An Electronic Nose System Using Artificial Neural Networks with an Effective Initial Training Data Set

Bancha Charumporn* Non-member
Michifumi Yoshioka* Member
Sigeru Omatu* Member

Nowadays there are several commercial electrical noses (ENs) applied in many applications, mainly in food and cosmetics industries. Most of them have been added with complicated mechanisms to control the measuring environment. Consequently, they are large in size and expensive. However, the reliability of those ENs can be achieved only at moderate levels. Therefore, a simple EN system with an effective method to analyze the data is proposed as an alternative way for classifying smells. The EN has not been added with a mechanism to control the measuring environment. Thus, the EN system is inexpensive, small and can be operated easily. However, a normalization method need to be utilized to reduce the effect of measuring environment. Then a method to select the representative training data for artificial neural networks (ANNs) based on a similarity index (SI) value is applied to reduce the training time. The results show the ability of the EN that is able to classify not only different kinds of smoke but also the same kind of smoke from different brands and different concentration levels quite precisely.

Keywords: Electronic nose, ANN, Similarity index, Effective training data, Smoke

1. Introduction

The conventional method for analyzing the smell in the industry generally uses a human expert or a gas chromatography (GC). The disadvantages of a human nose are poor reproducibility and low reliability. A human nose is also not able to test the hazardous odors. The GC needs an expert to analyze the complicated signals. Therefore, the artificial olfactory devices called electronic nose (EN) have been developed widely in many organizations [1]~[8] for the last two decades. Since early 1995, several commercial ENs have been launched to the market, mainly applied in food, drink, and cosmetic industries [2]. The main components of the ENs are the sensing devices that mostly use metal oxide gas sensors (MOGSs). Since an MOGS is sensitive to temperature and humidity, some commercial ENs have been added with complicated mechanisms to overcome this problem. Consequently, those ENs are large in size and expensive. Although those ENs have good systems to control measuring environment, their reliability can be achieved only moderate levels. Thus, a simple EN system with an effective method to analyze the data that have been affected by the measuring environment is proposed as an alternative olfactory analyzing system. The EN has not been added with a mechanism to control the measuring environment. Thus, this EN system is cheap, small,

and can be operated easily. However, a normalization method to handle the noisy data needs to be applied in order to obtain good classification rates. Then the data selection method based on the similarity index (SI) value is applied to reduce the training time for the artificial neural networks (ANNs). The EN has been successfully applied to classify three kinds of smoke, joss stick, cigarette, and mosquito coil by using an error back propagation neural network (BPNN) [9]. In this paper, the EN system is applied to classify not only different kinds of smoke, but also the same kind of smoke from different brands and different concentration levels.

Since the EN does not have a mechanism to control the measuring environment, the signals of the same smell measured under different weather environments are varied. Without a proper method to handle these data, a good result may not be obtained. Thus, all experimental data from the same data set that have been measured in the same day under a similar weather environment, are normalized in order to reduce the effect of measuring environment. After that a method to select a representative training data set based on the SI value is applied. All similar data are grouped together and the data that has the highest average SI value to the other data in its groups is used as a representative training data for the ANNs. Two kinds of ANNs, the BPNN and a learning vector quantization (LVQ) methods, are applied to classify these data. The results from both the BPNN and the LVQ method show the ability of the EN system to classify not only different kinds of smokes but also the same smoke from different brands and different concentration levels accurately by using a small number of data

* Division of Computer and Systems Sciences,
 Graduate School of Engineering,
 Osaka Prefecture University,
 Sakai, Osaka 599-8531, Japan.
 E-mail: bancha@sig.cs.osakafu-u.ac.jp

for training. This can prove that a simple EN with an effective method to handle the data can also be used as a reliable olfactory analyzing system.

The rest of the paper is organized as follows. Section **2.** explains the EN system. The experiment and data collection are briefly explained in Section **3.** Section **4.** shows a method to reduce the effect of measuring environment by the normalization method. The method for selecting the representative training data set for ANNs is explained in Section **5.** The BPNN and the LVQ method are briefly described in Section **6.** The results are discussed in Section **7.** Finally, the conclusion for this paper is summarized in Section **8.**

2. Electronic Nose System

The EN system shown in Fig. 1 has been designed based on the concept of a human olfactory system. Various kinds of MOGSs from the FIS Inc. listed in Table 1 are used as olfactory receptors. The main part of an MOGS is the metal oxide element on the surface of the sensor. When this element is heated at a certain high temperature, the oxygen is absorbed on the crystal surface with the negative charge. The chemical reaction between the negative charge of the metal oxide surface and deoxidizing gas makes the resistance of the sensor vary as the partial pressure of oxygen changes. Based
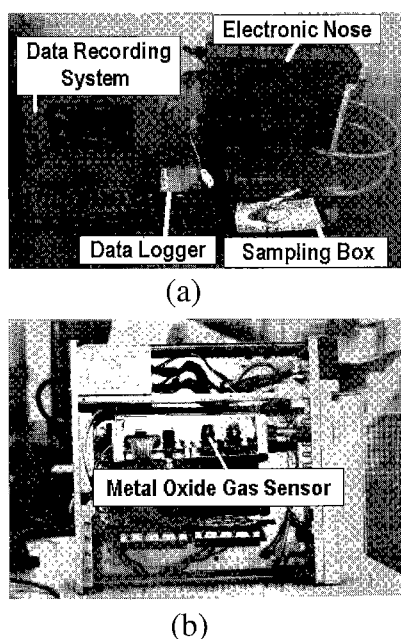

(a)

(b)

Fig. 1. (a) The electronic nose system; (b) Inside of the EN containing various metal oxide gas sensors.

Table 1. List of MOGSs from the FIS Inc. used in this experiment.

Sensor Model	Main Detecting Gas
SP-53	Ammonia, Ethanol
SP-MW0	Alcohol, Hydrogen
SP-32	Alcohol
SP-42A	Freon
SP-31	Hydrocarbon
SP-19	Hydrogen
SP-11	Methane, Hydrocarbon

on this characteristic, we can measure the total voltage changes during the sensors absorb the tested odor. After that the data logger is used to convert the signals from analog to digital before recording into the computer. Finally, the ANNs that imitate the function of human brain are applied to analyze these data.

In the EN, the oxygen is simply provided by the normal air at room temperature. After measuring each odor, the MOGSs need to be cleaned by flowing the clean air into the EN just like a human nose which need to breathe fresh air before being able to recognize a new smell accurately. More information about this EN is provided elsewhere [9].

3. Experimental Data Collection

The EN is applied to test the smell of three kinds of general household burning materials. In order to evaluate the efficiency of the EN, three brands of joss stick are tested with three concentration levels. Each data set contains eleven data listed in Table 2. In order to observe the effect of weather environments, forty data sets were measured at different locations under different weather environments during winter season and summer season of the year 2002 and another twenty data sets were measured in the winter season of the following year. However, all eleven data in the same data set were measured in the same day under a similar environment.

For each data, the voltage signals of the normal air from all MOGSs are measured every second for one minute and their average value ($\bar{v}_{air,i}$) is used as air reference point. After that the tested burning material is put into the sampling box, then the voltage signals of the sensors when absorbing the tested smoke, $v_{smoke,t,i}$, are collected every two seconds for a period of two minutes. Finally, the total change in signals on each period, $V_{smoke,t,i}$ are calculated by

$$V_{smoke,t,i} = v_{smoke,t,i} - \bar{v}_{air,i} \cdots\cdots\cdots\cdots\cdots (1)$$

where t is the period from 1 to 60 and i denotes the sensor number 1 to 7.

After removing the tested smoke from the sampling box, the voltage signals of all sensors are gradually decreased. The voltage signals of all sensors must be returned to the constant level before being able to test a new smoke sample. The voltage signals of four sample data from different data sets are plotted together in

Table 2. List of burning materials in the experiment.

Burning Material	Type or Brand Name	Concentration Level	Symbol
Joss stick	Purple	3sticks	P3
		2sticks	P2
		1stick	P1
	Brown	3sticks	B3
		2sticks	B2
		1stick	B1
	Green	3sticks	G3
		2sticks	G2
		1stick	G1
Cigarette	Marlboro Light	1 cigarette	ML
Mosquito coil	Kincho	1 coil	MC

Green Joss Stick 1 Stick

Brown Joss Stick 1 Stick

Marlboro Light Cigarette

Mosquitio Coil

SP-53		SP-MW0	
SP-32		SP-42A	
SP-31		SP-19	
SP-11			

Fig. 2. Sample data from this experiment.

without a proper modification. As mentioned in Section 3, each data of a same data set is measured consecutively under a similar environment. Therefore, we assume that all of the data in the same data set have the same environmental effect. If we normalize these data, it should be able to reduce the effect of measuring environment. Thus, each data set is normalized by

$$norm(V_{smoke,t,i}) = \frac{v_{smoke,t,i} - \bar{V}_i}{\sigma_{v,i}} \ \dots\dots\dots (2)$$

where $\bar{V}_i$ and $\sigma_{v,i}$ are the mean and standard deviation of all data in each data set from sensor i, respectively, t is the period from 1 to 60, and i denotes sensor number 1 to 7.

In order to verify the effect of the normalizing method, the principal component analysis (PCA) is applied to transform the experimental data. The voltage signals from Eqs.(1)and (2) are used as the input data of the PCA for the data *before* and *after* normalizing cases, respectively. The input data size is equal to 420 which is equal to the measuring periods, 60, times the number of sensors, 7. Two main components of the PCA is plotted in Fig. 3. The PCA of data *before* normalizing has a lot of mixing zones between ML, MC, and the joss stick data measured in different year and all data are dispersedly distributed. In case of data *after* normalizing, data are more clearly separated into three main categories and all data are distributed densely in their own areas. Since two main components from the PCA of data *after* normalizing case can only represent 89% of accumulated variance, it cannot clearly identify different kinds of joss stick data. However, it can show that each kind of data
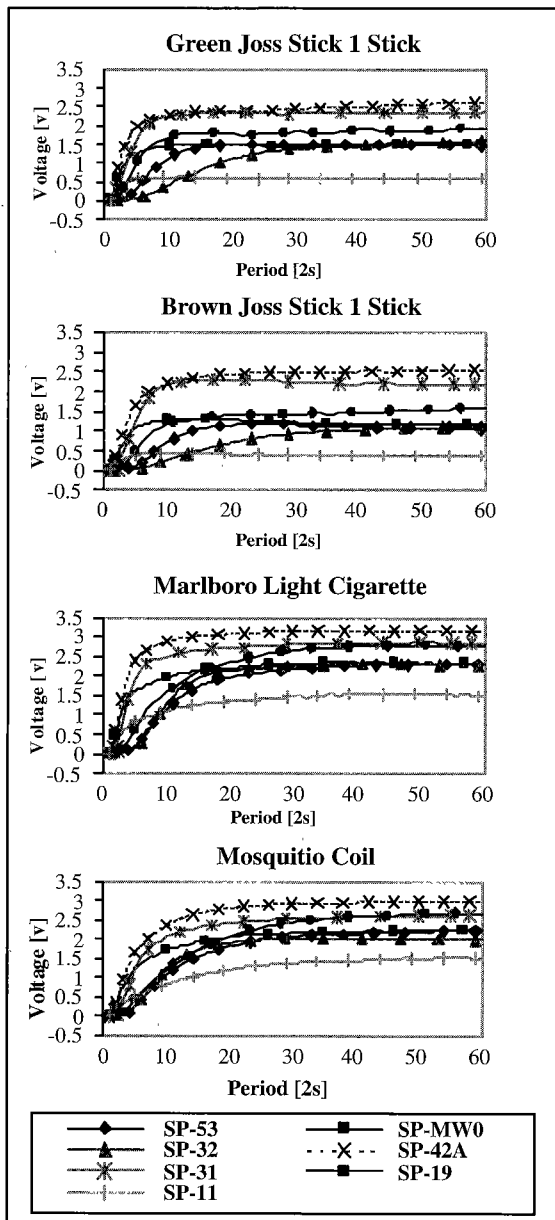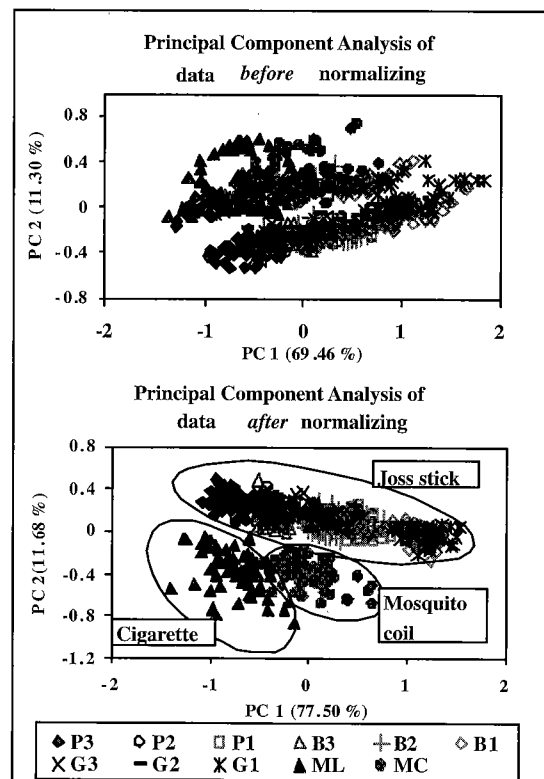
Fig. 2. Since MOGSs are sensitive to temperature and humidity, the signals from the Marlboro Light cigarette measured in an environment are similar to those of the mosquito coil measured under another environment. The signals from all kinds of joss stick are also similar due to the similar ingredients of these joss sticks, as shown by the sample data from the brown joss stick and the green joss stick in Fig. 2. Theoretically, the ANNs can be applied to analyze these data directly. However, the result may not yield a good classification rate. Therefore, a proper method to reduce the effect of measuring environment is necessary before training and testing these data by the ANNs.

4. Reducing Environment Effect

The sample data in Fig. 2 show the similarity of different kinds of smoke measured under different environments. These data are difficult to be classified correctly

| ◆ P3 | ◇ P2 | ☐ P1 | △ B3 | ╋ B2 | ◇ B1 |
| ✕ G3 | ━ G2 | ✖ G1 | ▲ ML | ✹ MC | |

Fig. 3. PCA of the data *before* and *after* normalizing.

1640

IEEJ Trans. EIS, Vol.123, No.9, 2003

are distributed densely in its own area. Thus, it is possible to use the ANNs to classify these data into specific categories as listed in Table 2.

5. Training Data Selection Method

5.1 Similarity Index (SI) In the statistical application, the correlation value developed mainly by Karl Pearson is widely used to find the relationship between two random variables [10] [11]. In this paper, we have applied this correlation value to measure the similarity of the training data set in order to group them together and find the effective representative training data set for the ANNs. We call this correlation value as a similarity index (SI). The SI value varies from -1 to $+1$. Two random variables with an SI of either -1 or $+1$ are highly correlated because knowledge of one provides precise knowledge of the other. However, the SI provides information only about linear relationships between random variables. Therefore, we assume that each data pattern has nearly linear relationship to the other data patterns. The similarity index matrix (**SIM**) of all training data is calculated by the following equation.

$$\mathbf{SIM} = \begin{bmatrix} r_{11} & \cdots & r_{1m} \\ r_{21} & \cdots & r_{2m} \\ \vdots & \cdots & \vdots \\ r_{m1} & \cdots & r_{mm} \end{bmatrix},$$

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad \cdots\cdots\cdots (3)$$

where x and y are the comparing data, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, n is the size of each data which equals 420 (60 periods $\times$ 7 sensors), and m is the number of training data which is equal to 30 (half of the experimental data).

In order to show the relationship between the SI value and the linear relationship of the data in this experiment, the voltage signals of all 420 data points from two data of P3 are plotted in Fig. 4. The SI value between these two data calculated by Eq.(3) is 0.998 and the scatter plot of these two data is almost a straight line. If data 1 or data 2 is added as the training data, the other data will be correctly classified. Thus, it is possible to apply the SI value as a guideline to select a representative training data set.

5.2 Training Data Selection Algorithm Half of the experimental data is randomly selected as a training data set, S_c. The training data set may contain a lot of similar data. Thus, a process to group the similar data base on their SI value is applied. Then the data that has the highest average value to all data in its group will be added into the representative training data set, S_r, and the rest of the data in its group will not be used for training.

Step1: Each data type in Table 2 contains 60 data. Thus, the number of training data set S_c is equal to 30.
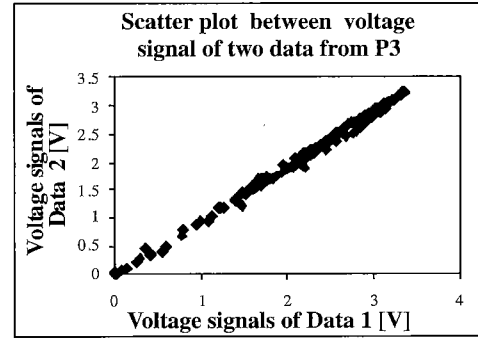

Fig. 4. Scatter plot of two data that has a high SI value.

We define $m = 30$ as the number of training data in S_c and Initialize the number of data in S_c and S_r equal to m and 0, respectively. Calculate $m \times m$ size **SIM** using Eq.(3).

Step2: From **SIM**, calculate the mean of the SI for each data, using Eq.(4).

$$\bar{r_i} = \frac{\sum_{j=1}^{m} r_{ij}}{m}. \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots (4)$$

Step3: Compute the overall mean, $\tilde{r}$, and a standard deviation, σ_r using Eqs.(5) and (6), respectively.

$$\tilde{r} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{m} r_{ij}}{m^2}. \quad \cdots\cdots\cdots\cdots\cdots\cdots (5)$$

$$\sigma_r = \frac{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{m}(r_{ij} - \tilde{r})^2}}{m}. \quad \cdots\cdots\cdots\cdots (6)$$

Step4: Find the upper threshold, T_U, and the lower threshold, T_L, using Eqs.(7) and (8), respectively.

$$T_U = \tilde{r} + \alpha\sigma_r \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (7)$$

$$T_L = \tilde{r} + \beta\sigma_r \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (8)$$

where α, β is the real value, $\alpha > \beta$, $T_L < T_U < 1$.

Step5: For all $i, j \in S_c$ and $i \neq j$, count the SI of data i by $Q_i = number\ of\ \{r_{ij}|r_{ij} > T\}$.

Step6: Group all data j that have $r_{kj} > T_U$ with data k, where $\bar{r_k} = max_i[\bar{r_i}]$ and $Q_k \geq n$. Here, n is the minimum group that is worthwhile to form a group, i.e., n higher than 2. Data k will be put into S_r as the representative training data. Data k and all data j related to data k will be moved out from S_c. Thus, the member of data in S_r is increased by 1 and the number of data in S_c is decreased by $1 + Q_k$.

Step7: Repeat Steps 5 to 6 until $Q_k < n$.

Step8: The data left in S_c will be grouped into two groups. For all data $j \in S_c$ and $k \in S_r$, move all data j, which have all $r_{jk} < T_L$ into S_r. These data are the data that have low correlation with the representative data. Therefore it is necessary to use them for training.

All of the remaining data in S_c will be verified after the first training process.

Steps 1 to 8 will be repeated for all data type listed in Table 2 before starting the first training process.

Step9: Begin the first training process by the ANNs. After that, test the remaining data in S_c. All misclassified data in S_c will be added into S_r. Then the final training process is continued until the termination condition such as the mean squared error (MSE), or the number of iteration, is reached.

6. Neural Networks Analysis

6.1 Error Back Propagation (BP) Method
A three-layered neural network consists of input, hidden, and output layers. The neuron i of the input layer is connected to the neuron j of the hidden layer with the weight v_{ij}, and the neuron j of the hidden layer is connected to the neuron k of the output layer with the weight w_{jk}. A reference pattern is given at the output layer to compare with the value from output layer, and the connecting weights are repeatedly adjusted until the MSE is minimized.

Let o_k denote the output reference value of the kth neuron of the output layer and let x_i, z_j, y_k denote the output of ith, j^{th}, k^{th} neuron in input, hidden, and output layers, respectively. The MSE to be minimized is given by

$$\text{MSE} = \frac{\sqrt{\sum_{k=1}^{m} (o_k - y_k)^2}}{M} \quad \cdots\cdots\cdots\cdots\cdots (9)$$

where M is the number of output node. The connection weights w_{jk} at iteration t is incremented by

$$\triangle w_{jk}(t+1) = \alpha \delta_k z_j + \mu \triangle w_{jk}(t) \cdots\cdots\cdots\cdots (10)$$

where $\alpha > 0$ is the learning rate, and $0 < \mu < 1$ is the momentum rate, and $\triangle w(t) := w(t) - w(t-1)$, $\delta_k := (o_k - y_k)y_k(1 - y_k)$.

This process is propagated back to the input layer and the weights v_{ij} at iteration t is updated by

$$\triangle v_{ij}(t+1) = \alpha \delta_j x_i + \mu \triangle v_{ij}(t) \cdots\cdots\cdots\cdots (11)$$

where

$$\triangle v(t) := v(t) - v(t-1), \; \delta_j := \sum_{k=1} \delta_k w_{jk} z_j (1 - z_j).$$

6.2 Leaning Vector Quantization (LVQ) Method
The Euclidean distance between the input pattern vector, $\mathbf{x} = [x_1, x_2,, x_n]$ and the reference vectors, $\mathbf{w_j} = [w_{ij}, w_{2j},, w_{nj}]$ is calculated and the two winner reference vectors, $\mathbf{w_{c1}}$ and $\mathbf{w_{c2}}$, which are the nearest neighbors to $\mathbf{x}$ are modified. However, the two winner units must satisfy the following conditions: First, one of these units must belong to the right category and the other is in a wrong category. Second, these two reference vectors must be in the window range that is measured by

$$\left(\frac{d_{c1}}{d_{c2}}, \frac{d_{c2}}{d_{c1}}\right) > S, \quad S = \frac{1-\varepsilon}{1+\varepsilon}$$

where d_{c1} and d_{c2} are the Euclidean distance from $\mathbf{x}$ to $\mathbf{w_{c1}}$ and $\mathbf{w_{c2}}$, respectively and ε is the window parameter.

When these conditions are met, the reference vector that belongs to the same category as $\mathbf{x}$, let say $\mathbf{w_{c1}}$, is updated at t iteration according

$$\mathbf{w_{c1}}(t+1) = \mathbf{w_{c1}}(t) + \alpha(t)(\mathbf{x} - \mathbf{w_{c1}}(t)) \quad \cdots\cdots (12)$$

and the reference vector that does not belong to the same category as $\mathbf{x}$ is updated by

$$\mathbf{w_{c2}}(t+1) = \mathbf{w_{c2}}(t) - \alpha(t)(\mathbf{x} - \mathbf{w_{c2}}(t)) \quad \cdots\cdots (13)$$

where $\alpha(t)$ is the training rate which continually decreases every iteration by $\alpha(t) = \alpha_0(1 - t/t_{\max})$, α_0 is a positive constant.

The updating process continues until $t_{\max}$ iteration

7. Classification Results and Discussion

The experimental data are randomly divided into two main groups called Group 1 and Group 2. One group will be used as the training data set and the other group will be used as the unseen testing data set. Since the input data size in this application is huge and the training data set may contain a lot of similar data, a method to select the training data explained in Section 5 is applied to choose the representative training data and the BPNN and the LVQ are applied to classify the data.

7.1 Result from the BPNN The structure of a BPNN in this application as shown in Fig. 5 has three layers. The input layer consists of four hundred and twenty input neurons, which equals to the number of measuring periods (60) times the number of sensors (7). For the hidden layer, we have tried with many values and finally thirty-five hidden neurons are the best value that shows a good accuracy with a reasonable training time. The output layer contains eleven output neurons. Each output neuron presents each data type in Table 2. The best training parameters getting from a trial and error method are, a learning rate, $\alpha = 0.1$ and a momentum rate, $\mu = 0.001$.

After the first training using the initial training data from Steps 1 to 8 in Section 5.2, the remaining data in the training data set, S_c, are tested. We assume that a pattern is classified correctly if the output node is greater than or equal to 0.6 while the output reference
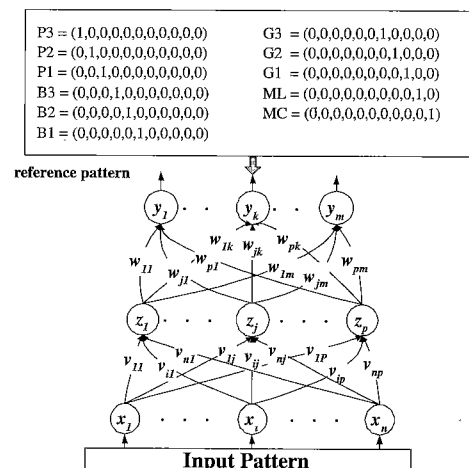

Fig. 5. Structure of the BP.

1642

IEEJ Trans. EIS, Vol.123, No.9, 2003

Table 3. Classification results of the experimental data using BPNN.

Data	No. of Training Data	Correct	Rate (%)	Misclassified	No. of Training Data	Correct	Rate (%)	Misclassified
	Group1 train, Group2 test				Group2 train, Group1 test			
P3	5	29/30	97		6	29/30	97	P2=1
P2	9	25/30	83	B3=1, B2=1	12	29/30	97	
P1	9	30/30	100		9	28/30	93	B2=1
B3	10	25/30	83	B2=1, G3 =1	13	27/30	90	P2=1
B2	12	28/30	93	B3=1, G2=1	12	27/30	90	G2=1
B1	8	28/30	93		10	27/30	90	G1=3
G3	7	27/30	90		9	27/30	90	B3=1
G2	7	28/30	93	B2=1	9	29/30	97	
G1	8	28/30	93	G2=2	8	30/30	100	
ML	7	29/30	97		8	30/30	100	
MC	6	30/30	100		8	29/30	97	

Table 4. Classification results of the experimental data using LVQ.

Data	No. of Training Data	Correct	Rate (%)	Misclassified	No. of Training Data	Correct	Rate (%)	Misclassified
	Group1 train, Group2 test				Group2 train, Group1 test			
P3	5	30/30	100		4	30/30	100	
P2	12	24/30	80	P3=1, B3=3,B2=2	11	30/30	100	
P1	8	29/30	97	G2=1	6	27/30	90	B2=2, B1=1
B3	13	23/30	77	P3=4, B2=1, G3 =2	11	24/30	80	P2=5, G3=1
B2	11	25/30	83	P1=2, B3=1,G2=2	9	24/30	80	G2=6
B1	8	25/30	83	B2=1, G1=4	6	29/30	97	G1=1
G3	11	26/30	87	B3=3, G2=1	8	25/30	83	B3=4, B2=1
G2	11	26/30	87	B2=4	10	27/30	90	B2=3
G1	8	27/30	90	B1=3	6	29/30	97	B1=1
ML	6	30/30	100		5	30/30	100	
MC	6	29/30	97	ML=1	5	30/30	100	

value is equal to 1, and the other output nodes are less than or equal to 0.3 while the output reference values equal 0. The misclassified data that are the data near the class boundaries, are added into the training data set and the networks are retrained again. The BPNN is trained until the average mean square error (MSE) of all training data is less than 0.0005. After that the BPNN is tested by the unseen testing data set. The final results using only a small number of training data are shown in Table 3.

7.2 Result from the LVQ method The LVQ structure in this application as shown in Fig. 6 consists of two layers. The input layer contains four hundred and twenty two nodes same as the input layer of the BPNN. The initial competitive layer when using Group 1 and Group 2 for training contains sixty six and sixty three nodes, respectively. This numbers are equal to the total number of initial training data obtained by Steps 1 to 8 of the algorithm in Section 5.2. The initial training rate, α, and the window parameter, ε, equal to 0.1 and 0.35, respectively. The LVQ was trained for 1000 iterations.

After completing the first training, the remaining data in the training data set, S_c, are tested. Some data that was classified into the wrong quantized region are added into the representative training data set. The number of nodes in the competitive layer is increased if both winner reference vectors of the new training data fall into the wrong categories. The final competitive layer size when using Group 1 and Group 2 for training is equal to seventy three and sixty eight, respectively. The final result using the LVQ method are shown in Table 4.
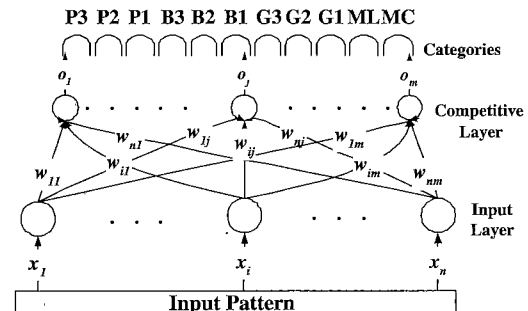

Fig. 6. Structure of the LVQ.

7.3 Discussion We have described the ability of a simple EN system with a normalization method to reduce the effect of measuring environments before being analyzed by the ANNs. The results from both the BPNN and the LVQ show the ability to classify the joss stick smoke from different brands and different concentration levels accurately. The overall accuracy, A_c of the classification in this application is calculated by

$$A_c = \frac{\text{Total number of correct classified data}}{\text{Total number of tested data}}.$$

In this application, A_c from the BPNN and the LVQ method using Group 1 for training and Group 2 for testing are 93% and 89%, respectively. When using Group 2 for training and Group 1 for testing, A_c from the BPNN and the LVQ method are 95% and 92%, respectively. We have tried to reduce the dimension of the input data by using only the average signals from each sensor as the input data in order to reduce the training time, but

the correct classification rates of similar data such as B3, G3, and P2 are lower than 50% and the overall classification is not satisfied. When using all data in the training data set for training, the overall classification rate is not higher than the classification rate when using only a small number of training data from the algorithm in Section 5. However, the training time is much longer than the case of using the selected training data set.

We have reviewed some applications [5]~[8] using the ENs that have various mechanisms to control the measuring environment. The distribution of data from those applications are also scattered as shown in the plot of the PCA in their applications. The overall accuracy of those applications is around 80% to 90 % that is not better than the accuracy of the EN shown in this paper.

Actually, there are a lot of applications that require the ENs to replace the human nose or the GC. Even though there are several commercial ENs available in the market, these ENs are still not widely used due to their high cost, large size, and the accuracy that can be achieved only the moderate level. The results from the application show an alterative way for olfactory classification by using a simple EN with an effective method to analyze the data.

8. Conclusions

This paper has presented a simple EN system with an effective method to analyze the data that are varied due to the effect from the measuring environment. The results from this experiment have demonstrated that without an expensive mechanism to control the measuring environment, a simple EN is also able to classify various kinds of similar smell precisely by utilizing the normalization method to reduce the effect of measuring environment and a proper method to select a proper training data.

The results of this application give us a hope to develop an inexpensive but reliable EN system for other useful applications, especially for the environmental monitoring system in our future work.

Acknowledgment

References

(1) P. Keller, L. Kangas, L. Liden, S. Hashem, and R. Kouzes: "Electronic Noses and Their Applications", Proceedings of World Congress on Neural Networks (WCNN'96), San Diego, CA, USA, pp.928–931 (1996)

(2) Paul E. Keller: Physiologically Inspired Pattern Recognition for Electronic Noses, Pacific Northwest National Laboratory (available as: http://www.emsl. pnl.gov:2080/proj/neuron/ papers/ keller.spie99a.html).

(3) G. Olafsdottir et al.: Application of an Electronic Nose To Predict Total Volatile Bases in Capelin (Mallotus villosus) for Fishmeal Production, Journal of Agriculture Food Chemistry,
Vol.48, pp.2353–2359 (2000)

(4) Sameer Singh, Evor L. Hines, Julian W. Gardner, Fuzzy Neural Computing of Coffee and Tainted Water Data from an Electronic Nose, Sensors and Actuators B, Vol.30, issue3 pp.190–195 (1996)

(5) Julian W. Gardner, et al.: AN Electronic Nose System to Diagnose Illness, Sensors and Actuators B, Vol.70, pp.19–24 (2000)

(6) Penza, G. Cassano, F. Tortorella and G. Zaccaria: "Classification of food, beverage and perfumes by WO3 thin film sensors array and pattern recognition techniques", Sensors and Actuators B, Vol.73, pp.76–87 (2001)

(7) Pascal Chatonnet and Denis Dubourdieu, Using Electronic Odor Sensors to Discriminate among Oak Barrel Toasting Levels, Journal of Agriculture Food Chemistry, Vol.47, pp.4319–4322 (1999)

(8) Charon Zondervan et al.: Controlling Maillard Reactions in the Heating Process of Blockmilk Using an Electronic Nose, Journal of Agriculture Food Chemistry, Vol.47, pp.4746–4749 (1999)

(9) B. Charumporn and S. Omatu: Classifying Smokes using an Electronic Nose and Neural Network, SICE Annual Confe. 2002, pp.2871–2875, Osaka, Japan, (2002)

(10) W.L. Carlson and B. Thorne: Applied Statistical Methods, Prentice Hall International, (1997)

(11) A. Stuart and J. Keith Ord: Advanced Theory of Statistics Vol.2, 5th Edition, Edward Arnold, (1991)

(12) L. Fausett: Fundamental of Neural Networks, Prentice Hall International, (1994)

Bancha Charumporn (Non-member) was born in Bangkok, Thailand, on March 27, 1969. He received his B.E. degree in Production Engineering from King Mongkut's Institute of Technology Thonburi, Thailand in 1990 and M.S degree in Industrial Engineering from The University of Texas at Arlington, USA in 1992. Now he is a Ph.D candidate at Department of Computer and Systems Science, Graduate School of Engineering, Osaka Prefecture University.

Michifumi Yoshioka (Member) was born in Osaka, Japan, on December 10, 1968. He received the B.E., the M.E., and the Ph.D. degrees in Geo System Engineering from the university of Tokyo, Japan, in 1991, 1993, and 1996, respectively. He has joined Osaka Prefecture University since 1996 as a research associate. Now he is an assistant professor in the Graduate School of Engineering. His current interests center on image processing method using neural networks.

Sigeru Omatu (Member) was born on December 16, 1946. He received his B.E. degree in Electrical Engineering from Ehime University in 1969 and the M.E, and the Ph.D. degrees in Electronics Engineering from Osaka Prefecture University in 1971, and 1974, respectively. He was with Tokushima University as a research associate since 1974, a lecturer since 1975, an associate professor since 1980, and a professor since 1988. Then he joined Osaka Prefecture University in1995, where he is currently a professor in the Graduate School of Engineering.