A Reduced Complexity Two-Stage Neural Classifier for Quantified Identification of Gases/Odors Using an Integrated Thick Film Sensor Array

N. S. Rajput¹, V. N. Mishra² and K. P. Singh³

Department of Electronics Engineering,
Institute of Technology,
Banaras Hindu University,
Varanasi (U.P.) 221005 INdia
¹nsrajput.ece@itbhu.ac.in, ²vnmishra.ece@itbhu.ac.in, ³kpsingh.ece@itbhu.ac.in

Abstract

In this paper, a reduced complexity two-stage feed-forward neural network has been employed for on-line classification and quantification of gases/odors. At the classification stage, sensor response for an unknown gas/odor sample is processed to identify its class. Now, at the quantification stage, concentration of this sample is predicted by an expert neural network, dedicatedly trained for quantification of that kind of gas/odor. Experimentally, a total of 21 neurons (4 neurons as the input nodes, 5 neurons in the classification network and 12 neurons in the quantification network) were used considering Thick-film Tin-Oxide sensor array responses for four gases/odors (viz. Acetone, Carbon Tetra-chloride, Ethyl Methyl Ketone and Xylene). It is reported that a reduced complexity two-stage ANN, designed following the proposed scheme, can reproducibly discriminate varieties of gases/odors. Classification of test samples has been ‘all correct’ while a very low mean squared error (0.0115 only) in the predicted concentration was observed.

Keywords: Gas/odor Identification, Intelligent Gas Sensors, Multistage Neural Network Architecture, Two-stage Neural Classifier

1. INTRODUCTION

For making fine discrimination amongst complex odors, mammalian olfactory systems make use of feature detection using broadly tuned receptor cells organized in a convergent neuronal pathway, without needing the use of highly specific receptors and exhibit high sensitivity for odors and high discrimination between them [1]. The traditional statistical techniques for classification and quantification have already been studied. The demerits of various traditional pattern recognition techniques like Partial Model Building, Fourier Transform Techniques, Cluster Method, Transformed Cluster Analysis, Multiple Regression Method, and Discriminant Function Analysis etc. have already been reported [2]. There is a considerable interest in the recent literature in developing an intelligent gas sensor (IGS) using semiconductor sensor array and pattern recognition (PR) (or discrimination) technology [3]–[5]. Artificial neural networks (ANNs) are now being used efficiently to analyze the vectorial signals of an IGS for quantified identification of gases [6]. D. Gnani et al. [7] have scaled a nine dimensional raw data into four dimensional manifold using a complex auto-associative
neural network comprising of 42 neurons in five layers, gaining better performance than respective principal component analysis (PCA) and projection pursuit network (PPN) methods for data information compaction, and they finally trained a neural network for IGS using second-order scaled conjugate gradient (SCG) method and refined through vanilla back-propagation algorithm. M. Pardo et al. [8] have gained enhancement through data pre-processing for classification of different brands of Espresso coffee with an electronic nose using PCA, up to 87.5% of classification performance, being the best than using Self Organizing maps (SOM). T. Roppel et al. [9] have established that data pre-processing using pulse coupled neural network (PCNN), rank order filtering (ROF) and centroid proximity metric (CPM) out-performs principal component analysis (PCA) for data pre-processing for neural quantifier in succeeding stage and have compared respective improvement in concentration identification up to 96% of correct identification. This is also recommended that while classifying a data, the process of classification be divided in two steps: super-class and sub-class, and then the super-class should be explored and this approach has been quoted as the 'spirit of hierarchical classification' [10], [11]. Better generalization ability of a two-stage network has also been utilized by Henkel et al. [12] for monitoring carbon dioxide and relative humidity but with step quantification for relative humidity in seven classes (from 0 to 60%) and gradual quantification of CO2 from 0 to 650 ppm facing variety of constraints and processed the data set for each stage with discordant reservations.

In recent literature, Acetone, Carbon Tetra-chloride, Ethyl Methyl Ketone and Xylene have been shown to score high hazard rankings among major industrial pollutants, considering its toxicity, persistence and exposure potential, evaluated through many parameters like Worker Exposure Hazard, Human Health Effect, Environmental Hazard Value and Total Hazard Value Score. Thereby, the need of accurately identified quantification of these gases using minimal complexity networks has gained importance for the development of real-time high-speed IGS for its use in hazard prone industrial situations [13] – [16].

Considering various aspects of previously published literature, and human and ecological health ranking of aforesaid industrial pollutants, we are motivated to attempt an n-stage neural network to achieve higher degree of quantified identification of the gases/odors with much simpler network architecture, involving only a few neurons and layers.

2. THE EXPERIMENT

To take up an experiment over the concept of effectiveness in ‘reduced-complexity over convergent neuronal pathway’, an n-stage neural network for quantified identification of individual gas/odor is considered. We have utilized the responses of sensor array for Acetone, Carbon Tetra-chloride, Ethyl Methyl Ketone and Xylene, as reported by Nayak et al. [17]. The sensor used in their experiments consisted of an array of four sensors with an integrated heater on a substrate. Four different sensors forming an array were fabricated with SnO2 as a base material and doped with different materials, namely, ZnO (sensor A), undoped SnO2 (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 reference [17]. The test rig used to collect sensor array response data and experimental conditions in the laboratory have also been described in the above 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 mentioned earlier, data used in the present analyses are derived from the results of [17], therefore, the sensor characteristics from [17] have been reproduced here for ready reference in Figure 1(a) to 1(d).

Using ‘Precision Image Digitizer ver. 1.3.0.0’, the graphs of Fig. 1(a) to 1(d) were imported and three categories of data were generated viz. Data set I (with 8 vectors for ace, 10 vectors for car, 12 vectors for emk and 12 vectors for xyl), Data set II (with 8 vectors for ace, 9 vectors for car, 7 vectors for emk and 8 vectors for xyl) and Data set III (with 3 vectors for ace, 4 vectors for car, 6 vectors for emk and 5 vectors for xyl). All of these vectors were uncorrelated and were generally uniformly spaced all over the concentration axis of the graphs at Figs. 1(a) to 1(d). From raw data, normalized data were developed by shifting the origin to match the graph’s origin and the data were scaled over the axes ranges of the graph, as the data points were in terms of respective pixel values of the software window. For all data points and graph windows, origin shifting and axes scaling was carried through (1), (2) and (3),

\[
(x, y) = (x', y') = (x_0, y_0) \ldots \quad (1),
\]

\[
\text{Conc.}(x_i) = \frac{x_i}{x_{\text{max}} - x_{\text{min}}} \cdot \text{X-axis concentration range} \quad \ldots (2)
\]

and

\[
\text{Response}(y_i) = \frac{y_i}{y_{\text{max}} - y_{\text{min}}} \cdot \text{Y-axis concentration range} \quad \ldots (3).
\]

Here, for all data pixels, \( x \) is the true pixel coordinate w.r.t. the origin of the respective graph, \((x', y')\) is the pixel coordinate w.r.t. the origin of the software window where the graph is imported and \( (x_0, y_0) \), is the origin of the actual graph while \((0,0)\) is always the origin of the window where the graph is imported for digitization.

\( x_{\text{max}} \) and \( y_{\text{max}} \) and \( x_{\text{min}} \) and \( y_{\text{min}} \) are the true maximum values of x-axis pixel and y-axis pixel, respectively. \( x_{\text{axis concentration range}} \) and \( y_{\text{axis concentration range}} \) are the represented as maximum ranges of percentage concentration along the x-axis and the maximum percentage change in resistance represented along the y-axis, respectively; taken appropriately for each gas/odor. \( \text{Conc.}(x_i) \) and \( \text{Response}(y_i) \) is the concentration of a particular gas/odor and the percentage change of sensor element resistance for the \( i \) th data point of the graph, plotted along x-axis and y-axis for each gas/odor and for each sensor element individually.

The normalized ‘sensor response vs. gas concentration’ data was hence intrinsically obtained for all the four gases/odors viz. ace, car, emk and xyl. These data were used for training and testing of the neural network thus conceived.

3. THE CONCEPT ANN

Inspired by the general architecture of Mammalian Olfactory System [1] and following the principles of odor coding [18] and the spirit of hierarchical classification [11], we divided the task of classified quantification into a two stage task wherein first stage should classify the data into its classes. Thence, the duly classified data from individual class should be quantified using respective quantifier trained for that class and the neural net used for this should be of simplest configuration in terms of perceptron.
Accordingly, Mean Square Error (mse) is taken to be the performance parameter as it is the average of squared error between the network output vectors $y(n)$ and the targeted output vectors $d(n)$. The mse is calculated as below:

$$
\text{mse} = \frac{1}{N} \sum_{n=1}^{N} e(n)^2 = \frac{1}{N} \sum_{n=1}^{N} (d(n) - y(n))^2
$$

... (4).

Here, $N$ is the total number of input vectors used in training/testing, while the $n$th training pattern. ‘mse’ is especially chosen as the performance parameter since this performance index is a quadratic function and will either have one global minimum or a weak minimum or even, no minimum, depending on the characteristics of the input vectors [19]. The gradient of the performance function should be used to determine the way to adjust the weights to minimize performance function and the weights are moved in the direction of the negative gradient, through the chain rule of calculus. Iteration used in this algorithm is written as:

$$
w_j(n+1) = w_j(n) - \alpha_n g_n \quad \ldots (5),
$$

where, $w_j$ is a vector of current weights and biases, $g_n$ is the current gradient, and $\alpha_n$ is the learning rate. For straightforward convergence, gradient descent is chosen and, is implemented in batch mode. Simpler yet faster training of the neural network is achieved using LM algorithm utilizing standard numerical optimization technique [20] to approach second-order training speed without having to compute the Hessian matrix [21].

4. THE CLASSIFYING NETWORK

The neural network was planned in two-stages, first of which was to classify data into the known categories of gases/odors (4 as in our case). A two-layer minimal-complexity neural classifier with only four neurons (viz. 5, 6, 7 and 8) with Hyperbolic Tan-sigmoid function in the hidden layer and only one neuron (viz. 9) with Linear function in the output layer is shown in Fig. 2.

Fig. 2. Classifier.

Data set I and Data set II were used to train the aforesaid classifier for known classes of gas/odor using LM algorithm with performance function as mse and the input vectors were taken without any modification as were obtained in sec. 4. Weights and biases of the classifier network were initialized with random numbers between 0 and 1 for 100 epochs with mse goal to be 0.0001. For early stopping minimum gradient was chosen to be $10^{-6}$, else

Fig. 3. Classifier’s performance over test data.

Data set I and Data set II were used to train the aforesaid classifier for known classes of gas/odor using LM algorithm with performance function as mse and the input vectors were taken without any modification as were obtained in sec. 4. Weights and biases of the classifier network were initialized with random numbers between 0 and 1 for 100 epochs with mse goal to be 0.0001. For early stopping minimum gradient was chosen to be $10^{-6}$, else
optimally trained set of weights and biases was obtained when the weight and bias update starts to raise mse in five successive iterations. The parameter $\eta$ 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 $\eta$ as in (5), is multiplied with the negative of the gradient to determine the changes to the weights and biases. For this network, when the minimum gradient of update descent is reached, the training stops. Optimality of weights and biases is the condition for early stopping, subject to, if the error in successive iterations typically begins 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 is retained. As previously described, the network is trained with LM algorithm [20].

In here, we have initialized the parameter $\eta$ with a value of 0.001. It is multiplied by 0.1, whenever the performance function is reduced by a step whereas It is multiplied by 10, whenever a step would increase the performance function. When $\eta$ becomes larger than $10^9$ the training is stopped.

Once, the network was trained, it was exported and named as Classifier_1. Multiples of such classifiers were created taking variety of training and learning criteria and each classifier was tested with the data set III for better performance. It has been observed that for the best performing classifier, the classifier could be trained in just 18 epochs giving 100% accurate performance for the data set III (consisting of 18 samples) when used as test data (Figure 3).

5. THE QUANTIFYING NETWORK

Once input data vectors were identified to be associated with a specific class of gas/odor, we attempted to quantify the concentration of that gas/odor through its neural quantifier. As the data had already been associated with its respective gas/odor, individual quantifiers for specific gas/odor were, then, designed with a two-layer minimal-complexity neural net involving only two neurons with Hyperbolic Tan-sigmoid function in the hidden layer and only one neuron with Linear function in the output layer [25] for each class of gas/odor viz. quantifier ace (10, 11, 18), quantifier car (12, 13, 19), quantifier emk (14, 15, 20) and quantifier xyl (16, 17, 21). Input vectors were fed to these, through the neurons 1, 2, 3 and 4 present in the input layer. Hierarchical classification is achieved having formed super and sub-class by partitioning and training the hidden and output layer, in such a way that the hidden neurons 10 and 11 feed the output neuron 18 and don’t feed neurons numbered 19, 20 and 21. Similarly, other neurons of the hidden layer were also fed selectively to their respective output neurons, as is depicted in Fig. 4.

Similar kind of training was attempted when aforesaid partitioning was not carried.
out and the network had simply failed to learn, emphasizing that a very simple network can work well if data fed to it was pre-classified at classifier stage. The quantifier’s network structure is depicted in Figure 4, for each gas. A series of quantifiers for respective gases/odors were developed by training aforesaid ‘elementary network’ for concentration quantifiers, with variety of training and learning parameters and having extracted respective training vectors from data set I and data set II. We have finally tested the quantifier’s performance with respective test data obtained from data set III. Here, the aforesaid quantifiers for known classes of gas/odor were trained using LM algorithm with performance function as ‘mse’ and the input vectors were taken without any modification as were obtained in sec. II. For the quantifier’s networks weights, biases and the parameter $\theta$ were initialized and updated in the similar manner as was done for the classifier’s network in Sec. 5. The Learning rate $\alpha$, as in eq. (5), is is multiplied with the negative of the gradient to determine the changes to the weights and biases. For the networks, when the minimum gradient of update descent is reached, the training stops.

Optimality of weights and biases is the condition for early stopping, subject to, only if the error in successive iterations typically begins 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 is retained. Once a network was trained, it was exported and named as quantifier_ace/car/emk/xyl_n. As a result, the best performing quantifier for respective gas/odor was identified. It has been observed that for the best performing quantifiers for their respective gases/odors, the quantifiers could be trained in just 8 to 12 epochs delivering mean-square-error of 0.0115 over accurate quantification for the data set III (consisting of 18 samples), when used as test data (Figure 5).

6. RESULT AND DISCUSSIONS

We have attempted quantified identification of four of the predominantly hazardous gases/odors with 100% correct identification while the quantification error for individual gas/odor has been found to have a mean squared error of 0.0115 only.

Fig. 5. Quantifier’s performance over test data.

This nominal mean square error of 0.0115 may be attributed to the exactitude in the characterization of the sensor at nano-level precision and the interpolations used while drawing those graphs obtained from our experimental data points and curve fitting.

In the present work, we have observed that the network structure used is universal in nature and is gas independent enabling one to extend it for any number of gases/odors with same network simplicity. The better accuracy has been seen with only 21 neurons of simpler transfer function and in no more than 3 layers including one input layer with four neurons. It has been demonstrated that with hierarchical design e.g. a 2-stage network, the network complexity has been found to be minimal and responds better than its consonantal networks reported earlier. We
have also processed our data twice. In the classifier stage, the classifier classified individual vectors of the data set and associated it with its constituent class. We have once again classified the data through the quantifier stage to obtain the gas/odor concentration. This has resulted into a definite inference that multiple networks can derive variety of information from the same set of data and this kind of multi-staged processing can keep individual network structure much simpler.

7. CONCLUSIONS

It has been possible for us to demonstrate through present research that the quantified identification of gases/odors may be carried very effectively and correctly, with very elementary neural networks, when quantified identification is carried in successions, following the ‘spirit of hierarchical classification’. It has been very encouraging that the success rate in identifying category of gases/odors is 100% and the mean-square-error for accurate quantification is as low as 0.0115 for all the four highly hazardous and persistently toxic gases viz. Acetone, Carbon Tetra-chloride, Ethyl Methyl Ketone and Xylene which highly affects the industrial workers as well as the ecology.

In the present work, we have developed system for four gases which can very well be extended to multiple gases/odors retaining its simplicity owing to its universal nature of extendible structure. Our result establishes that with very elementary neural net, complex gases present even in small concentrations, can well be identified utilizing hierarchically classified quantification.

ACKNOWLEDGMENT

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


