Paper

Detecting Household Burning Smell Using a Neuro-Electronic Nose System

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

Conventional fire detecting systems lack the ability to detect fire in the early stages since they trigger the alarm by the high density of smoke or the high air temperature. In this paper, a new electronic nose (EN) system is proposed as an alternative way to detect various sources of fire from the burning smells. The EN is added with a mechanism to reduce the effect of the temperature and the humidity. Consequently, the time series data from the same smell in every repetition data are highly correlated. We have selected only a single data from each source of smell that has the highest average similarity index (SI) value to be a training data for the error back-propagation neural networks (BPNN). Generally, the time series data can be used as the input data for the BPNN directly. However, it will consume a lot of time for training due to the huge dimensional data. A new method called slope max mean (SMM) is applied to reduce the dimension of the input data. By using the SMM data, the training time is reduced and the overall classification rate of 99.8% is achieved which shows the high feasibility to apply the EN as a precision fire detecting system.

Keywords: electronic nose, fire, back-propagation, k-means algorithm, slope max mean, similarity index

1. Introduction

The damage from fire disasters brings about not only the severe loss to property assets, but also the physical and psychological injuries of the people. Although most of the residences have installed some fire detecting systems, those devices lack the ability to detect fire in the early stages since they trigger the fire alarm based on the high smoke density or the high air temperature. Some researchers [(1)~(3)] have proved that conventional fire detecting systems are not reliable to detect fire in the early stages and can not discriminate between the smoke and the odor sources. Jame A. Mike [(1)] and Susan L. Rose-Pehrsson, et al [(2)]. have applied the metal oxide gas sensors (MOGSs) to identify the different stages of fire. Although their results have proved that the MOGSs are better than the photo electric and the ionization smoke detectors, their devices are still not precise enough to classify the smells into specific categories. In this paper, we extend their works by using an electronic nose (EN) that has been successfully applied to classify not only the same smoke from different brands, but also the same smell at different concentration levels [(4)(5)] to identify various burning smells. It has been known that the MOGSs are sensitive to the temperature and the humidity. Thus, the EN that uses various MOGSs as the olfactory receptors is put into a chamber that has a mechanism to reduce the effect from the temperature and the humidity of the air. The air flow system and the sampling headspace are also redesigned in this experiment.

The smells from various sources of fire are measured in this experiment by connecting the headspace of the EN directly over the tested smell. By controlling the temperature and the humidity in the tested chamber, the time series signals from the same smell in every repetition data are highly correlated. Therefore, only a single data that has the highest average similarity index (SI) to its repetition data is selected to be the training data for the error back-propagation neural networks (BPNN). Although only a single time series data is used for training, the training time is quite long due to the huge dimension of the input data. Thus, a method to reduce the input dimension from the time series called slope maximum mean (SMM) is applied before feeding the data to the BPNN. By using only a single data of each smell from the SMM for training, the testing results of 99.8% correct classification can be achieved and the training time is much less than the case of using the full time series data. The k-mean algorithm method is also applied to analyze the time series data, and 98.3% of the overall classification rate is achieved which also shows the high ability of the EN to identify various burning smells. The results show the high feasibility to apply the EN as a precision device to detect the smells of burning materials before turning to be a severe fire disaster.

This paper is organized as follows. Chapter **2.** explains the EN system. The experiment and data col-

* Division of Computer and Systems Sciences,
Graduate School of Engineering,
Osaka Prefecture University, Sakai, Osaka 599-8531, Japan.
Telephone: +81-72-254-9278, Fax: +81-72-257-1788
Email: bancha@sig.cs.osakafu-u.ac.jp

lection are briefly explained in Chapter **3.** Chapter **4.** describes the correlation of the data using the SI value. The SMM method is explained in Chapter **5.** The BPNN are briefly described in Chapter **6.** The results are discussed in Chapter **7.** Finally, the conclusion for this paper is summarized in Chapter **8.**

2. Electronic Nose System

The EN shown in Fig.1 is designed based on the concept of a human olfactory system [6][7] by using eight kinds of MOGSs from FIS Inc., listed in Table 1 as the olfactory receptors. The main part of the 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 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 [8]. Based on this property, we can measure the total voltage changes during the sensors absorbing the tested smell.

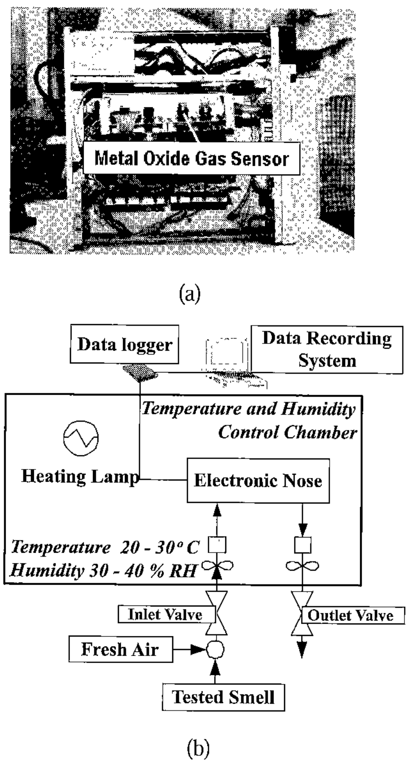


Fig. 1. (a) Inside of the electronic nose containing various MOGSs. (b) Schematic diagram of the EN system.

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
SP-MW1	Cooking vapor

In the previous works [4][5], the EN has been applied to classify various kinds of smokes. We have found that the EN is sensitive to the temperature and the humidity. Therefore, the EN is put into a small chamber that has a heating system to increase the air temperature during the winter season. The heating system can also decrease the humidity from the air. Pure water is manually sprayed into the chamber when the humidity drops below the control level. The temperature and the humidity in the chamber during the tested period are controlled at 20–30 °C, and 30–40 %RH, respectively. The air flow system and the sampling headspace of the EN are redesigned by making a conducting pipe to flow the fresh air from outside of the measuring room. Another conducting pipe is connected over the tested smell to the inlet valve during the test period. Then the tested smell is sucked and mixed with the fresh air at the inlet valve before being fed to the EN. The distance from the headspace to the EN is about 1.5 m which approximately equals the distance from the working surface to the ceiling of the house. The data are converted from analog to digital by the data logger and stored into the computer. Finally, the artificial neural networks (ANN) such as the BPNN, or the statistical analysis such as the k-means algorithm, can be applied to analyze these data.

3. Experimental Data Collection

In this experiment, the smell from twelve sources of fire listed in Table 2 are measured by the EN system explained in Chapter 2. The experiments have been done during the winter season. In order to verify the reproducibility of the EN, each smell has been tested with forty repetition data.

For each data, the voltage signal of the normal air before feeding the tested smell is measured every second for one minute and its average value, $\bar{v}_{air}$, is used as an air reference point. After that, the voltage signals of the sensors when absorbing the tested smell, $v_{smell,t}$, are collected every two seconds for a period of two minutes on each smell sample. During the tested period, the tested material is continually burnt and the smell is continually sucked into the tested chamber. Finally, the total change in signals on each period, $V_{smell,t}$, is calculated by

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

Table 2. List of tested smell in the experiment.

Sources of fire	Abbreviation
Steam from boiling water	Steam
Burning joss stick	Joss
Burning mosquito coil	Mos
Aroma oil	Aroma
Aroma candle	Candle
Flame from liquid petroleum gas(LPG)	Flame
Leakage of LPG	LPG
Steam from Japanese soup called "oden"	Oden
Boiling vegetable oil	Oil
Toasted bread	Toast
Burning paper	Paper
Burning wood	Wood

where t is the period from 1 to 60.

After testing each smell, the EN needs to be cleaned by removing the tested smell and supplying only the fresh air until the signals of all sensors return to a stable stage before being able to test the new sample. This process is just like the human nose which needs to breath fresh air before being able to recognize a new smell accurately. The samples of time series data from the experiment are plotted in Fig. 2.

Figure 2 shows that each source of smell has a unique data pattern. Each smell activates the responses of all eight MOGSs in different ways. During the experiment, we have found that the consistent burning smells such as the joss stick and the mosquito coil, have highly correlated repetition data. In the other ways, the inconsistent tested smells such as the paper and the wood tend to have scattered data. Therefore, the method to investigate the correlation of the data is applied before analyzing the data by the BPNN and the k-means algorithm. The detail of this topic will be discussed in the next chapter.

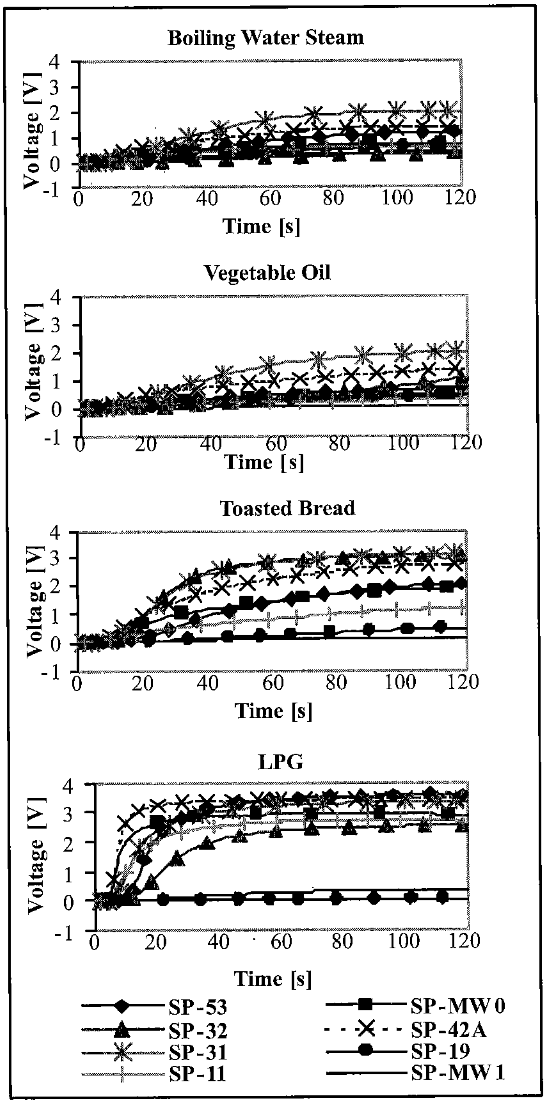


Fig. 2. Sample data from this experiment.

4. The Correlation of Data

Before analyzing each source of data, the correlation between the repetition data of the same smell is investigated by using the SI value and the principal components analysis (PCA). The SI value can show the linear relationship between the compared data and the PCA can show the distribution of the data. By using the information from the SI value and the PCA, a proper training data for the BPNN and the initial center vectors for the k-means algorithm can be obtained by choosing the data that has the highest average SI value.

4.1 Similarity Index In the statistical application, the correlation value developed mainly by Karl Pearson is widely used to find the relationship between two random variables. In this paper, we call the 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. Random variables could have a nonlinear relationship but still have an SI close to 0 [9]. Therefore, we assume that each data pattern in the experiment has nearly linear relationship to the other data patterns. The SI value between two data is calculated by

$$r_{xy} = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sqrt{\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right)\left(\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right)}} \quad \cdots\cdots (2)$$

where r_{xy} is the SI value, $\sum$ denotes $\sum_{i=1}^{n}$, x_i and y_i are the comparing data, and n is the dimension of data which equals to 480 (8 sensors $\times$ 60 periods).

The SI values of the data that have high consistent burning rates such as the joss stick and the mosquito coil are higher than 0.995 in every compared data. The data that have inconsistent burning rate like the paper and the wood have lower SI value. This means that the repetition of these data types are less correlated than the data that have high SI value. In Chapter 7, we will show that the classification rate of the data which have high SI value is higher than those of the data that have low SI value. For each smell, only a single data that has the highest average SI value to its own repetition data will be used as a training data for the BPNN, and the initial vector for the k-means algorithm.

4.2 Principal Components Analysis The principal components analysis (PCA) is a popular method to transform huge dimensional data vectors into a small number of principal components [10]. The PCA is applied to transform the time series data in this experiment from the data size of four hundred and eighty into two main components as shown in Fig. 3.

The principal components 1 and 2 represent about 64% and 16% of the accumulated total variance, respec-

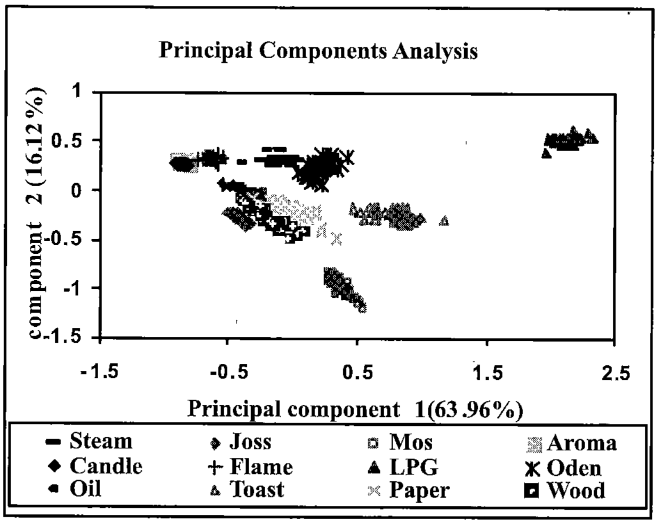


Fig. 3. Two main components of the experimental data using the PCA.

tively. Each kind of smell is distributed into its own region with some overlap zones. The distributions of the data that have inconsistent burning rate such as the paper and the wood are more scattered than the other data that have better consistent burning rate such as the joss stick. The scattered data has high risk to be incorrectly classified as will be discussed in Chapter 7.

5. The Slope Max Mean (SMM) Data Reduction

The SMM is applied to transform the information of the time series data vectors into a smaller size data with the aim to reduce the training time of the BPNN without decreasing the classification accuracy. The SMM method is explained by using the sample signals from the sensor SP-32 of the LPG data shown in Fig. 4. The original data from each MOGS is divided equally into n intervals. The signals of each interval are replaced by the estimated regression line and the *slope* of all regression lines are used as the first n components of the SMM data. Several time intervals are tested and the interval that gives the best classification rate is equals to 40s ($n = 3$). Thus, the first three components of SMM data are obtained from the the *slope* of the regression line in each interval. Then the *maximum* value and the *mean* value of the data over the measuring periods (60 periods = 120s) are added as the fourth and the fifth components of SMM data, respectively. By using the SMM data, the data size is reduced to forty (8 MOGSs × 5 SMM components).

The estimated regression model of each interval is specified by

$$y_i = \beta_0 + \beta_1 x_i \qquad i = 1, 2, ..., 20 \cdots\cdots\cdots\cdots (3)$$

where y is the approximated voltage, β_0 is the intercept, β_1 is the slope of the regression line, and x is the time period.

The slope, β_1, using the least square estimation method [11] can be approximated by

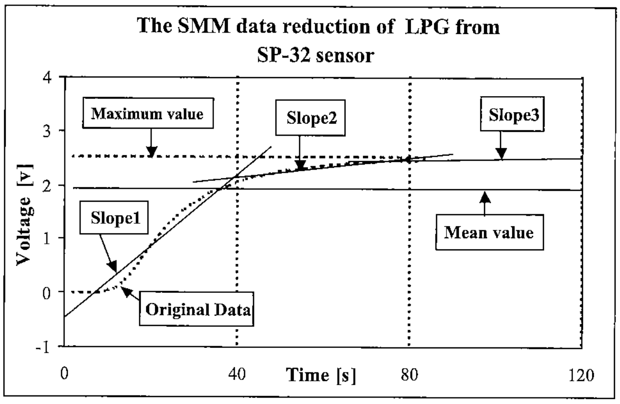


Fig. 4. SMM data reduction.

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \cdots\cdots\cdots\cdots\cdots\cdots (4)$$

where $\overline{x} = \frac{\sum_{i=1}^{n}(x_i)}{n}$, $\overline{y} = \frac{\sum_{i=1}^{n}(y_i)}{n}$, and $n = 20$.

6. Error Back Propagation 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 connects to the neuron k of the output layer with 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 output mean square error (MSE) is minimized.

Let o_k denote the output reference value of the k^{th} neuron of the output layer and let x_i, z_j, y_k denote the output of i^{th}, j^{th}, k^{th} neuron in the input layer, hidden layer, and output layer, respectively. The MSE to be minimized is given by Eq.(5).

$$\mathrm{MSE} = \frac{\sqrt{\sum_{k=1}^{m}(o_k - y_k)^2}}{m} \cdots\cdots\cdots\cdots\cdots\cdots (5)$$

where m is the number of the output node. The updating rule for the connection weights w_{jk} at iteration t of the learning process is incremented by Eq.(6).

$$\triangle w_{jk}(t+1) = \alpha\delta_k z_j + \mu\triangle w_{jk}(t) \cdots\cdots\cdots\cdots (6)$$

where $\alpha > 0$ is the learning rate, $0 < \mu < 1$ 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 connection weights v_{ij} at iteration t is updated by Eq.(7)

$$\triangle v_{ij}(t+1) = \alpha\delta_j x_i + \mu\triangle v_{ij}(t) \cdots\cdots\cdots\cdots (7)$$

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

7. Classification Results and Discussion

7.1 Experimental Results The BPNN and the k-means algorithm are applied to analyze the data. For the BPNN, three cases of data are used for analyzing. Case 1 uses the full time series data. Case 2 uses the average data from each MOGS. Case 3 uses the data from the SMM method as explained in Chapter 5. The k-means algorithm is applied only for the full time series data. All cases use only a single data from each smell that has the highest average SI value as explained in Section 4.1 to be the training data for the BPNN and the initial vector for the k-means algorithm. The structure of the BPNN in this application shown in Fig. 5 has three layers.

The input layer of Case 1, Case 2, and Case 3 consists of four hundred and eighty nodes (8 MOGSs $\times 60$ periods), eight nodes (8 MOGSs), and forty nodes (8 MOGSs $\times$ 5 components of SMM), respectively. For the hidden layer, we have tried with many values. Finally, forty nodes, forty nodes, and thirty nodes turned out to be the best for Case 1, Case 2, and Case 3, respectively. The output layer contains twelve output neurons. Each output neuron presents each data type in Table 2. The best training parameters getting from a trial and error method for all cases are, a learning rate, $\alpha = 0.1$ and a momentum rate, $\mu = 0.001$. The training process for all cases is done until the MSE ≤ 0.0003. After finishing the training process, the rest of the data are tested. We assume that a pattern is classified correctly if (output$\geq$ 0.6 and target =1) or (output $\leq$ 0.3 and target =0).

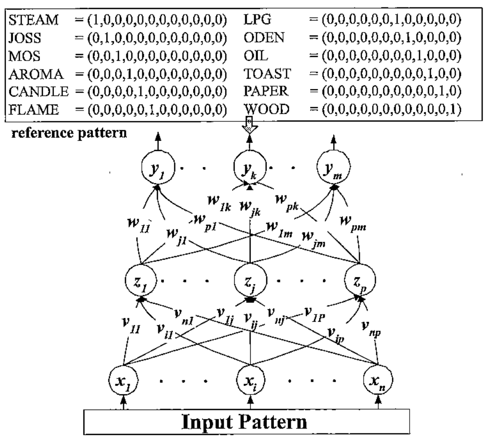


Fig. 5. Structure of the BPNN.

For the k-means algorithm, the training data from Case1 of the BPNN are used as the initialize data and then assigns the data patterns to the nearest cluster center by calculating the Euclidean distance. After that, the new cluster center is recalculated. The process continues until the position of the cluster center is not changed.

The final results of this experiment are shown in Table 3. The overall accuracy of 99.6% is achieved in Case 1 while using the training time almost 15 hours. The results from Case 3 show 99.8% correct classification which is nearly the same as the results from Case 1. However, the training time is only 2.5 hours which is much less than the training time of Case 1. Case 2 requires only around 1.5 hours of training time, but the classification rates are not as good as those of the other cases since we use only average value of data for analyzing. The results of the data that have consistent smell such as the joss stick, the mosquito coil, or the LPG are perfectly classified in every case. These data are the data that have high correlation as can be noted by the high SI value and the PCA in Chapter 4. However, in case of inconsistent smell like the paper and the wood which have low SI values as described in Chapter 4, the results from Case 2 is much worse than the other cases due to the variation signals of these data. The results from the k-means algorithm reaches 98.3% of average correct classification rate which is also quite high while using only a minute for calculation.

7.2 Discussion There are two data of Case 1, and one data of Case 3 from the burning paper smell that are incorrectly classified. As mention in Chapter 4, we use only one data that has the highest average SI value to its own repetition data from each smell as the training data. In case of paper burning smell which has inconsistent burning rate, the signals of this smell are varied from data to data. Thus, it is possible to have some incorrect classified data for this kind of smell. In this paper, the SMM method can be applied to reduce

Table 3. Experimental Results

Sources of fire	Error back-propagation (BP)									k-means	
	Case 1, Full Time Series Data			Case 2, Average Data			Case 3, SMM Data				
	Correct	%	Avg. Output	Correct	%	Avg. Output	Correct	%	Avg. Output	Correct	%
Steam	39/39	100	0.988	38/39	97.4	0.986	39/39	100	0.991	40/40	100
Joss	39/39	100	0.997	39/39	100	0.998	39/39	100	0.996	40/40	100
Mos	39/39	100	0.995	39/39	100	0.997	39/39	100	0.998	40/40	100
Aroma	39/39	100	0.995	39/39	100	0.991	39/39	100	0.995	40/40	100
Candle	39/39	100	0.963	39/39	100	0.954	39/39	100	0.958	40/40	100
Flame	39/39	100	0.978	39/39	100	0.989	39/39	100	0.987	40/40	100
LPG	39/39	100	0.998	39/39	100	0.998	39/39	100	0.998	40/40	100
Oden	39/39	100	0.992	39/39	100	0.991	39/39	100	0.995	40/40	100
Oil	39/39	100	0.987	38/39	97.4	0.952	39/39	100	0.975	40/40	100
Toast	39/39	100	0.982	38/39	97.4	0.931	39/39	100	0.995	40/40	100
Paper	37/39	94.9	0.935	31/39	79.5	0.782	38/39	97.4	0.932	35/40	87.5
Wood	39/39	100	0.964	32/39	82.1	0.824	39/39	100	0.990	37/40	92.5
Average		99.6	0.982		96.2	0.949		99.8	0.984		98.3

the dimension of the input data for the BPNN without decreasing the classification accuracy. There are still many factors that could reduced the classification accuracy, such as the combination of the smell from various sources, the concentration of the smell, the distance from the burning sources, and so on. From the correlation of the data in this experiment, it can be concluded that the EN has a high reproducibility and high reliability. We believe that even though the tested factors are varied, the EN will be able to respond to those smells perfectly and the BPNN using the data from SMM method will be able to clarify the smells accurately. However, the sample of those smells in every repetition should have consistent smell.

8. Conclusions

This paper has presented an EN system that is able to identify various burning smells precisely. The EN has a high reproducibility qualification as can be noticed from the high correlation of the same smell in every repetition data. By using only a single SMM training data, the overall classification rate of 99.8% is achieved. Therefore, it can be concluded that the EN system is suitable to be applied as a precision fire detection system since it is able to detect the smell of burning materials in the early stages when the smoke density is not high enough to trigger the alarm of a smoke detector.

The EN system can be applied not only as the fire detecting system, but also as a device to detect the smell in many applications, especially, in the case of identifying the toxic smell that human nose cannot test. Our future work is to develop both the mechanical and the analytical system of the EN to be more reliable and applied in other useful applications.

Acknowledgment

The authors would like to thank Mr. Nobuaki Murakami of FIS. Inc. for the technical information on making this electronic nose and for the data logger equipment.

(Manuscript received Feb. 19, 2003,
revised Aug. 20, 2003)

References

(1) Jame A. Mike: "Application of Neural Networks for Discriminating Fire Detectors", *Proc. of 10th international conference on Automatic Fire Detection*, AUBE'95, Duisburg,Germany, pp.213–222 (1995)

(2) Susan L. Rose-Pehrsson, et al.: "Real-Time Probabiliastic Neural Network Performance and Optimization for Fire Detection and Nuisance Alarm Rejection", *Proc. of 12th international conference on Automatic Fire Detection*, AUBE'01, Maryland, USA, pp.176–188 (2001)

(3) Milan Dj. Blagojević and Dejan M. Petković: "Detecting Fire in Early Stage- a New Approach", *Working and Living Environment Protection*, Vol.2, No.1, pp.19–26 (2001)

(4) B. Charumporn, M. Yoshioka, T. Fujinaka, and S. Omatu: "An Electronic Nose System Using Back Propagation Neural Networks with a Centroid Training Data Set", *Proc. of the Eight International Symposium on Artificial Life and Robotics*, Oita, Japan, pp.605–608 (2003)

(5) B. Charumporn and S. Omatu: "An Electronic Nose System with Neural Networks for Smoke Classification", *Record of the 2002 Kansai-Section Joint Convention of Institute of Electrical Engineering*,Osaka, Japan, pp.G69 (2002)

(6) 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)

(7) Peter L. Williams: "Gray's Anatomy thirty-eighth edition", ELBS, *Churchill Livingstone Publishing, London* (1995)

(8) General Information for TGS sensors, *Figaro Engineering*, available at www.figarosensor.com/products/general.pdf

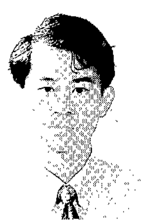
(9) William L. Carlson and Betty Thorne: "Applied Statistical Methods", *Prentice Hall International* (1997)

(10) R.A. Johnson and W.D. Wichern: "Applied Multivariate Statistical Analysis, Fifth-Edition", *Prentic Hall International* (2002)

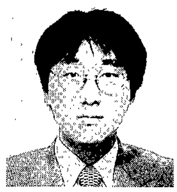
(11) Sanford Weisberg: "Applied Linear Regression, second edition", *John Wiley and Son* (1985)

(12) Laurence Fausett: "Fundamental of Neural Networks", *Prentice Hall International* (1994)

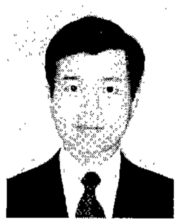
Bancha Charumporn (Non-member) was born in Bangkok, Thailand, on March 27, 1969. He received his B.E. degree in Production Engineering from King Mongkut's Institute of Technology Thonburi, Thailand in 1990 and M.S degree in Industrial Engineering from The University of Texas at Arlington, USA in 1992. 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.



Toru Fujinaka (Member) received the B.E., the M.E, and the Ph.D. degrees, all in Electrical Engineering from Kyoto University, Japan, in 1979, 1981, and 1990, respectively. He was with Kyoto University as a research associate from 1986 to 1990, and then he joined Osaka Prefecture University, where he is currently an associate professor in the Graduate School of Engineering. His research interests include theory and application of optimal and intelligent control.



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 in 1995, where he is currently a professor in the Graduate School of Engineering.