



Selection of the Most Efficient Sensors for Methanol Detection in Fake Alcohol

Zou XIAOBO* and Zhao JIEWEN

School of Biological and Environmental Engineering,
Jiangsu University, Zhenjiang, Jiangsu 212013, China

*Phone: +86-0511-8780174 *E-mail: zou_xiaobo@ujs.edu.cn

Received: 26 July 2004 / Accepted: 14 August 2004 / Published: 20 August 2004

Abstract: High concentration ethanol was always found in adulteration fake alcohol. And almost all tin oxide gas sensors are high sensitive not only to methanol but also to ethanol. Furthermore, high concentration of ethanol in wine is very easy to saturate sensor responses and mask the response to other aroma compounds. In order to optimize the system cost, speed and reliability and then to select the best sensors suitable for this application, a method called organization feature parameter based on formulae expression tree by using genetic algorithms has been proposed in this paper. It could solve the problem how to getting optimum sensors, and make the genetic algorithm more convenient and straight Three optimum gas-sensors were selected, and combination the three sensors feature parameters by the new method could distinguish different fake alcohol samples.

Keywords: Gas sensor array; feature parameters; genetic algorithm; methanol; fake alcohol

1. Introduction

Alcohol is a popular drink in the world. But poisoning with methanol may be accidental and intentional. There were 700 severe accidents caused by fake alcohol in recent years in China, in which lawbreaking persons adopted industrial alcohol to make fake alcohol caused 129 death and 6000 poisoning and 33 losing their sight. A case of fake alcohol happened in Suozhou City, Shanxi Province

in 1998 was a case shocking by the sight with 27 dead and around 700 have been hospitalized after drinking poisoned alcohol made from industrial alcohol by money-motivated mobsters [1]. There have been epidemics of methanol toxicity in cases where illicit whiskey has been sold to large populations or when the less expensive methanol was substituted for ethanol in drinks in west countries.

Methanol, CH₃OH, Mr 32.042, also termed methyl alcohol or carbinol, is one of the most important chemical raw materials. Methanol is produced from the distillation of wood and is a clear, colorless, volatile liquid with a weak odor that is somewhat sweeter than ethanol. Methanol is well absorbed from the gastrointestinal tract mucosa as well as through the skin and lungs. Doses as low as 25 cc of 40% methanol have been reported as causing toxicity. Most sources consider the minimal lethal dose to be around 100 cc (1 g/kg). The overall mortality of methanol poisoning is approximately 20% and among survivors the rate of permanent visual impairment is 20-25% [2]. However, there is no different for human's visual and taste between methanol and ethanol. And it is very difficult for man to detect the fake alcohol by sensory analysis. Therefore, methanol is a very dangerous ingredient in alcoholic drinks for public health and is a must be examination item in commercial alcoholic drinks. By GB5009-96, there are two methods for determination of methanol, i. e. fuchsine-sulfurous acid colorimetry and gas chromatography. The two methods have their advantages, such as high sensitive, and disadvantages, such as gas chromatography is too expensive and not portable, it can only be used in lab. And fuchsine-sulfurous acid colorimetry is miscellaneous, and it needs a lot of time to prepare the experiment. Furthermore, the two methods are inappropriate to examine colored drinks.

Recently there is an increasing interesting in research, development and application of gas sensors with an associated pattern-recognition technique to identify the odors given out from food, as it can monitor the quality of foods and presents a new food aroma analysis method compared to the sensory analysis by natural olfaction [3-9]. Over the past few decades, tin oxide based films are widely used as gas sensors due to their high sensitivity in the presence of small amounts of some gases of interest viz. carbon monoxide, ethanol, methane etc. [3, 10, 11]. And intensive research in the field of tin-oxide gas sensor array technologies has lead to multi-parameter extraction techniques [3]. Many sensors of the same type can be found although they produce only slightly different responses, as they are still non-specific. One of the most promising applications of this approach is in the development of application specific instruments, by selecting the optimum sensors. Research should not only focus on sensor technology development and powerful classification algorithms, but parameter selection, which is considered to be an important intermediate step as a subset of parameters can often provide better discrimination [4, 12, 13]. Hence, a systematic or structured method for selecting the best sensors, identifying the optimal array configuration or finding the key parameters would be desired to enhance the overall system performance. The main motivation for tin oxide gas sensor array based devices is the development of a low-cost, real-time, and portable method to perform reliable, objective, and reproducible measures of alcohol volatile compounds and odors.

In this paper, a series of tests were carried out on good and fake alcohol samples. In Section 2, we present the gas sensor system and the experimental procedure. The sensor selection techniques and methods were considered in Section 3. Results are also presented in Section 3. The conclusions in Section 4, discuss the advantages of the technologies.

2. Experiment

High concentration ethanol was always found in adulteration fake alcohol. And almost all tin oxide gas sensors are high sensitive not only to methanol but also to ethanol. Furthermore, high concentration of ethanol in wine is very easy to saturate sensor responses and mask the response to other aroma compounds [5]. Therefore, how to setup an experiment and how many sensors should we use in the system that could detect methanol in high concentration fake alcohol on-line? The experiment and the gas sensor selection system were as follows.

2.1. Gas Sensor Selection System

Nine tin-oxide sensors are commercially available Taguchi-type gas sensors obtained from Figaro Co. Ltd. (Sensor 1, TGS 813; Sensor 2, TGS 821; Sensor 3, TGS 824; Sensor 4, TGS 880; Sensor 5, TGS 822; Sensor 6, TGS 822TF; Sensor 7, TGS 825; and Sensor 8, TGS 812 Sensor 9, TGS 800). These sensors are heated to a constant temperature, holding the sensor heater voltage at 5V. Exposure of a tin-oxide sensor to methanol or ethanol vapor produces a large change in its electrical resistance [13, 14]. The gas sensor selection system (Fig.1) is composed of an 89C51 as CPU, a strip gas sensor chamber, a 12-bit AD/DA converter, and an air filter for suppressing humidity, a suction pump, and a personal computer. The temperature of the whole system was controlled. The system was used to select sensor and to combine a low cost, real-time, and portable device that could detect methanol in alcohol on line.

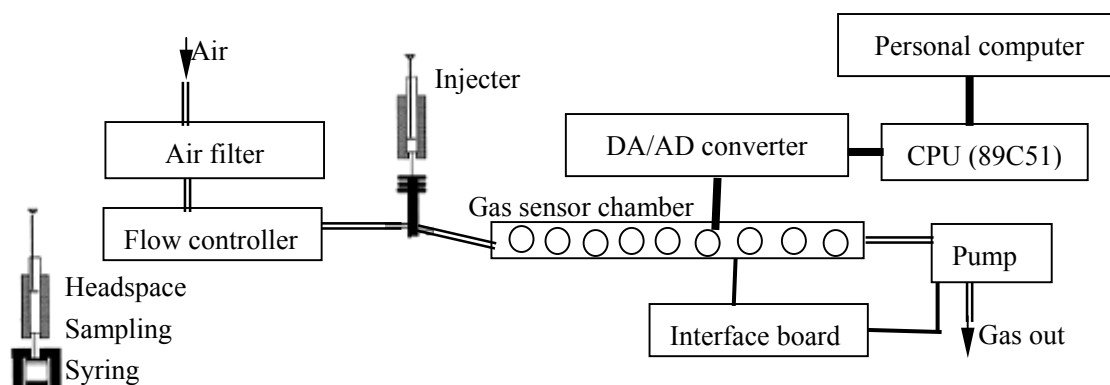


Fig.1. Schematic diagram of gas sensor selection system

2.2. Adulteration Fake Alcohol Samples and Experiment Procedure

The system applied to the discrimination good alcohol and adulteration fake alcohol. Currently in the China market, the most saleable alcohol contained ethanol (v/v) about 50% (Wuliangchun spirit). And the concentration of methanol in adulteration fake alcohol is higher than 1%.

The mother (original) solutions are prepared: in the two groups, one solution composed of 50% ethanol alone, representing 'good vine', the other one with 1% methanol dissolved in 50% ethanol mimicking the adulteration 'fake alcohol'.

In order to generate one headspace, 10 ml of liquid sample is drawn from one of the original solutions and injected into a 100 ml vial; the headspace is generated over 30 min at 24. 20 ml of headspace is injected with a syringe as figure 1 shown. Then, the gas sample is conveyed to the measurement cell by a carrier gas. This carrier gas is the atmospheric air, thermostatically controlled, filtered on active charcoal and dehydrated with silica gel. Its flow-rate is controlled at 700 ml/min, either for cleaning the measurement cell or for the static injection. Exposure of a tin-oxide sensor to vapor produces a change in its electrical resistance. During operation the sensor array "smells" the gas from the head-spaces of one alcohol sample, the sensor signals are digitized and fed into computer 3 times every second, and the whole signal is exploited, from the absorption beginning to the desorption phase, and stationary phase of equilibrium between reversible adsorption and desorption, the process lasts 3~4 min. At last, as the carrier gas cleaning the measurement cell the sensors were recovered. These conditions have proved to be the most repeatable and give the best agreement between volatile concentration and gas sample generation duration. Each measurement (including headspace generation from a renewed liquid sample) is repeated 20 times. Fig.2 is the typical response curve for the gas sensor array reaction, the curve is smoothed and the baseline is subtracted. Here we cite the experiment of the sensors response to one 'fake alcohol' sample solution that contained 1% methanol and 50% ethanol.

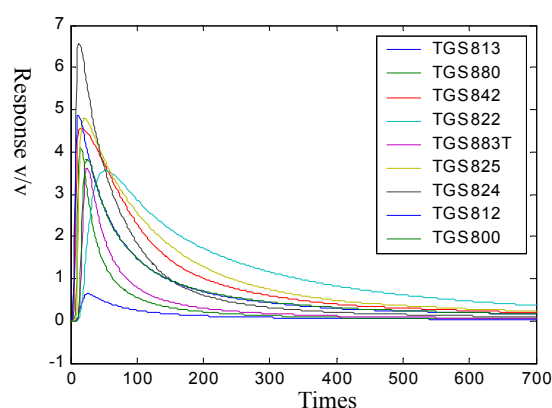


Fig.2. Typical response curve for the gas sensor array reaction.

3. Sensor selection

For tin oxide sensors, Sylvie Roussel, Gustaf Forsberg, etc. demonstrated a methodology of output feature parameter (FP) selection: From each smoothed and the baseline subtracted curve, 29 features were extracted, 13 from the curve smoothed by a moving average of size 9 points, and 13 from the primary and secondary derivatives [5]. The 29 features were then sorted by repeatability, discriminant distance values and their ratio q_i , Signal maximum and adsorption maximum slop show high level of efficiency. Therefore, in this paper, we extracted two features parameters that were the signal maximum (SM) and adsorption maximum slop (AMS) from each sensor response curve. Then, $2 \times 9 = 18$ features parameters were extracted from the 9 tin oxide gas sensors array. Because 18 features were not in the same scale they must be normalized as follows:

$$Ps_j(i) = \frac{Sm_j(i) - \min(Sm_j(1:n))}{\max(Sm_j(1:n)) - \min(Sm_j(1:n))}, i = 1, 2, \dots, n \quad j = 1, 2, \dots, 9 \quad (1)$$

$$Pa_j(i) = \frac{Ams_j(i) - \min(Ams_j(1:n))}{\max(Ams_j(1:n)) - \min(Ams_j(1:n))}, i = 1, 2, \dots, n \quad j = 1, 2, \dots, 9 \quad (2)$$

Here i is the number of experiment, and j is the sensor. Then $Ps_j(i), Pa_j(i)$ are variables with $[0, 1]$ distribution. Then we can use formula (1) and (2) to make unified normalization for the 18 different features.

3.1 The Combination Method of Sensor Features

PCA was always used to combine the sensor features in gas sensors array study [6-9]. But in this paper, we want to selected the best sensors for distinguish fake alcohol samples. The feature parameters formulae expression tree was used to combine the sensor features. A feature parameter formula that their feature parameters come from sensor 1 and sensor 2 can be expressed, for example, the formula (1) can be expressed by the expression tree shown as Fig.3 (a).

$$\{(Ps_1 - Ps_2) + (Pa_1/Pa_2)\} \times \sqrt{Ps_1 + Ps_2} , \tag{3}$$

Here, Ps_i, Pa_i ($i=1,2,\dots, 9$) are the normalized feature parameters.

Another example of formula and expression tree is shown as formula (4) and Fig. 3(b). In this case, n is the exponential factor with numerical value and is put on a tree node.

$$\{(Ps_1)^n\}^n - \{Ps_1 \times Ps_2 + (Pa_1 - Pa_2)\} , \tag{4}$$

The expression tree can be translated to a structure variable of C language for GA operation.

3.2 GA Operation

The feature parameter (OFP) has been expressed by the expression tree, such as Fig.3. The GA operation can be carried out as follows:

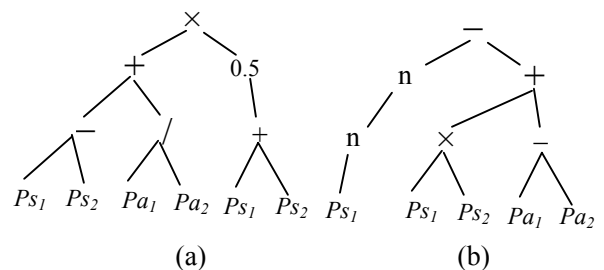


Fig.3. Expression tree of formulas (3) and (4).

3.2.1 Fitness of genotype

The fitness of a feature parameter (FP), which will be used to distinguish two states, such as ‘Good vine (only 50% ethanol)’ or ‘Fake alcohol (1% methanol and 50% ethanol)’, is derived in the following way.

For distinguishing two states (state 1 and state 2), the failure distinction ability of feature parameter can be evaluated by the “Distinction Rate (D.R.) P0” [14] defined in the following formula:

$$p_0 = \int_R f_i(x) dx \quad i = 1, 2, \dots, n \tag{5}$$

Here, $f_i(x)$ is the probability density function measured in the state i , R_i is decided by the following formula:

$$\int_{R_1} f_1(x)dx = \int_{R_2} f_2(x)dx, \tag{6}$$

For example, when $f_i(x)$ is the standard density function, R_i ($x \sim \mu_1, \sigma_1$ and $x \sim \mu_2, \sigma_2$) can be derived as follows:

$$\frac{1}{\sqrt{2\pi}\sigma_1} \int_{-\infty}^{x_0} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} dx = \frac{1}{\sqrt{2\pi}\sigma_2} \int_{x_0}^{\infty} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} dx, \tag{7}$$

Here μ_1 and μ_2 are the mean values of the feature parameters calculated by the signals measured in state 1 and state 2. σ_1 and σ_2 are their standard deviations. x_0 can be worked out as follows:

$$x_0 = \frac{\mu_1\sigma_2 + \mu_2\sigma_1}{\sigma_1 + \sigma_2}, \tag{8}$$

Fig. 4 shows p_0 and x_0 . With the substitution $z=(x-\mu_1)/\sigma_1$ or $z=(x-\mu_2)/\sigma_2$ to formulae (5), (6), the “Distinction Rate P_0 ” can be obtained in following way:

$$p_0 = \frac{1}{2\pi} \int_{-D.I.}^{\infty} e^{-\frac{z^2}{2}} dz, \tag{9}$$

$$p_0 = \frac{1}{2\pi} \int_{-\infty}^{D.I.} e^{-\frac{z^2}{2}} dz, \tag{10}$$

Here, D.I. is called the “Distinction Index” and calculated by the following formula:

$$D.I. = \frac{\mu_2 - \mu_1}{\sigma_2 + \sigma_1}, \mu_2 > \mu_1 \tag{11}$$

It is obvious that the larger the value of D.I., the larger the value of “Distinction Rate P_0 ” will be, and therefore, the better the feature parameter will be. So D.I. can be used as the fitness of the genotype.

3.2.2 Initial group of genotype

For starting GA operation, the initial group of

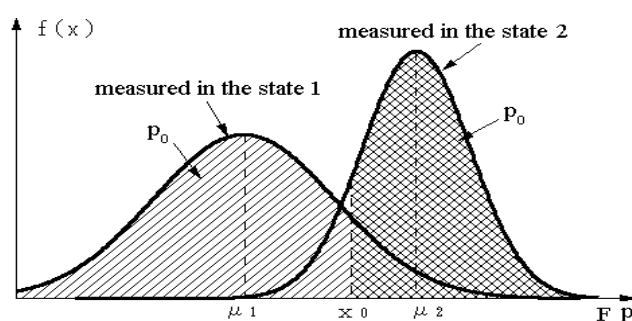


Fig.4. An example of x_0 and p_0

genotype must be generated in a random way. The generation procedure of the initial group is as follows:

- (1) deciding the size of genotype population (POPU); in this paper POPU=25;
- (2) deciding the size of node for each genotype (NODE) in a random way; in this paper $5 \leq \text{NODE} \leq 30$;
- (3) generating the leaf node (node="Psi or Pai") for each genotype in random way; the number of the leaf nodes is less than $\text{NODE}/2+1$;
- (4) generating the parent nodes for each leaf node;
- (5) finishing all initial groups of genotypes.

Table 1. Initial OFP's for GA to distinguish vinegars.

OFP	D.I.	D.R.
$Ps_6 Pa_7^{-1} Ps_8^{-0.33} + (Ps_2 - Pa_7)$	1.23	85.4%
$Ps_5 Pa_3 Ps_4^{-1} Ps_1^{-3}$	1.18	84.8%
⋮	⋮	⋮
$Pa_7^{-1} (Ps_2 + Ps_1) + Pa_2 Ps_5 - Pa_4 Ps_5$	0.60	70.9%

Table 1 shows the example of the initial group and their D.I., D.R. Though the initial group of genotypes is generated in the random way, the GA operations, which we have carried out up to now, always converge. Here, the convergence means that the DI of the optimum OFP obtained by the organization is always larger than 2.0 (DR > 97%).

3.2.3 Crossover and mutation

Genotypes are evolved by one-point-crossover. Figure 5 shows an example of the crossover of two genotypes exchange with each other.

The formulae expressed by the new expression trees shown in Fig.5 are

$$\{(Ps_1 + Ps_2) + (Ps_2/Pa_2)\} \times \{Ps_1 \times Ps_2 + (Pa_1 - Pa_2)\} \tag{12}$$

$$\{(Ps_1)^n\}^n - \sqrt{Ps_1 - Ps_2} \tag{13}$$

The genotypes for cross-breeding are chosen by the probability •j shown in the following formula:

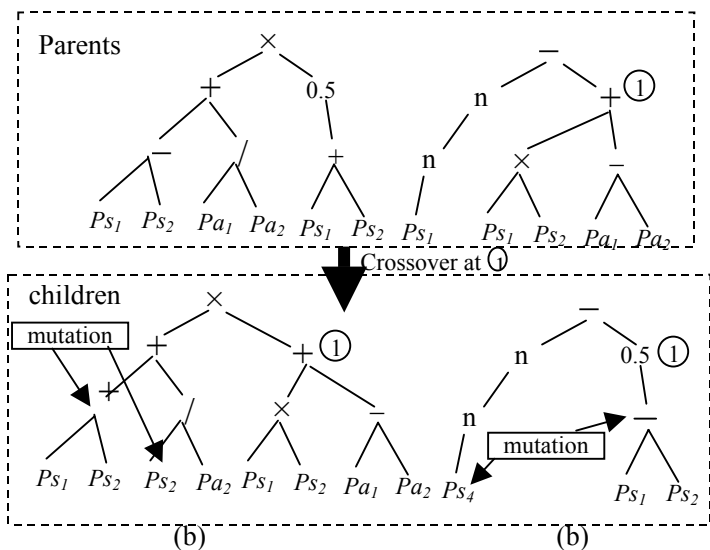


Fig.5. Crossover and mutation.

$$\lambda_j = q_j / q_{\max} ; q_j = \left[\frac{D.I._j}{\sum_{j=1}^G D.I._j / G} \right] \quad (14)$$

Here, $D.I._j$ is “Distinction Index” of the genotype j . G is the total number of genotypes.

The mutation operation is used to change the operator (+, -, *, /) or value (Psi, Pai, n) of a node into another one, as in Fig. 5. The mutation rate is generally decided by experience [15]. We have tried repeatedly to decide the mutation rate. In the cases of the examples shown in this paper, the mutation rate is 2 %. In each generation, the crossover and mutation do not occur in the elite genotype, which has the maximum value of fitness.

Furthermore, if there are some meaningless terms in the formulae of SP, such as $(Ps_4^2)^{-2}$, " $Pa_7 - Pa_7$ ", $(-Ps_2)^{0.5}$ etc., other meaningful terms will be substituted for the meaningless terms.

3.2.4 The terminating conditions

The GA operation is terminated when satisfying one of the following conditions:

- (1) The maximum value of the fitness (DI) does not change during certain generations (for example, 50 generations).
- (2) DI is larger than a given value (such as 60), because if $DI > 70$, the $DR(P_0) = 100\%$.

3.3 Results

The new method discussed here has been used to distinguish between ‘good vine’ and ‘fake alcohol’. The D.I. and D.R.(P₀) of each feature parameter (Psi Pai) defined as sensor’s value are shown in Tables 2.

Table 2. Each feature parameters for distinguishing between ‘good vine’ and ‘fake alcohol’.

Sensor	1	2	3	4	5	6	7	8	9
SM	Ps_1	Ps_2	Ps_3	Ps_4	Ps_5	Ps_6	Ps_7	Ps_8	Ps_9
D.I.	1.491	1.217	0.849	1.28	1.422	0.703	0.756	1.321	1.074
D.R.	90.5	84.94	79.5	85.1	89.0	74.3	76.1	86.1	83.5
AMS	Pa_1	Pa_2	Pa_3	Pa_4	Pa_5	Pa_6	Pa_7	Pa_8	Pa_9
D.I.	0.6589	1.0728	0.0934	0.1156	1.1106	0.0824	0.0394	0.6987	0.6211
D.R.	69.20	83.92	54.92	57.58	84.2	54.3	53.02	69.64	68.75

Table 2 shows that the D.I.’s are less than 1.5 and D.R’(P₀’s) are less than 91%. Consequently the feature parameters are not good enough to distinguish between ‘good vine’ and ‘fake alcohol’. Fig.6 exhibits the results of principal component analysis (PCA) for the two samples. PCA is a simple method to project data from several feature parameters to a two-dimensional space. The values of 87.66% of PCA1 (Fig.6), 4.35% of PCA2 (Fig.6) indicate contribution rate to pattern separation. It shows that the pattern separation is not sharp.

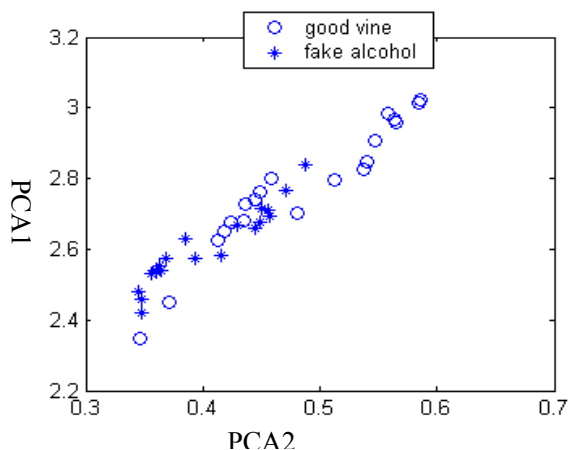


Fig. 6. Results of the PCA of the gas sensor array for experiment samples.

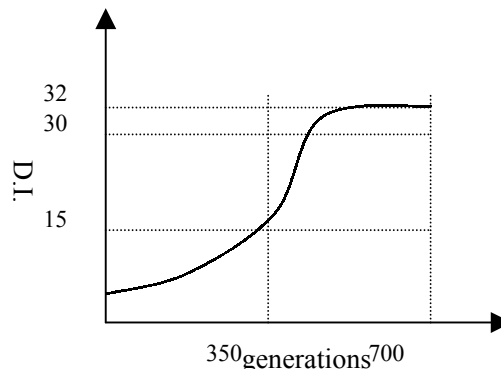


Fig. 7. Maximum D.I. values with each generation for distinguishing between 'good vine' and 'fake alcohol'.

The initial OFPs generated in the random way have been shown in Table 1 for the organization of feature parameters. Table 3 shows the optimum OFP's which are the results of the "organization of feature parameters". The "Distinction Rate (P0)" (=100%) of the optimum OFP's shown in Table 3 is much larger than those of the FP's shown in table 2. Figure 7 shows the increases in the maximum D.I. values with each GA generation. From the formulae in table3 sensor1: TGS813, sensor2: TGS821 and sensor 5: TGS882T should be selected as the best sensors for distinguishing 'fake alcohol' and 'good vine'.

Table 3. Optimum OFP for distinguishing.

Optimum OFP	D.I.	DR
$Pa_5^{-1}(Ps_1 + Ps_2)(Ps_1^{-0.5} - Ps_5 + Pa_2)$	32.5	100%

4. Conclusions

Fake alcohol is very harmful to human healthy. However the two methods for determination of methanol in commercial alcoholic drinks, i. e. fuchsine-sulfurous acid colorimetry and gas chromatography are not feasible in on-line detection, though they have high sensitive. A three gas-sensors based system was developed and the experiment was setup to detecting methanol in high concentration ethanol solution that was always found in adulteration fake alcohol. And almost all tin oxide gas sensors are high sensitive not only to methanol but also to ethanol. A new method called organization feature parameter (OFP) based on formulae expression tree by using genetic algorithm was proposed in this paper. It could solve the problem how to getting best sensors and fusion their features. For evaluating the goodness of an OFP, the "Distinction Index" (D.I.) has been defined by statistical theory. And D.I. is also used for fitness in the GA operation.

The sensor selection method discussed here has been used to distinguishing two different solutions. And the experiment demonstrated that it has good nonlinear approximation ability, short convergent time and high discrimination ratio. Although the example shown here is for distinguishing between

two states, respectively, the OFP can also for distinguishing among many odours. And the optimum OFP can be used for the “Multi-step distinguishing algorithm”. By this algorithm, the excellent OFP’s can be sequentially extracted, then the sensors can be selected, and the solutions can be sequentially distinguished.

Acknowledgements

This work is supported by the “National Natural Science Fund and 863 High-tech Fund”. We are grateful to Dr. Peng Cheng who has helped us. We also wish to thank many of their academic colleagues for many stimulating discussion in this field.

References

- [1] <http://www.chinadaily.com.cn>
- [2] <http://www.Sinocnet.com>
- [3] Julian W. Gardner, Philip N.Bartelet, *Electronic nose: Principles and Applications*, Oxford University Press 1999.14, 185-207.
- [4] Christophe S et al Potential of semiconductor sensor arrays for the origin Authentication of pure valencia orange Juices *Journal of Agricultural & Food Chemistry*, 2001(49), 3151-3160.
- [5] Sylvie Roussel, Gustaf Forsberg, etc. Optimisation of Electronic Nose Measurements. Part I: Methodology of Output Feature Selection. *Journal of Food Engineering* 37 (1998) 207-222.
- [6] Vernat-Rossi V, Garcia C.Talon R, et al., Rapid discrimination of meat products and bacterial strains using semi-conductor gas sensors, *Sensors and Actuators B*, 1996(37), 43-48.
- [7] Natale C D, Davide F A M, Amico A D, et al., Complex chemical pattern recognition with sensory array the discrimination of vintage years of wine, *Sensors and Actuators B*, 1995(24-25), 801-804.
- [8] Singh S., Hines E L, Gardner J W, et al., Fuzzy neural computing of coffee and tainted-water from an electronic nose, *Sensors and Actuators B*, 1996(30), 185-190.
- [9] Huyberechts G., Simultaneous quantification of carbon monoxide and methane in humid air using a sensor array and an artificial neural networks, *Sensors and Actuators B*, 1997(45), 123-130.
- [10] Zou X.B., Wu S.y. Evaluating the quality of cigarettes by an electronic nose system, *Journal of Testing and Evaluation*, Vol. 30, No. 6, 2002(12).
- [11] Zou X.B. Study of electronic nose on simulation inspecting the process of ferment in the brewage of vinegar (in Chinese), Master’s degree thesis, Jiangsu University, Zhenjiang, Jiangsu , China, 2001, 3.
- [12] Laurent R., Tekin K., Thomas M., etc. A comparative study of signal processing technique for clustering microsensor data (a first step towards an artificial nose), *Sensors and Actuators B*, 1997(41), 105-120.
- [13] Wang P. Artificial nose and Artificial tongue (in Chinese). science book concern. Beijing, China, 2000,12.
- [14] Peng Chen, Masam Na, Toshio T., Self-reorganization of symptom parameters in frequency domain for failure diagnosis by genetic algorithms, *Journal of Intelligent and Fuzzy Systems*, 6 (1998) 27-37.
- [15] Srinivas M., Patuaik L M, Adaptive probabilities of crossover and mutation in GA, *IEEE Trans. Syst. Manand Cybernetics*, 1994, 24(4), 656-667.

Biographies

Zou Xiaobo was born in 1975 in Hunan, China. He received the B.S., M.S. degree from the College of Biological Environment Engineering, Jiangsu University of science and Technology, Zhenjiang,

China, in 1997, 2000, respectively. His research interests include artificial olfactory system, gas sensor arrays and neural networks.

Zhao Jiewen, received his Ph.D in telecommunication engineering in 1989 from the Jiangsu University of science and Technology (China). His main areas of interest are in non-destructive analysis in foods, semiconductor chemical sensors and simulation.
