more accurate classification results than each of the individual classifiers [18]. In such systems, as shown in Fig. 1, the classification task can be solved by integrating different classifiers, leading to better performance. However, the ensemble approach depends on the assumption that single classifiers make errors on different samples, known as classifier diversity [19]. The intuition is that if each classifier makes different specific errors, then the total errors can be reduced by an appropriate combination of these classifiers. There are three general approaches to create an ensemble of classifiers. The most straightforward approach is using different learning algorithms for the base classifiers or variations of the parameters of the base classifiers. For example, different initial weights or different topologies of a series of neural network classifiers can be utilized as different base classifiers. Another approach is using different training sets to build different base classifiers. Such sets are often obtained from the original training set by re-sampling techniques [20,21].

The third approach, which is employed in this work for classification of the response patterns of a thermally modulated gas sensor, is to train the individual classifiers with data that consist of different feature subsets, or so-called ensemble feature selection [22]. While traditional feature selection algorithms seek to find an optimal subset of features, the goal of ensemble feature selection is to find different feature subsets to generate accurate and diverse classifiers. In the random subspace method [6], the ensemble system is built by randomly choosing the feature subsets. These feature subsets are generated by randomly selecting m features from the n-dimensional feature space ($m < n$). Then, each feature subset is fed into an individual classifier. Finally, all classifiers are aggregated by an appropriate combination rule. While common classification systems suffer from the high-dimension data, the ensemble feature selection approach takes advantage of large number of features [6].

Regardless of the base classification algorithm used, the diverse classifiers are then fused by a combination technique such as voting methods, fuzzy integral, Markov chains, Dempster-Shafer rule, and behavior knowledge space [18].

3. Experimental

Experiments were carried out to establish a database of response patterns recorded in air contaminated with three different volatile organic compounds each at different concentrations. The sensor used is a generic tin oxide-based general sensitivity gas sensor (SP3-AQ2, FIS Inc., Japan), which is operating temperature-modulated

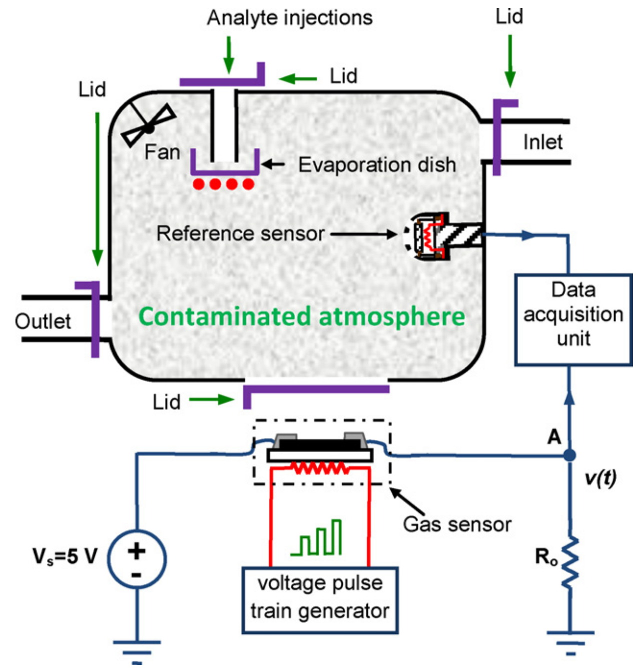

Fig. 2. The schematics of the experimental setup.

with different voltage waveforms applied to its microheater. The microheater of the sensor is a ruthenium oxide-based thick film with a room temperature resistance of ~57 Ω. The experimental setup is schematically shown in Fig. 2.

An atmosphere-controlled chamber of 8-l volume (Fig. 2) is utilized for response recordings. The background gas is clean air, and the analytes examined are methanol, ethanol, and 1-butanol vapors at concentration levels ranging from 100 to 2000 ppm. The clean air in the chamber is contaminated by injecting predetermined volumes of the liquid chemicals with a sampler, which are evaporated in the closed chamber. The chamber air is mixed with a small electric fan for homogenization. The contamination level is continuously monitored using a reference sensor. Data obtained from the reference sensor is used only for the purpose of comparison of the experimental conditions; analyte concentrations given in relation to the response patterns are calculated based on the amount of the liquid analyte injected to the chamber. No specific control was imposed, and no compensation measure was taken [23], on the temperature and humidity level of the chamber atmosphere. Throughout all the experiments; the respective ranges of temperature and relative humidity were 22–27 °C and 30–50%. The chamber has a gas impermeable lid which opens to allow the insertion of the temperature-modulated gas sensor. The sensor is mounted on a ceramic probe through which the heating and response signal carrying wires are connected to the related peripheral electronic units.

Voltage pulses of different temporal profiles are generated using a computer programmable "multifunction card" (PCI-1711L, Advantech Co., USA) connected to the microheater of the gas sensor via a custom-designed buffer circuit which prevents overloading of the card. The circuit used for obtaining the responses, shown in Fig. 2, converts the variations of the resistance of the sensor pallet to a voltage signal. This includes a 2.2 kΩ resistor (R_o) connected in series to the gas sensor and a DC voltage source. The voltage drop on R_o is approximately proportional to the conductance of the gas sensor. Regardless of this approximate relationship, the temporal signal, $v(t)$, measured at point A, is considered as the response pattern related to the prevailing contaminated atmosphere. $v(t)$ is sampled at a rate of 100 s^{-1}, digitized and recorded.

Please cite this article in press as: A. Amini, et al., Improving gas identification accuracy of a temperature-modulated gas sensor using an ensemble of classifiers, Sens. Actuators B: Chem. (2012), http://dx.doi.org/10.1016/j.snb.2012.10.140