Sensors and Actuators B 148 (2010) 550–558

Contents lists available at ScienceDirect

Sensors and Actuators B: Chemical

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

A neural net implementation of SPCA pre-processor for gas/odor classification using the responses of thick film gas sensor array

N.S. Rajput*, R.R. Das, V.N. Mishra, K.P. Singh, R. Dwivedi

Institute of Technology, Department of Electronics Engineering, Banaras Hindu University, Varanasi, UP 221005, India

ARTICLE INFO

Article history:
Received 11 March 2010
Received in revised form 3 May 2010
Accepted 20 May 2010
Available online 1 June 2010

Keywords:
Standardized principal component analysis (SPCA)
SPCA pre-processor
SPCA transforming neural net
Electronic nose

ABSTRACT

In this paper, an artificial neural net (ANN) implementation of SPCA pre-processing is presented for its use with a neural classifier trained with SPCA transformed data. Here, a SPCA transforming neural stage (*Net* I_{SPCA}) is placed before a SPCA trained neural classifier stage (*Net* II_{SPCA}). Accordingly, newer sensor array response of respective gas/odor can now be classified, more precisely, using *Net* II_{SPCA} fed through (*Net* I_{SPCA}). This way, newer sensor response gets transformed to corresponding SPCA transformation, with conformity, for its classification using *Net* II_{SPCA}. Efficacy of this scheme is demonstrated by considering thick film tin oxide sensor array response data for four gases/odors (viz. acetone, carbon tetra-chloride, ethyl methyl ketone and xylene). Experimentally, respective nets, *Net* I_{SPCA} and *Net* II_{SPCA} comprising of 11 and 04 neurons were, respectively, trained in just 78 and 12 epochs of 42×4 training vectors for the aforesaid gases/odors. The SPCA transformation derived mathematically and obtained through *Net* I_{SPCA}, carried a mean squared error (MSE) of 1.81×10^{-9} for 18 test sensor responses not used for training *Net* I_{SPCA}. As well as, 100% correct classification was achieved for the aforesaid 18 independent samples using *Net* II_{SPCA} with a MSE of 1.22×10^{-9}.

1. Introduction

Development of electronic nose (e-nose) has attracted major interest of researchers mainly for their applicability and necessity in variety of real life applications [1,2]. The ability of e-nose to precisely and continuously monitor and measure the presence of hazardous gases in real time ambience is crucial to prevent respective accidents and for accounting the short term and long term exposure of the workers to major industrial pollutants [3–7]. A simpler and effective e-nose with improved real time performance is therefore 'the need of the day'. For developing high performance e-noses, standard methodology consists of an array of gas sensors followed by stages for signal conditioning, pre-processing and pattern recognition for classification of gases/odors [1,2]. A basic block schematic of such an e-nose is illustrated in Fig. 1.

The raw as well as the conditioned sensor responses of the gas sensor array are usually complex and carry mutually correlated parameters in high dimensional hyperspace. The sensor response is therefore processed for its feature enhancement by translating the sensor responses along mutually orthogonal principal axes (along which the transformed data is uncorrelated) using many a mapping methods such as principal component analysis (PCA), standardized principal component analysis (SPCA), linear discriminant analysis (LDA), blind source separation (BSS), independent component analysis (ICA), neuroscale, kernel PCA, exploratory projection pursuit, etc. Nuances of these techniques and steps in pattern recognition using e-nose has been widely reviewed and referred in Refs. [1,2]. The data transformed into mutually uncorrelated form along its principal components, still remains an equivalent of the original sensor response data and classifiers perform better when they are trained using this rather more clarified (uncorrelated) data.

Typically, application of PCA/SPCA technique requires that the sensor response data must be first normalized either for zero mean (called PCA) or should be standardized for zero mean and unit variance (called SPCA). A feature vector is then obtained for transformation of this normalized/standardized data set. Finally, the normalized/standardized data set is transformed onto the principal component axes of PCA or SPCA, respectively, by multiplying the respective feature vector with the normalized/standardized data set. This transformed data set is partitioned to obtain training, validation and testing data sets.

Along these principal components (PCs) transformed data is uncorrelated and a few of its major principal components carry most of the information of the original data while dimensionality reduction can be obtained by not considering its less important principal components.

In general, the classification performance of neural classifiers when trained with PCA/SPCA transformed data is very promising justifying its popular use [8].

* Corresponding author. Tel.: +91 542 6701247; fax: +91 542 2368428.
E-mail address: nsrajput.ece@itbhu.ac.in (N.S. Rajput).

0925-4005/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.snb.2010.05.051

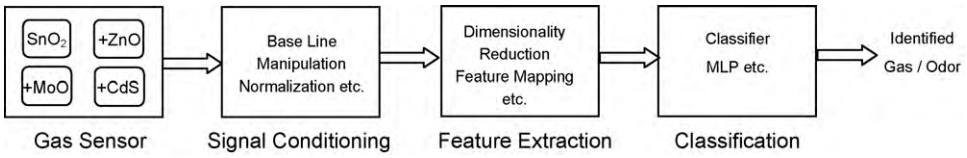

Fig. 1. Basic block schematic of an e-nose.

In recent literature, the traditional techniques such as PCA and SPCA still find a popular and effective use for developing better implementations for e-nose. Xua et al. used a 5-element nano-structured gas sensor array to classify 4 volatile organic compounds. They transformed 120 vectors of sensor response data using PCA. Out of these 120 vectors, 90 PCA transformed vectors considering its 3 major principal components were used to train a fuzzy-ARTMAP network while testing was carried out with the remaining 30 PCA transformed vectors [9]. Also, Alizadeh and Zeynali employed a 3-element sensor array for classification and quantification of binary mixture of two gases using a neural net trained with 20 vectors using 2 principal components of SPCA transformed sensor responses and tested the classification performance using 21 response vectors. These 41 vectors used for training and testing were created typically in a single batch of SPCA processing [10]. A neural classifier was also trained and tested by Siripatrawan using 2 principal components of their 108 vectors of SPCA transformed data set. This SPCA processed 108 data set was then divided in the ratio of 70:30 for training and testing, respectively [11]. Lv et al. trained a neural classifier for gas classification using 1175 training and 375 testing vectors in the form of 2 principal components of the SPCA transformed data set of their 8 element SnO_2 sensor responses [12]. Similarly, Beltran et al. first PCA processed their primary data set with 1000 vectors and utilized 20 major principal components for training a neural net to identify three classes of Chilean wines. Again, their batch processed PCA transformed data containing 1000 vectors was divided into 900 vectors for training and 100 vectors for testing [13].

As can be inferred from Refs. [8–13], in PCA and SPCA, the transformation is obtained in one go. This batch transformed data is then divided into training, validation and testing sets for respective tasks in pattern recognition using neural network. Mandatory requirement of data normalization/standardization and subsequent transformation with 'data set specific feature vector' inhibits the end user to test another sample which is not available within the batch transformed set. Any newly generated data, therefore, cannot be fed to the e-nose for its classification.

The primary limitation of implementing PCA/SPCA is that if a newer response $x_{n+1} \notin$ *data set* $\{x_1, \ldots, x_n\}$ of the gas sensor for the considered gases/odors is generated; it cannot be tested using the earlier trained neural network as the PCA/SPCA transformation of x_{n+1} cannot be obtained which conforms to the PCA/SPCA transformation of *data set* $\{x_1, \ldots, x_n\}$ Further, if we wait to gather freshly generated m samples for test, where this *test set* $\{t_1, \ldots, t_n\} \notin$ *data set* $\{x_1, \ldots, x_n\}$, the m samples must be covering whole experimental ranges for the gases/odors under consideration, to achieve the same zero mean and unit variance as that of n samples.

Accordingly, in an attempt to develop an interface, whereby, freshly generated samples could be tested using a PCA/SPCA trained neural classifiers, Saleh and Hoyle [14] developed a pre-processor, a post-processor and trained a multi-layer perceptron (MLP) classifier neural net. They first obtained 2 principal components of the transformed SPCA version of their training data using a mathematical per-processor and retained the 'training data SPCA transformation feature vector'. Next, they gathered additional samples, standardized them to their zero mean and unit variance and transformed them mathematically using a post-processor to obtain a conformal SPCA making use of the 'training data SPCA transformation feature vector'. Now suppose they had only one sample instead of m samples, then the suggested scheme cannot essentially handle the situation because this single sample cannot be normalized to zero mean and unit variance. Again, if m samples with non-uniform distribution are to be classified, the transformation obtained using the post-processor proposed will not conform to original distribution as the normalizing mean and standard deviation will be biased because of non-uniform distribution. Lastly, in practical situations, there will always be only one sample, at a particular point of time, accurate classification of which will be desired.

In this paper, we are therefore presenting an artificial neural net (ANN) implementation of SPCA pre-processor for use with neural classifiers for gas/odor classification, which have been trained with SPCA transformed data. We propose and demonstrate a simple implementation of ANN classifier utilizing statistical SPCA transformation. This implementation has the capability to classify even those samples which are freshly generated from the gas sensor array. The proposed implementation has been demonstrated using SPCA transformation for the fact that the performance of SPCA is usually better than PCA data transformations [15].

The basic block schematic of the proposed '*SPCA trained e-nose*' is shown in Fig. 2. It can accept a freshly generated sensor response for its correct classification.

For designing *Net* I_{SPCA}, implementation was carried out using raw sensor response data which covered the whole experimental range for considered gases/odors. The raw sensor responses have been taken as the training inputs while the respective mathematically transformed SPCA data were designated as the targets for training the neural net *Net* I_{SPCA}. *Net* II_{SPCA} has been designed by taking the duly transformed SPCA data as the training input and respective classes of the considered gases/odors have been taken as the training target. Evaluation of performance of *Net* I_{SPCA} and *Net* II_{SPCA} is then carried out for 18 independent test responses which were not used for training these nets.

2. Materials and methods

2.1. The data selection

To test the proposed methodology, authentic, already published and accepted data was sought for. Accordingly, responses of a gas sensor array for acetone, carbon tetra-chloride, ethyl methyl ketone and xylene has been taken [16]. The sensor referred in their experiments consisted of an array of four sensor elements with an integrated heater on a substrate. The array was fabricated with SnO_2 as a base material and doped with different materials, namely, ZnO (sensor A), undoped SnO_2 (sensor B), MoO (sensor C) and CdS (sensor D). The sensor array was fabricated using thick film technology. The fabrication process is described in detail in Ref. [16]. Their test set up used to collect the sensor array response data and the experimental conditions in the laboratory has also been described in the reference. The four gases/odors used in the referred experiment are acetone (ace), ethyl methyl ketone (emk), carbon tetra-chloride (car) and xylene (xyl) and the outputs measured were percent change in resistance of different sensors when exposed to various concentrations of these gases. As men-

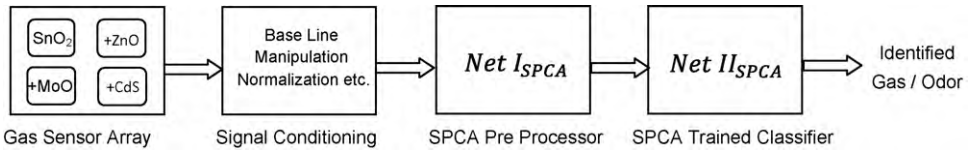

Fig. 2. Basic block schematic of proposed SPCA based implementation of e-nose.

tioned earlier, data used in the present analyses are derived from the results of Ref. [16] and therefore, the sensor characteristics from Ref. [16] have been reproduced here for ready reference in Fig. 3(a)–(d).

2.2. The data extraction

The graphs of Fig. 3 (a)–(d) were imported into the software named 'Precision Image Digitizer (PID)' ver. 1.3.0.0′ and two categories of data were extracted as Data Set I having 42 vectors (with 8 vectors for ace, 10 vectors for car, 12 vectors for emk and 12 vectors for xyl) and Data Set II consisting of 18 vectors (with 3 vectors for ace, 4 vectors for car, 6 vectors for emk and 5 vectors for xyl). The sample vectors were generally extracted with uniform spacing to cover the whole concentration axis of the graphs. The data extracted from PID was initially in terms of the pixel values of the coordinate plane of the PID window. The actual gas sensor array response data were then developed using (1)–(3), i.e. by shifting the origin of PID's window to match the origin of the graph and by scaling the pixel coordinates w.r.t. the graph's axes. An explanatory diagram for data extraction using PID is shown in Fig. 4.

For all data points, origin shifting and coordinate scaling was carried out using (1)–(3),

$$(x_i' y_i') = (x - x_0, y - y_0), \tag{1}$$

$$Conc.\ (x_i) = \frac{x_i'}{x_{\max} - x_0 / x_{axis\ range\ of\ graph}} \tag{2}$$

and

$$Response\ (y_i) \frac{y_i'}{y_{\max} - y_0 / y_{axis\ range\ of\ graph}} \tag{3}$$

Here, for all pixel coordinates (x, y) of various data points, x_i', y_i' are respective coordinates in terms of the graph parameters. Also, (x_0, y_0) is the pixel coordinate of the origin of the graph while $x_{\max}$ and $y_{\max}$ are the respective maximum values of x-axis and y-axis pixels on the graph axes. In Fig. 4 '*x-axis range of graph*' and '*y-axis range of graph*' are the represented ranges of change in percentage concentration and the range of percentage change in resistance, respectively; taken appropriately for each gas/odor.

Finally, at the ith data point x_i', y_i' is the true coordinate while *Conc.* (x_i) and *Response* (y_i) is the true response of the sensor array on the graph.

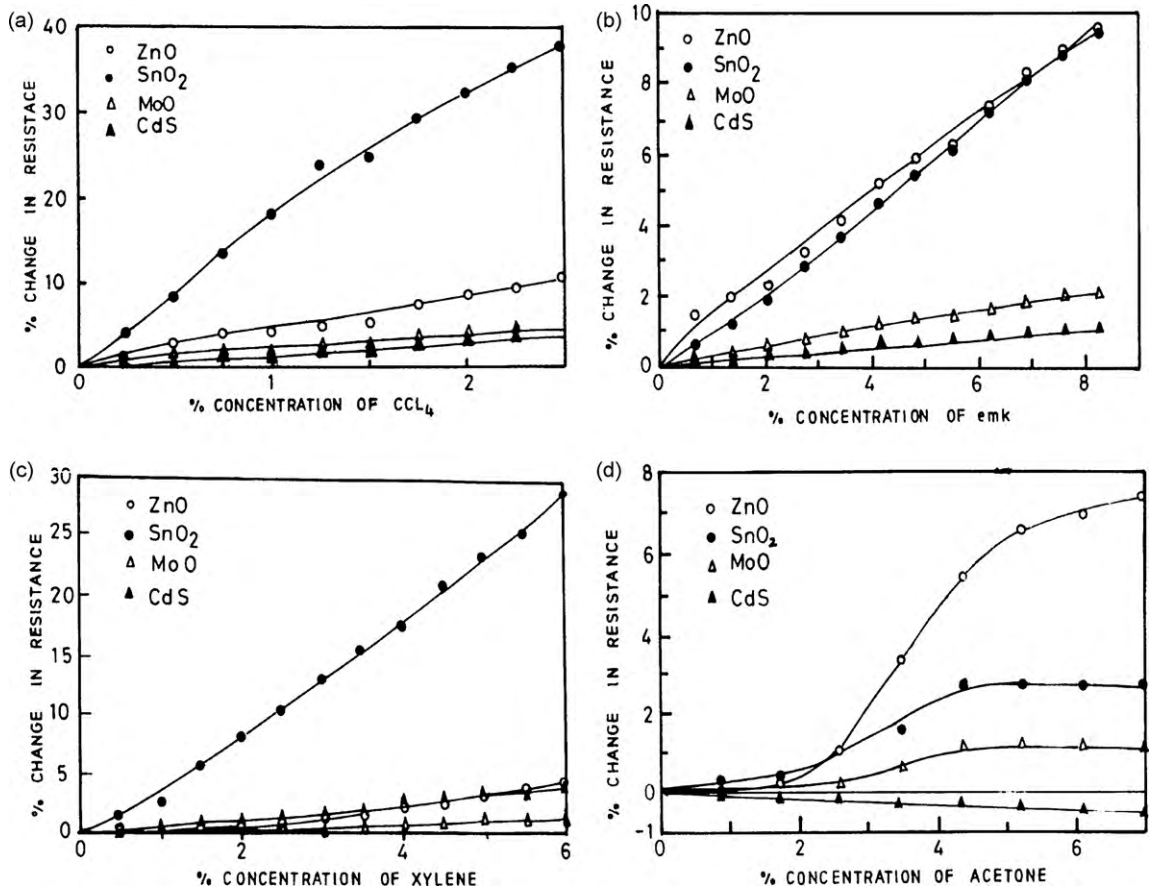

Fig. 3. Sensor characteristics for (a) Carbon Tetra-chloride (b) Ethyl Methyl Ketone (c) Xylene and (d) Acetone exposures, in N_2 ambience, under energized conditions (2 W Supply).

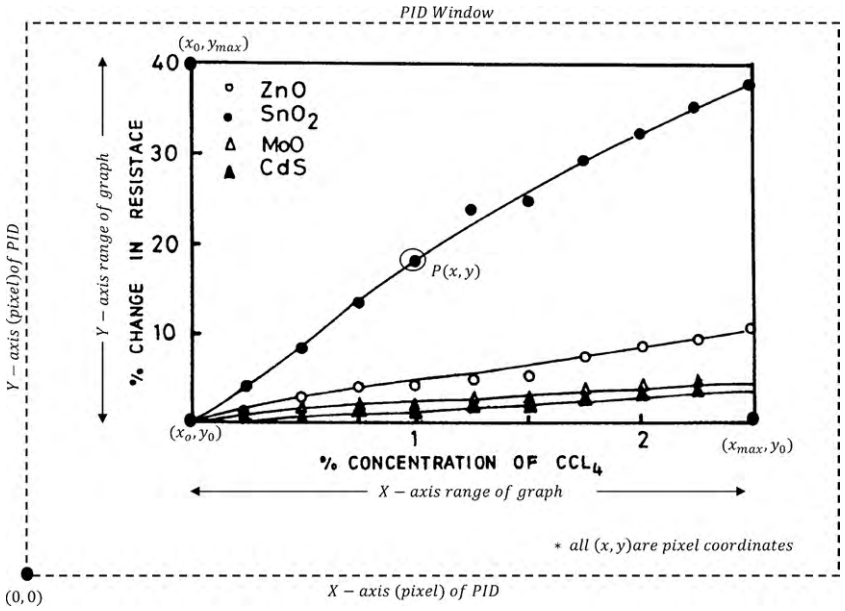

Fig. 4. Graphical demonstration of data extraction using PID tool.

The true gas sensor array response data was hence intrinsically obtained for all the four gases/odors viz. ace, car, emk and xyl. The Data Set I was used for training and Data Set II was used for testing of the proposed e-nose development methodology. The experimentation has been carried out following the standard validation practices and recommendations, to ensure that the experiment is valid for any engineering application [2].

2.3. Contextual outline of SPCA

Principal component analysis (PCA) and standardized principal component analysis (SPCA) are techniques for multivariate analysis to transform a set of correlated variables into a set of uncorrelated variables. Consider a data set $\boldsymbol{Y}$ consisting of m variables with n number of observation samples. Each data element of $\boldsymbol{Y}$ can be denoted by $y_{i,j}$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$. This data matrix $\boldsymbol{Y}$ is normalized to carry zero mean and unit variance so that the centroid of the whole data set shifts to origin and follows normal distribution. For zero mean and unit variance normalization (called standardization), first, the mean ($\bar{y}_j, j = 1, \ldots, m$) of each column is subtracted from each element of respective column. Now each element of respective column is divided by the standard deviation (s_j, $j = 1, \ldots, m$) of the respective columns. The data set $\boldsymbol{Y}$ standardized to carry zero mean and unit variance will hereafter be called as $\boldsymbol{X}$. Each data element of $\boldsymbol{X}$ can be denoted by $x_{i,j}$ for all $i = 1, \ldots, n$ and $i = 1, \ldots, m$. It is the $\boldsymbol{X}$ matrix that undergoes SPCA transformation and the principal components of SPCA form m linear combinations [15] for all $i = 1, \ldots, n$.

$$\left.\begin{aligned} PC_1(i) &= w_{11}x_{i1} + w_{21}x_{i2} + \cdots + w_{m1}x_{im} \\ PC_2(i) &= w_{12}x_{i1} + w_{22}x_{i2} + \cdots + w_{m2}x_{im} \\ &\vdots \\ PC_m(i) &= w_{1m}x_{i1} + w_{2m}x_{i2} + \cdots + w_{mm}x_{im} \end{aligned}\right\} \quad (4)$$

$$\sum_{i=1}^{n} x_{i,j}^2 = 1, \quad \text{for any } j = 1, \ldots, m; \quad \text{while} \quad (5)$$

$$w_{i1}w_{j1} + w_{i2}w_{j2} + \cdots + w_{im}w_{jm} = 0; \quad \text{for all } i \neq j \quad (6)$$

where new variables $PC_1(i), PC_2(i), \ldots, PC_m(i)$ are m principal components (PCs) of the transformation.

$$\boldsymbol{FV} = \begin{bmatrix} w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mm} \end{bmatrix}^{\mathrm{T}} \quad (7)$$

The matrix at (7) is called the feature vector ($\boldsymbol{FV}$) for $\boldsymbol{X}$ and is operated on $\boldsymbol{X}$ to obtain the SPCA transformation of $\boldsymbol{X}$.

The $m \times m$ correlation matrix $\Re$ of $\boldsymbol{X}$ is defined as

$$\Re_{\boldsymbol{x}} = \frac{\boldsymbol{X}^{\mathrm{T}} \times \boldsymbol{X}}{n-1} \quad (8)$$

Now the Eigen values and Eigen vectors are calculated for the matrix $\Re_{\boldsymbol{x}}$. The Eigen vectors are then arranged in descending order of respective Eigen values. Such a rearranged set of Eigen vectors becomes the feature vector ($\boldsymbol{FV}$) of $\boldsymbol{X}$ for carrying out SPCA transformation of $\boldsymbol{X}$. Multiplying the transposed matrix of both the $\boldsymbol{FV}$ and $\boldsymbol{X}$ creates the SPCA transformed data ($\boldsymbol{X_{SPCA}}$), in its transposed form. The columns of $\boldsymbol{X_{SPCA}}$ are now the principal components (mutually orthogonal principal axes) of the desired SPCA transformations along which the data is uncorrelated.

The SPCA transformation reserves full information of the original data while at the same time eliminates a certain amount of noise. It is also a widely used tool for dimensionality reduction of the standardized data. For dimensionality reduction of the standardized data set, depending on the amount of percentage information accounted by its m Eigen values, less information carrying principal components can be dropped. This discarding of less important principal components amounts to loosing certain calculable percentage of information.

2.4. Contextual outline of multi-layer perceptron (MLP)

MLPs are feed-forward ANN models with three adjacent layers; the input, hidden and output layers [17]. Each layer consists of several neurons. The general structure of a MLP ANN is illustrated in Fig. 5.

MLPs learn from the input–output training samples and become capable of giving outputs for inputs which it had not seen earlier. The training process employs a learning algorithm that helps the

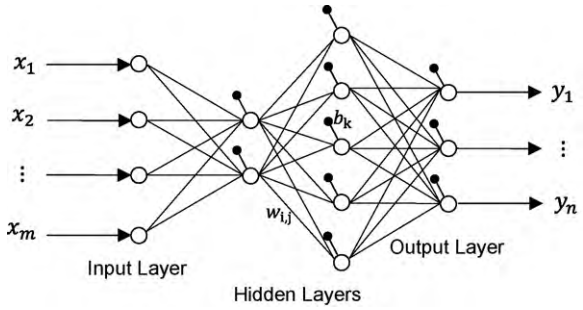

Fig. 5. General architecture of a MLP ANN.

MLP develop a mapping function between the inputs and outputs. Basically, during the learning process, the input neurons receive the training input data from the external environment (denoted as $x_1, x_2, \ldots, x_m$) and pass them to the neurons of hidden layer. Each neuron of the hidden layer performs mathematical computation using its weighted inputs and bias. The cumulative response is then mapped onto appropriate activation function of each neuron and an output is produced. The outputs of the hidden layer are fed to the neurons of the next hidden layer and similar computation is carried out until final outputs at the output layer is produced (denoted as $y_1, y_2, \ldots, y_n$).

Now during the training process each computed output is then compared with its targeted output for respective input and the error is calculated. This error is then minimized in several iterations through weight and bias updation until error value is reduced within an acceptable range. On completion of the training process, the MLP should now be able to produce appropriate outputs for any input data of the data set based on the generalized mapping which it has developed.

The performance of the MLP very much depends on its generalization capability, which in turn depends upon the training data representation. One important characteristics of data representation is its uncorrelatedness. Better performance using simpler nets may be obtained if the training data is uncorrelated. This suggests a need for eliminating correlation in the training and validation data before they are used for MLP learning. Once MLP is trained using uncorrelated data, although better performance with simpler nets is gained but it also puts a constraint that a newer input to the MLP should also be in uncorrelated form.

In present work, the uncorrelated form of the Data Set I is obtained using mathematical SPCA transformation while the Data Set II for testing is transformed into uncorrelated form using SPCA transforming neural net *Net* I_{SPCA}.

Further details of MLPs and ANNs can be found in Ref. [17]. For training such a MLP, the Levenberg–Marquardt (LM) back propagation learning algorithm is a quite popular and suitable technique for e-nose applications [18].

2.5. Data transformation using SPCA

Following the procedures described in Sections 2.2 and 2.3 the original data (Data Set I) represented by matrix $\boldsymbol{Y}$, carrying 42 observations of the gas sensor array having four sensing elements for four gases viz. acetone, carbon tetra-chloride, ethyl methyl ketone and xylene was accordingly first standardized to obtain $\boldsymbol{X}$. This data was then transformed to obtain $\boldsymbol{X_{SPCA}}$ without any dimensionality reduction. The scattering diagram for response data of three sensor elements and for first three principal components of SPCA transformed data $\boldsymbol{X_{SPCA}}$ are shown in Fig. 6(a) and (b).

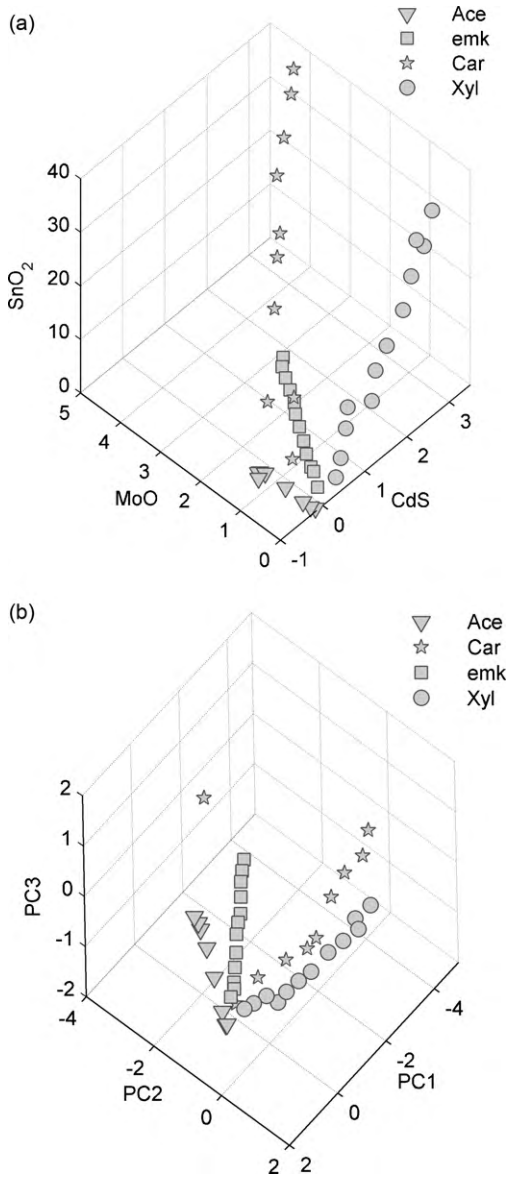

Fig. 6. Scatter plot of (a) raw sensor response in Data Set I (b) SPCA Transformed Data for Data Set I.

2.6. The SPCA transforming neural network

Observe that if a new sample's signal generated from the sensor array was obtained, it cannot be transformed to conforming SPCA equivalent as it is not a member of the SPCA set obtained earlier; then, such a freshly generated sample cannot be classified using a SPCA trained classifier.

For classification of such a sample, a neural network for SPCA transformation entitled *Net* I_{SPCA} is designed to transform the sensor data into equivalent SPCA version. For training *Net* I_{SPCA}, the training input was chosen to be the original data set (Data Set I) while the output target was chosen to be the SPCA transformed data ($\boldsymbol{X_{SPCA}}$).

The idea is that, the duly trained *Net* I_{SPCA} will receive the freshly generated data (not standardized) from the gas sensor array to translate it into compatible SPCA form to feed the classifier *Net* II_{SPCA}. For selecting model for *Net* I_{SPCA}, an eight-fold cross-validation using Data Set I was carried out for a three layer MLP taking log-sigmoid activation function for the neurons in the hidden layer and Purelin function for neuron of the output layer. The optimal number of neurons in the hidden layer of this three layer

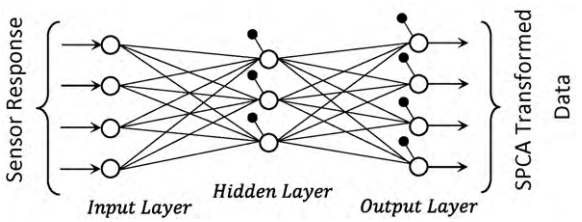

Fig. 7. Architecture of *Net* I_{SPCA}.

MLP was found to be 3. Accordingly, MLPs with 4 neurons in the input layer, 3 neurons in the hidden layer and 4 neurons in the output layer were developed. The ANN structure for *Net* I_{SPCA} is depicted in Fig. 7.

Data Set I (***Y***) was taken as input while $\boldsymbol{X_{SPCA}}$ was taken as target and the LM learning algorithm with performance function as MSE was chosen for the training of *Net* I_{SPCA}. Multiple networks using MATLAB® ver. 7.0.1 were initialized with different sets of weights and biases with random numbers between 0 and 1 and for 100 epochs of training with MSE goal to be10^{-6}. For each network, for early stopping minimum gradient was chosen to be e^{-10}, else optimally trained set of weights and biases were obtained when the weight and bias update starts to raise MSE in five successive iterations. The scalar parameter δ was initialized with a value of 0.001, which was incremented or decremented by multiplying with 10 or 0.1, respectively, up to a maximum of 10^9. The learning rate (α_n) was multiplied with the negative of the gradient to determine the changes to the weights and biases. The training of the network stopped when minimum gradient of update descent is reached. Optimization of weights and biases had been the condition for early stopping, subject to, if the error in successive iterations typically began to rise, for a predefined number of iterations (five, as described earlier) and the set of weights and biases at the minimum of the error was retained. Each trained network was then exported and the best performing network was accepted as *Net* I_{SPCA}.

Once trained, *Net* I_{SPCA} was fed the *i*th *test vector* (i = 1–18) form Data Set II, one at a time, to obtain correspondingly transformed SPCA versions of the Data Set II. The scattering diagram for first three principal components of the SPCA transformed data for Data Set II (obtained using *Net* I_{SPCA}), and the SPCA transformed data for Data Set I (obtained mathematically) is presented in Fig. 8.

2.7. Development of the SPCA trained classifier

As already enumerated at Section 2.4, an ANN classifier can be designed either by taking the correlated (Data Set I, ***Y***) or the uncorrelated data (*Net* $\boldsymbol{X_{SPCA}}$) as training input and respective classes of the gases/odors as the training targets and using LM training algorithm. Accordingly, the classifier trained with original data is named as Net_{RAW} while *Net* II_{SPCA} is the classifier which is trained with $\boldsymbol{X_{SPCA}}$. For model selection, an eight-fold cross-validation with Data Set I was carried out for a three layer MLP with log-sigmoid activation function for the neurons in the hidden layer and Purelin function for the output neuron. The optimal number of neurons in the hidden layer of this three layer MLP was found to be 4 and 3, respectively, for Net_{RAW} and *Net* II_{SPCA}. Lastly using LM learning parameters, random initial weights and biases, as already enumerated in Section 2.6, multiple classifiers of both varieties were developed. These classifiers were then tested by classifying 18 independent test vectors of Data Set II and the best performing net for each category was accordingly chosen as Net_{RAW} and *Net* II_{SPCA}.

The schematic of Net_{RAW} and *Net* II_{SPCA} is presented at Fig. 9(a) and (b).

The MSE for 18 samples for Net_{RAW} was 8.827×10^{-2} with individual error in class prediction was as high as 0.24 and as low as 1.45×10^{-2}. In contrast, the MSE for 18 samples for *Net* II_{SPCA} was 1.22×10^{-9} while individual error in class prediction was as high as 1.45×10^{-3} and as low as 7.25×10^{-10}. The classification performance of the networks Net_{RAW} and *Net* II_{SPCA} for Data Set II is shown in Fig. 10(a) and (b).

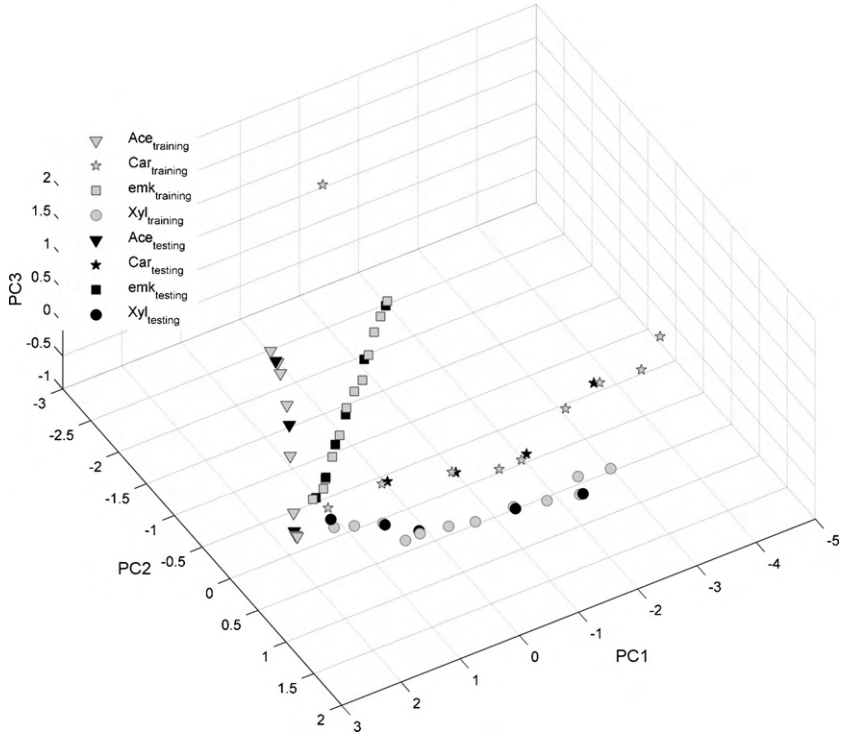

Fig. 8. Scatter plot of neurally transformed SPCA data for Data Set I and Data Set II.

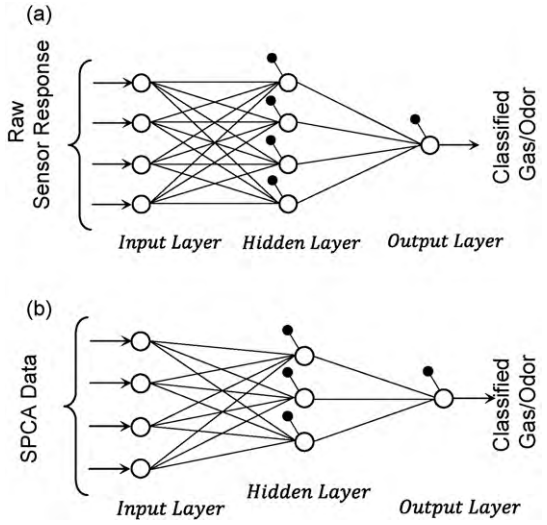

Fig. 9. Architecture of (a) Net_{RAW} (b) *Net* II_{SPCA}.

2.8. The proposed SPCA trained classifier

Once *Net* I_{SPCA} and *Net* II_{SPCA} had been developed and tested separately, they were now connected in succession to receive any freshly generated gas sensor response from the external environment. The overall classification performance of the proposed SPCA trained classifier for 18 independent samples from Data Set II is shown in Fig. 10(b). The final architecture of the SPCA trained classifier is presented in Fig. 11.

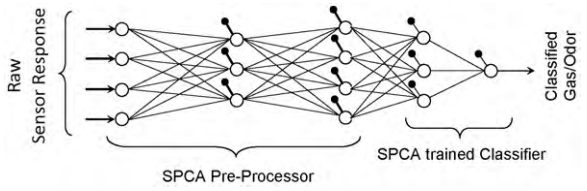

Fig. 11. ANN architecture of SPCA trained Classifier implemented using *Net* I_{SPCA} and *Net* II_{SPCA}.

3. Results and discussion

3.1. The statistical suitability of data transformation

In Section 2.6, a fresh data not belonging to Data Set I can be transformed into its corresponding SPCA equivalent using the mean and standard deviation of Data Set I is in accordance with the central limit theorem (CLT) [15]. It states, the conditions under which the

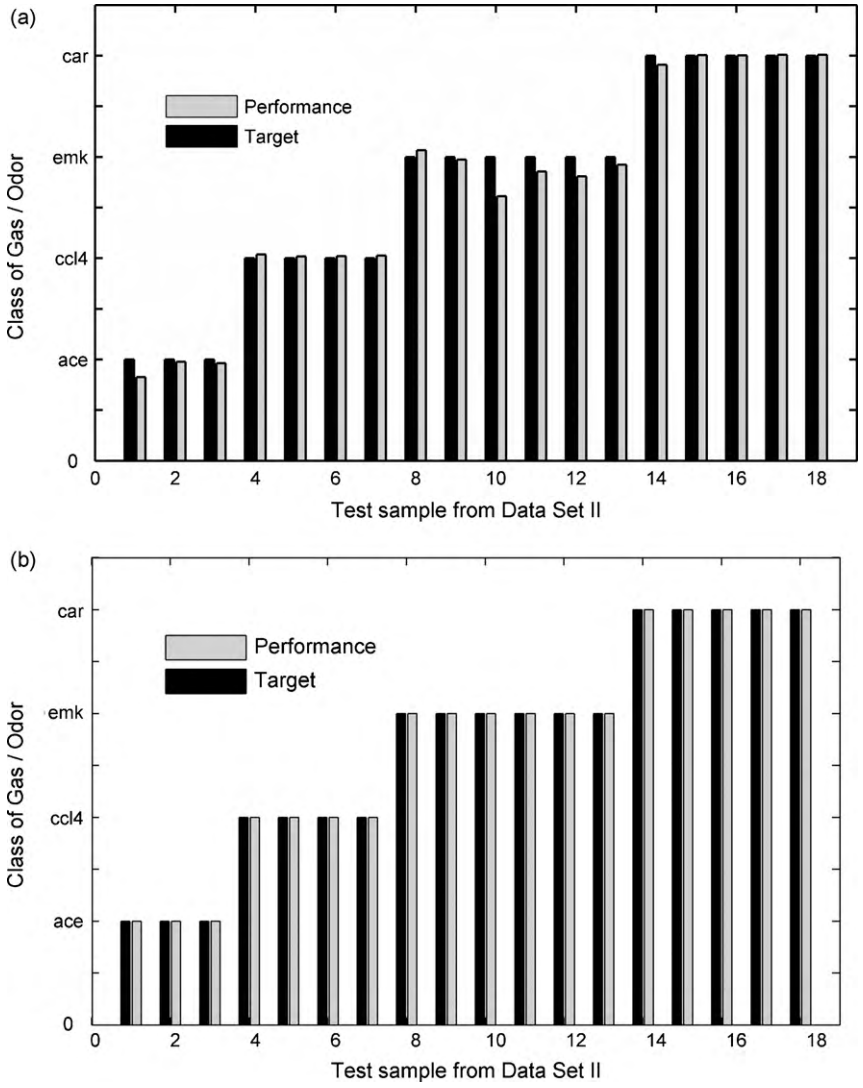

Fig. 10. Performance of (a) Net_{RAW} Classifier for test samples in Data Set II (b) *Net* II_{SPCA} Classifier for SPCA transformed test samples in Data Set II.

sum of a sufficiently large number of independent random variables, each with finite mean and variance, will be approximately normally distributed. Normal distribution is attained in our case for Data Set I (represented as matrix $\boldsymbol{Y}$) by its standardization with the column mean ($\bar{y}_j$, for all $j = 1, \ldots, m$) and standard deviation (s_j, for all $j = 1, \ldots, m$) which ensures that the centroid of the whole data set is now at origin of the new transformation yet the relative location of points remains same.

Accordingly, any new sensor response produced within the experimental sensor response range will be an integral entity of this distribution. It can, therefore, always be standardized w.r.t. the statistical characteristics of the Data Set I distribution viz. mean and standard deviation.

3.2. Performance of Net I_{SPCA} and Net II_{SPCA}

The neural network *Net* I_{SPCA} has been trained using data in Data Set I. It is used for obtaining SPCA equivalent of any freshly generated sensor response in conformity with the statistical distribution of Data Set I. The SPCA equivalent of the fresh sensor response can now be correctly classified using *Net* II_{SPCA}. Also, since the transformation has already been mapped onto the *Net* I_{SPCA}, the end user need not know the measures of the transformation with which the whole ANN system had been designed.

Net I_{SPCA} had been designed with the Data Set I containing 42 raw sample vectors.

To estimate the performance of *Net* I_{SPCA} and its conformity with the mathematically obtained SPCA transformation of the ith vector from Data Set II, the ith vector was processed using the statistical characteristics of the Data Set I distribution viz. mean, standard deviation and the feature vector ($\boldsymbol{FV}$). This process generated an 18×4 matrix ($\boldsymbol{T_{math}}$) carrying SPCA transformation of 18 test vectors present in Data Set II. Each test vector was also neurally transformed to obtain its respective SPCA using *Net* I_{SPCA}. Accordingly, for all the 18 test vectors, a neural SPCA transformed 18×4 matrix ($\boldsymbol{T_{neural}}$) was obtained. Elements of $\boldsymbol{T_{math}}$ and $\boldsymbol{T_{neural}}$ can be represented as $PCm_{i,j}$ and $PCn_{i,j}$, for $i = 1$–80 and $j = 1$–4, respectively. The mean squared error (*mse*) between respective vectors of $\boldsymbol{T_{math}}$ and $\boldsymbol{T_{neural}}$ is

$$mse_i = \frac{1}{4}\sum_{j=1}^{4}(PCm_{i,j} - PCn_{i,j})^2, \quad \text{for } i = 1, \ldots, 18. \tag{9}$$

The overall MSE for the 18 samples of $\boldsymbol{T_{math}}$ and $\boldsymbol{T_{neural}}$ is therefore

$$\text{MSE}_i = \frac{1}{18}\sum_{i=1}^{18} mse_i. \tag{10}$$

In the present work *mse* for individual test vectors were found to lie between 1.45×10^{-8} and 5.53×10^{-11} with an average of MSE = 1.81×10^{-9} for a set of 18 test samples.

A plot of *mse* for 18 vectors is presented at Fig. 12.

3.3. The overall classification performance with Net I_{SPCA} and Net II_{SPCA}

As shown in Fig. 11 a SPCA trained classifier with cascaded architecture using *Net* I_{SPCA} and *Net* II_{SPCA} offers a very good classification performance. All correct classification was obtained for 18 independent samples from Data Set II as is shown in Fig. 10(b).

The MSE for 18 samples for Net_{RAW} was 8.827×10^{-2} with individual error in class prediction was as high as 2.362×10^{-1} with a lowest error of 1.45×10^{-2}. In contrast, for the proposed classifier represented in Fig. 11 where *Net* I_{SPCA} and *Net* II_{SPCA} are cascaded together, the MSE for 18 samples was 1.22×10^{-9}. Individual error in class prediction was as low as 7.25×10^{-10} and the highest error observed was 1.45×10^{-9}. Thus proposed scheme in the work offers a clear advantage over the classifiers trained with raw sensor responses.

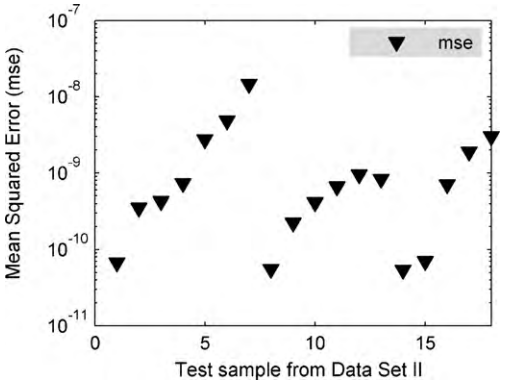

Fig. 12. Plot of *mse* between mathematically and neurally obtained SPCA transformation for Data Set II.

3.4. The accommodative ability due to the inclusion of Net I_{SPCA}

An e-nose comprising of SPCA transforming net essentially becomes accommodative to a newly generated gas sensor array response and it can classify it correctly gaining the advantages associated with SPCA trained classifiers.

4. Conclusions

The objective of this work was to demonstrate the development of an e-nose using feature enhancement technique such as SPCA for use in realistic situations. In the proposed work inclusion of a SPCA transforming neural net, i.e. *Net* I_{SPCA} preceding *Net* II_{SPCA} classifier has been suggested. It has been demonstrated that the classifier can now accommodate and classify correctly a freshly generated sensor response for the considered gases/odors. Interestingly this fresh sensor response was not a part of the training and validation data set however it is restricted within the experimental range.

Also, with this scheme, e-noses can be designed and trained under factory conditions for its use in practical situations. The end user of such an e-nose can straightaway start its use for classification of gases/odors in new ambience. The user need not worry about the statistical parameters used it its training. All he needs to know is the range of concentrations and the kinds of gases/odors for which it has been designed and developed by the manufacturer.

The results of classification has been all correct for the four gases we had considered and could be implemented using simpler network architectures owing to the use of feature enhancement technique such as SPCA, in our case. It has also been shown that this SPCA trained classifier shows a superior performance than the classifiers trained using raw sensor responses.

The universal nature of this methodology enables an e-nose designer to extend the system for any number of gases and odors.

Acknowledgement

The financial support in part from the Department of Science & Technology, Ministry of Science & Technology, Government of India (Grant no. R/Dev/P-07-313) is gratefully acknowledged.

References

[1] A. Bermak, S.B. Belhouari, M. Shi, D. Martinez, Pattern recognition techniques for odor discrimination in gas sensor array, Encyclopedia Sens. 10 (2006) 1–17.
[2] R. Gutierrez-Osuna, Pattern analysis for machine olfaction: a review, IEEE Sens. J. 2 (2002) 189–202.
[3] D.S. Lee, D.D. Lee, S.W. Ban, M. Lee, Y.T. Kim, SNO_2 gas sensing array for combustible and explosive gas leakage recognition, IEEE Sens. J. 2 (2002) 140–146.
[4] X. Liu, J. Hu, B. Cheng, H. Qin, M. Jiang, Acetone gas sensing properties of $SmFe_{1-x}$ Mg_xO_3 perovskite oxides, Sens. Actuators B 134 (2008) 483–487.
[5] G. Mehmetçik, G. Özdemirler, N.K. Toker, U. Çevikbaş, M. Uysal, Effect of pretreatment with artichoke extract on carbon tetrachloride-induced liver injury and oxidative stress, J. Exp. Toxicol. Pathol. 60 (2008) 475–480.
[6] R.R. Kotdawala, N. Kazantzis, R.W. Thompson, Molecular simulation studies of hydrogen cynide and methyl ethyl ketone on zeolite NaX and activated carbon, J. Hazard. Mater. 159 (2008) 169–176.
[7] M. Farhadian, D. Duchez, C. Vachelard, C. Larroche, Accurate quantitative determination of monoaromatic compounds for the monitoring of bioremediation processes, Bioresour. Technol. 100 (2009) 173–178.
[8] E. Garcia, PCA and SPCA Tutorial, Personal Publication, March 2008, available at http://www.miislita.com/information-retrieval-tutorial/pca-spca-tutorial.pdf.
[9] Z. Xua, X. Shia, L. Wangb, J. Luob, C.J. Zhongb, S. Lua, Pattern recognition for sensor array signals using fuzzy ARTMAP, Sens. Actuators B 141 (2009) 458–464.
[10] T. Alizadeh, S. Zeynali, Electronic nose based on the polymer coated SAW sensors array for the warfare agent simulants classification, Sens. Actuators B 129 (2008) 412–423.
[11] U. Siripatrawan, Rapid differentiation between *E. coli* and *Salmonella typhimurium* using metal oxide sensors integrated with pattern recognition, Sens. Actuators B 133 (2008) 414–419.
[12] P. Lv, Z. Tang, G. Wei, J. Yu, Z. Huang, Recognizing indoor formaldehyde in binary gas mixtures with a micro gas sensor array and a neural network, Meas. Sci. Technol. 18 (2007) 2997–3004.
[13] N.H. Beltrán, M.A.D. Mermoud, V.A. Soto, S.A. Salah, M.A. Bustos, Chilean wines classification based only on aroma information, Int. J. Intell. Syst. Technol. Winter (2006) 85–90.
[14] J.M. Saleh, B.S. Hoyle, Improved neural network performance using principal component analysis on Matlab, Int. J. Comput. Internet Manage. 16 (May–August (2)) (2008) 1–8.
[15] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, John Wiley & Sons, 2001.
[16] K.K. Shukla, R.R. Das, R. Dwivedi, Adaptive resonance neural classifier for identification of gases/odors using an integrated sensor array, Sens. Actuators B 50 (1998) 194–203, and references therein.
[17] S. Haykin, Neural Network: A Comprehensive Foundation, Macmillan, New York, 1994.
[18] B.G. Kermani, S.S. Schiffman, H.T. Nagle, Performance of the Levenberg–Marquardt neural network training method in electronic nose applications, Sens. Actuators B 110 (2005) 13–22.

Biographies

N.S. Rajput received the M. Eng. degree in communication systems in 1997. He is pursuing for Ph.D. degree from the Department of Electronics Engineering, Institute of Technology, Banaras Hindu University (IT-BHU), Varanasi, India. He is presently working as an assistant professor in the Department of Electronics Engineering, IT-BHU. His research interests include neural networks and PR techniques.

R.R. Das received the M.Sc. Eng. degree in electronics in 1975 and the Ph.D. degree in 1999 from the Institute of Technology, Banaras Hindu University (IT-BHU), Varanasi, India. He is presently working as a professor in the Department of Electronics Engineering, IT-BHU. His research interests include neural networks and digital systems.

V.N. Mishra received the Ph.D. degree in electronics engineering in 1998 from the University of Roorkee, Roorkee, India. He is presently working as reader in the Department of Electronics Engineering, IT-BHU. His research interests include design and fabrication if IC compatible gas sensors, neural networks, and PR techniques.

K.P. Singh received the Ph.D. degree in electronics engineering in 1973 from the Institute of Technology, Banaras Hindu University (IT-BHU), Varanasi, India. He is presently working as professor in the Department of Electronics Engineering, IT-BHU and as the Director, Institute of Technology, Banaras Hindu University (IT-BHU). He has over 105 research publications in various international/national journals. His research interests include microwave remote sensing, plasma physics, electromagnetic wave interactions with various media, communication engineering and neural network applications.

R. Dwivedi received the Ph.D. degree in electronics engineering in 1978 from the Institute of Technology, Banaras Hindu University (IT-BHU), Varanasi, India. He is presently working as reader in the Department of Electronics Engineering, IT-BHU. He has over 150 research publications in various international/national journals and in the proceedings of symposia. His research interests include silicon-based and thick-film sensors, MOS devices photovoltaics, and LSI/VLSI.