Optical and Acoustical Methods in Science and Technology

Gas Sensing Supported by Pattern Recognition

C. Tyszkiewicz, A. Szpakowski and T. Pustelny

Department of Optoelectronics, Silesian University of Technology, Krzywoustego 2, 44-100 Gliwice, Poland

The system composed of the array of eight semiconductor, chemoresistive gas sensors was used for the classification of hydrogen, methane and carbon oxide gaseous samples. The classification task was performed by pattern recognition methods applied to the multivariate response of the array. The pattern recognition scheme used for classification uses a feature subset selection algorithm coupled with an objective function designed for clustering and a multilayer perceptron classifier.

PACS numbers: 02.50.Sk, 07.07.Df, 84.35.+i

1. Introduction

From the practical point of view, chemoresistive gas sensors based on SnO_2 belong to the most important sensor type due to their high sensitivity, good accuracy, short response time, good stability and suitability to portable instruments [1]. The major disadvantage of a tin-oxide based sensor is its non-selectiveness and cross--sensitivity [2]. Selectivity can be enhanced by doping of the surface of the sensing layer with noble metals as palladium or platinum [1, 2]. Metallic surface dopants act as specific adsorption sites for the oxygen species and catalytic oxidation of reducing gases taking place on the surface of the sensor layer. On the other hand, selectivity can be increased by means of pattern recognition methods [3, 4]. Pattern recognition is a process in which raw data are collected and then processed basing on the category to which data belongs. Cross-sensitivity is in this case a very important feature of the sensors because the system which uses the pattern recognition methods demands superfluous input information from the array of sensors.

2. Pattern recognition scheme

In the case of qualitative analysis of the sensed environment the pattern recognition procedure is composed of four steps: pre-processing, dimensionality reduction, classification and identification. In most cases, the aim of pre-processing is shifting, compressing and normalizing of the raw steady-state and transient-state signals from the sensor array in order to improve the performance of the subsequent steps. After pre-processing raw data are converted into the feature vector in high dimensional space which is a descriptive parameter of the sensor array response. In most cases after pre-processing the dimensionality of the feature vector is in the order of millions. Processing of such data demands a reduction of dimensionality to hundreds or at most thousands dimensions. The initial set of features is mapped into the low-dimensional feature space that preserves most of the information accumulated in the original feature set. This reduction is usually achieved by means of feature extraction techniques or feature selection techniques [4, 5]. Reduced feature vectors are used then in the classification step. In this step most important is the procedure that clusters together the feature vectors related with the given chemical composition of the sensed environment. This step can be quite extensive and can be extended by means of further dimensionality reduction techniques working in union with feature space search techniques [4, 6]. One of the most popular types of classifier is a multilayer perception which establishes the classification boundaries in a reduced feature space [4]. In the last step, the identification step, the signal from the array of sensors is assigned to a class and the occurrence of a particular pattern is recognized.

3. Experimental setup

The sensor array used in the experiment is composed of chemoresistive gas sensors based on SnO_2. In our setup we are using a group of eight Figaro sensors, whose list is given in Table.

TABLE

Sensors forming the array and gases for which the sensors are destined in compliance with data sheets provided by Figaro Engineering Inc.

Sensor	Destination	Sensor	Destination
TGS800	general air contaminations	TGS2602	C_3H_8, C_4H_{10}, CH_4
TGS813	CH_4, C_3H_8, C_4H_{10}	TGS2610	CH_4, C_4H_{10}, H_2
TGS842	H_2, CO	TGS2611	C_3H_8, C_4H_{10}, CH_4
TGS2600	general air contaminations	TGS2620	C_2H_5OH, H_2, C_4H_{10}

The sensors worked in a continuous flow controlled system at a constant temperature mode. The time of exposition and purging periods was sufficient to achieve a stationary response for each sensor. The measurements of the conductivity of the sensors were made automatically and the data were sampled with a frequency of 10 Hz. The pre-processing procedure consists of a logical analysis of the signal from the sensor array on the basis of the signal from the gas controller in order to find the starting and ending points of adsorption/desorption cycles. The registered signal was filtered using the Hanning window. Finally, the baseline compensation for two modes, differential and proportional, was performed and the baseline conductivity of each of the sensors was calculated.

4. Experimental results

The sensor array was exposed to cyclic interaction with different concentrations of hydrogen (H_2), methane (CH_4) and carbon oxide (CO). All gases were diluted in synthetic air of a known relative humidity. In all experiments the total gas flow rate was 100 sccm. The concentrations of the gases were respectively H_2: 0–20000 ppm, CH_4: 0–500 ppm, CO: 0–500 ppm.

Up to 70 features for each sensor were calculated for each gas sample of a given concentration. These yields up to 560 dimensions in feature space. In the feature extraction phase this dataset was reduced to three dimensions. The reduction was performed using a combination of feature extraction and feature selection techniques. The features were selected by means of filter approach. The employed search strategy was based on the plus-L minus-R selection algorithm (LRS) [7, 8], with the independent component analysis (ICA) as a dimensionality reduction technique [9]. In our software the JADE implementation of the ICA algorithm was used [10]. For the purpose of evaluation our own method was used. The method is based on the combination of scattering matrices, calculating variances of the distances of the cluster centres and employing the penalty functions. The method strongly favours the situation where the clusters corresponding to each class (in our case each gas) are evenly distributed and are of the even size. Additionally favoured are situations where clusters are small and distant from each other. Such a distribution is favourable for each subsequent classification step.

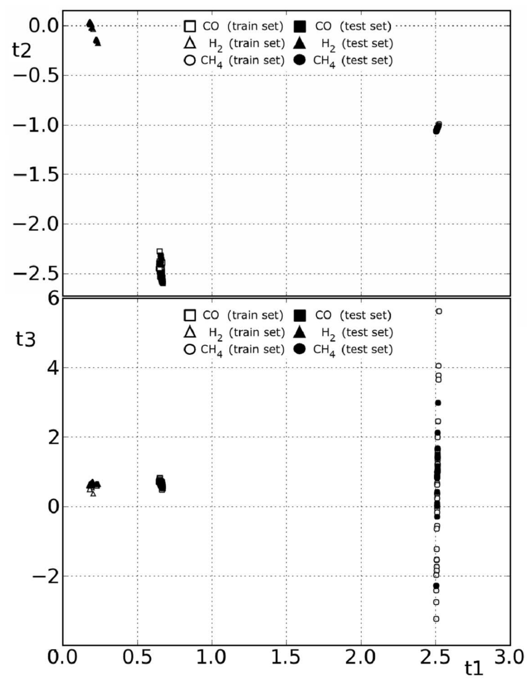

Fig. 2. Set of feature vectors in an optimized three-dimensional secondary feature space (t1, t2, t3 — independent components).

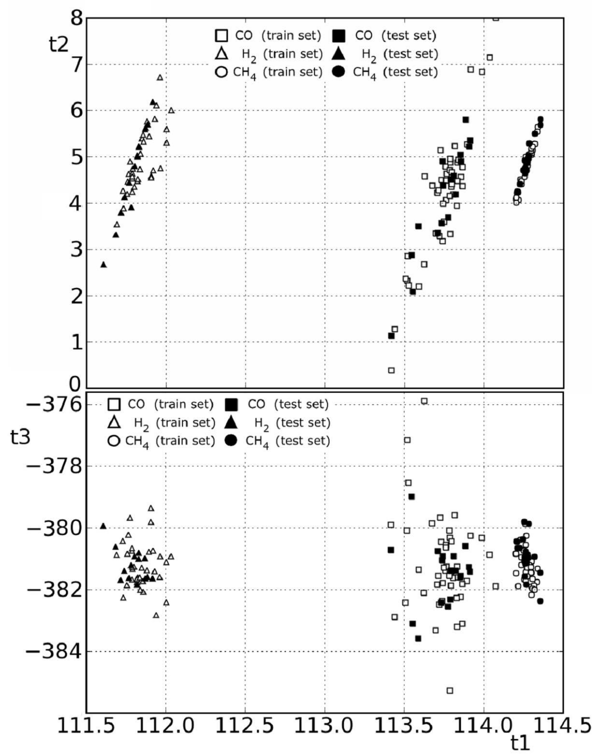

Fig. 1. Set of feature vectors in a unoptimized three-dimensional secondary feature space (t1, t2, t3 — independent components).

As a result, our FSS implementation was able to perform accurately a separation of three gases under consideration regardless of their concentration. The results of optimization of low-dimensional feature space are shown in Figs. 1 and 2. Set of features in unoptimized feature space is shown in Fig. 1, whereas the optimized feature space is shown in Fig. 2. It can be seen that the clusters in the optimized feature space are far more compact than in the unoptimized one. The complexity of the classifier depends strongly on the secondary feature space optimization.

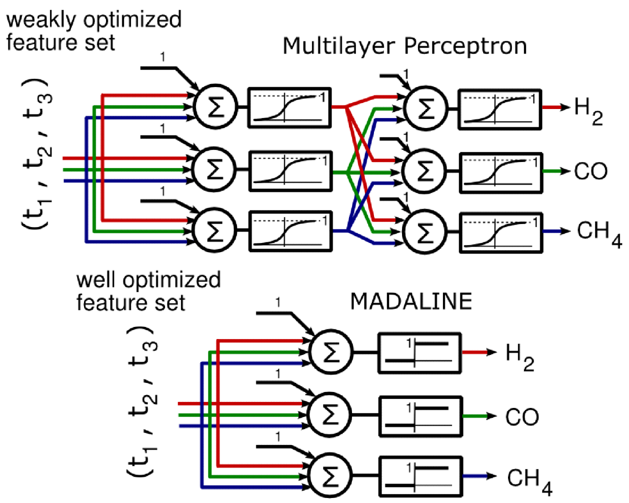

Fig. 3. The structure of the classifiers for unoptimized and optimized secondary feature spaces.

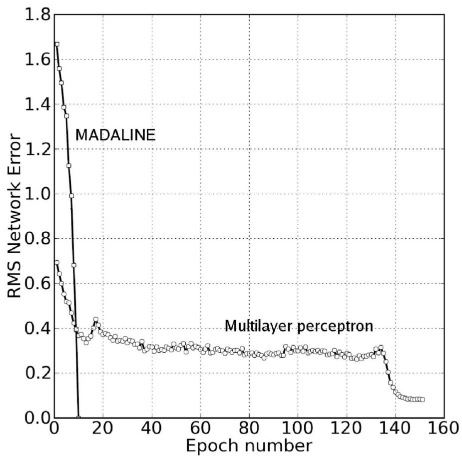

Fig. 4. The dependence of the RMS error of multilayer perceptron and the MADALINE network in the subsequent learning epochs.

In order to find decision boundaries two types of neural network classifiers were used. Their structures are shown in Fig. 3. In the case of the unoptimized feature space the multilayer perception learned by using the back propagation algorithm was used. In the case of the optimized feature space the simple MADALINE network was able to find the decision boundaries and classify correctly all the features from test set. The numerousness of learning and test sets were 40 and 25 samples, respectively. The comparison of learning process of both networks is shown in Fig. 4. Optimization of secondary feature space simplifies significantly the classifier and makes the learning process shorten.

5. Conclusions

Using the array of eight, non-selective, general-purpose SnO_2 sensors, the employed pattern recognition methods allow us to construct the simple non-parametric classifier, which is able to recognize and classify hydrogen, methane and carbon oxide gas samples. The sensors were working at constant temperature mode in a continuous-flow chamber. Future works will involve using the sensors in the temperature modulation mode, in closed chambers and in free air. Additionally the numerical analysis will be extended in regression methods in order to classify gas mixtures and to evaluate their concentrations.

References

[1] G. Korotcenkov, *Mater. Sci. Eng. B* **139**, 1 (2007).

[2] W. Göpel, K.D. Schierbaum, *Sensors Actuators B* **26-27**, 1 (1995).

[3] C. Di Natale, F. Davide, A. D'Amico, *Sensors Actuators B* **23**, 111 (1995).

[4] R. Gutierrez-Osuna, *IEEE Sensors J.* **2**, 189 (2002).

[5] R. Gutierrez-Osuna, H. Troy Nagle, *IEEE Trans. Systems. Man Cybernetics-Part B Cybernetics* **29**, 626 (1999).

[6] E. Cantu-Paz, S. Newsam, C. Kamath, in: *Proc. Tenth ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, Seattle (WA) USA 2004*, p. 788.

[7] A. Perera, T. Sundic, A. Pardo, R. Gutierrez-Osuna, S. Marco, *Sensors J. IEEE* **2**, 235 (2002).

[8] A. Szpakowski, C. Tyszkiewicz, T. Pustelny, *Acta Phys. Pol. A* **114**, A-237 (2008).

[9] A. Szpakowski, T. Pustelny, K. Barczak, C. Tyszkiewicz, *J. Phys. IV (France)* **129**, 155 (2005).

[10] L. De Lathauwer, B. De Moor, J. Vandewalle, *J. Chemometrics* **14**, 123 (2000).