A Novel Feature Extraction Technique of Electronic Nose for Detecting of Wound Infection Based on Phase Space

Rudan Luo
College of Electronic and Information Engineering
Southwest University
Chongqing, China
e-mail: luorudan999@126.com

Pengfei Jia
College of Electronic and Information Engineering
Southwest University
Chongqing, China
e-mail:jiapengfei200609@126.com

Abstract—Rapid and timely monitoring of traumatic inflammation is conducive to doctors' diagnosis and treatment. It has been proved that electronic nose (E-nose) is an effective way to predict the bacterial classes of wound infection by smelling the odor produced by the metabolites, and it has also been found thatthe classification accuracy of E-nose is very different when different feature is extracted and put into the classifier. The gas sensor array of E-nose can be seen as a dynamic system whose response temporally evolves following the concentration of the odors. As the central concept in the analysis of dynamic systems, phase space is the first time to be employed by us to construct the feature matrix of wound infection data in this paper. Dynamic moments, the functions of time delay in phase space, is used as the feature of wound infection. The odors of four different classes of wound (wound uninfected, and infected with*P. aeruginosa*, *E. coli*and*S. aureus*) are used as the original response of E-nose.Experimental results prove that the classification accuracy of test data set is 96.43% when R2 is used as the feature, which is much better than M2P, M3P (other two dynamic moments), maximum value of the steady-state response and maximum value of the first-order derivative (two traditional feature of E-nose).

Keywords*-Wound infection; Electronic nose; Phase space; Dynamic moments*

*****

I. INTRODUCTION

Wound is a big pubic hazard in the world, and the traditional bacteriological diagnosis takes so long time to confirm infection that the treatment is often delayed. Rapid and timely monitoring of traumatic inflammation is conducive to doctors' diagnosis and treatment. But it is difficult for doctor to distinguish clinical infection in the early stage. In fact, the wounds infected by bacteria have a variety of special odor before obvious metabolitesappears.It will be significant tothe rapid diagnosis of wound infectionif this odor can be discriminated.

Electronic nose (E-nose) is anexpert systemwhich is composed of an array of gas sensors and a corresponding artificial intelligence technique. It is effective in dealing with problems analysis, and has been introduced to many fields such as food engineering [1-2], disease diagnosis [3-5]*et al.*

Previous work[6-8] hasconfirmed that it is feasible forE-nose to detect different bacteria by investigatingits odor. So E-nose is employed to detect the bacterial classes of wound infection in this paper.

As the first step of data processing, some feature should be extracted from the response of sensors to describe the pattern of odor. The classification results of E-nose are very different when different feature is extracted and put into the classifier. From a general point of view, a gas sensor can be considered as a dynamic system whose response temporally evolves following, with its proper dynamics, the concentration of the odors, and phase space is the central concept in the analysis of dynamic systems [9].

In this paper, for the first time, dynamic moments, functions of phase space, are used to construct the feature matrix of wound infection data obtained by E-nose, and support vector machine (SVM) is used as the classifier of E-nose. The rest of this paper is organized as follows, the bacteria and experiments are introduced in Section 2. PS, dynamic moments and SVM will be presented in Section 3. Different methods are used to extract the feature of wound infection data, and the classification results are analyzed in Section 4. Finally, we will give the conclusion of this paper in Section 5.

II. MATERIALS AND EXPERIMENTS

A. *Materials and preparation*

Three common bacteria, *P. aeruginosa*, *E. coli*and*S. aureus*,are used as the target in this paper. These three bacteria used in the odor sampling experiments are purchased from Chinese National Institute for the Control of Pharmaceutical and Biological Products, and the National Center for Medical Culture Collection. All species of bacteria grow in media at 37°C with shaking at 150 rpm in a gyratory incubator shaker for 24 h. The culture medium is ordinary broth medium and the main components include peptone, NaCl, beef extract and glucose. After 3 successive generations of subculture, the purchased bacteria becomes stable. Then they are inoculated into the test agar slant. Dynamic head-space method is adopted during all gas sampling experiments. The head-space gas in

ISSN: 2321-8169
21 – 25

each test slant containing the metabolic products of bacteria is imported into the E-nose for the sampling. The detail information of the three bacteria is shown in Table 1.

Bacteria	Metabolites
P. Aeruginosa	Butanol, Dimethyldisulfide, Dimethyltrisulfide, Esters, Methyl ketones, Isobutanol, Isopentanol, Isopentyl acetate, Pyruvate, Sulphurcompounds,Toluene,1-Undecene,2-Aminoacetophenone, 2-Butanone, 2-Heptanone, 2-Nonanone, 2-Undecanone
E. Coli	Acetaldehyde, Acetic acid, Aminoacetophenone, Butanediol, Decanol, Dimethyldisulfide, Dimethyltrisulfide, Dodecanol, Ethanol, Formaldehyde, Formic acid, Hydrogen sulfide, Indole, Lactic acid, Methanethiol, Methyl ketones, Octanol, Pentanols, Succinic acid, 1-Propanol
S. Aureus	Aceticacid,Aminoacetophenone,Ammonia,Ethanol, Formaldehyde, Isobutanol, Isopentyl acetate, Isopentanol, Methyl ketones, Trimethylamine, 1-Undecene, 2,5-Dimethylpyrazine isoamylamine, 2-Methylamine

TABLEI. PATHOGENS IN WOUND INFECTION AND THEIR METABOLITES

B. *E-nose for bacteria detection and measurement*

Duringthe sampling experiments,the head-space gas of the bacteria is sampled by an E-nose (Airsense-modelPEN3, Airsense Analytics, Schwerin, Germany), equipped with10different thermo-regulated (150~500°C)sensors made ofmetal oxidesemiconductors (MOS). The sensors which is positionedin a gas chamber (1.8 mL), are sensitiveto different classes of chemical compounds, and the detail information is shown in Table 2.

TABLEII. ESPONSE CHARACTERISTICS OF SENSORS IN PEN 3

Sensors	Response characteristics
W1C	aromatic
W5S	broadrange
W3C	aromatic
W6S	hydrogen
W5C	arom-aliph
W1S	broad-methane
W1W	sulphur-organic
W2S	broad-alcohol
W2W	sulph-chlor
W3S	methane-aliph

The schematic diagram of experimental system is shown in Fig.1. PEN 3 is used to sample the odor of different bacteria, and the feature extraction and pattern recognition is realized on a computer. The practical E-nose system of this paper is shown in Fig.2.

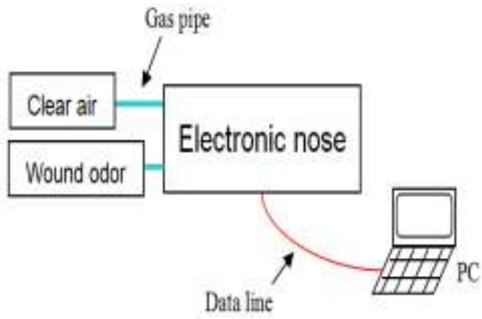

Figure 1 Schematic diagram of experimental system

Figure 2 Practical E-nose system

Before the measurement starts, the system (gas pipe andsensor array chamber) is cleaned by a cycle of 300 s, usingfiltered air (zero-air: air filtered on active carbon), to ensure theabsence of the residual odor molecules and to report sensorsto the baseline. The flow rate of the sampling gas is 300 mL/min, the sampling frequency is 1 Hz, one sampling experiment lasts 60 s. The response curves of 10 sensors on oneculture medium with *E. coli* are shown in Fig.3. One can find the obvious rise of each response curveappears when the gas containing VOCs of *E. coli* begins to pass over the sensor array, which is the result of cross sensitivity.

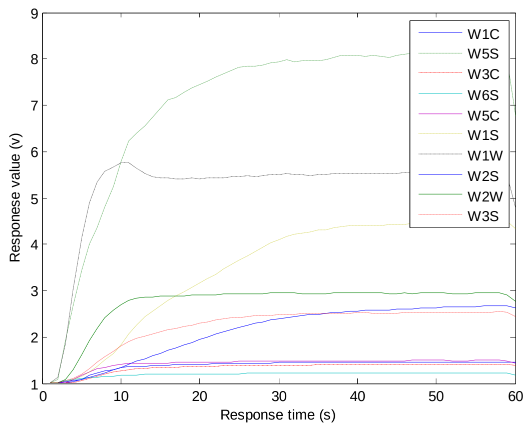

Figure 3 Response of E-nose on oneculture medium with*E. coli*

 ISSN: 2321-8169
21 – 25

III. PHASE SPACE, DYNAMIC MOMENTS AND SVM

A. *Phase space and dynamic moments*

Phase space is a central concept in the analysis of dynamics system, and its property is in the univocal correspondence between points in the space and states of the system [10]. Given a system whose state is completed described by *n* scalar variables, different states correspond to different points in a *n*-dimension vector space defined by an orthonormal basis where each direction correspond to one of the scalar variables. The fundamental property of phase space is the correspondence between each point and the instantaneous state of the system. A genetic phase space can be defined considering the Taken's embedding theorem [11]. Given an observable quantity $s(t)$ and defining a time delay τ , the space coordinates can be described as Eq. (1).

$$[s(t)\, s(t+\tau)\cdots s(t+(k-1)\tau)]. \quad (1)$$

The time evolution of $s(t)$ results in a trajectory containing the dynamic properties of the system. Trajectory assumes a large variety of shapes depending of the nature of the phenomena. The shapes of the trajectories are expected to be associated to the properties of the phenomena, so it is interesting to define some morphological description able to encode the shape of the trajectory. These morphological description can then be used to obtain information about the system dynamics.

Here, sets of morphological descriptors representing the parameters analogous to the second moments of the area of a geometrical figure in a 2-D space are considered. The functions of time delay τ in phase space are called the dynamic moments [12-13]. They are calculated considering both the coordinates and bisectors of the phase space.

Three dynamic moments, described by Eq. (2) to (4), are considered by us to represent the pattern of wound infection data.

$$\mathrm{M2P} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i \,, \quad (2)$$

$$\mathrm{R2} = \frac{1-\frac{1}{n}\sum_{i=1}^{n} x_i y_i}{\left|1+\frac{1}{n}\sum_{i=1}^{n} x_i y_i\right|} \,, \quad (3)$$

$$\mathrm{M3P} = \frac{\sqrt{2}}{2n}\sum_{i=1}^{n}(x_i^2 y_i - x_i y_i^2) \,, \quad (4)$$

where x_i and y_i are the axis of phase space, and we set the response of sensors (defined as $f(t)$) and its first-orderderivative(defined as $f'(t)$) as the axis. Each of the moments describes the different morphological feature of the trajectory.

The value of τ can influence the performance of dynamic moments, and the classification results will be very different if the value of τ is different.

B. *SVM*

SVM is known as an ideal technique for classification. It adopts structural risk minimization principle to get the best generalization ability according to the limited sample information[14]. The target of SVM is to find an optimal separating hyper-plane $y = \mathbf{w}\cdot\mathbf{x} + b$ for samples.

For linearly separable two classes' data $\{(\mathbf{x}_i, y_i)\}_{i=1}^{n}$, where $\mathbf{x}_i$ is the sample of which the label is +1 or -1, and itsdimension is *d*, y_i is the label of $\mathbf{x}_i$, and n is the number of samples in training dataset. As is shown in Fig.4, **H** is the optimal hyper-plane found by SVM to separate the samples from different classes. Multi-class problems can be achieved by constructing a multiple SVM.

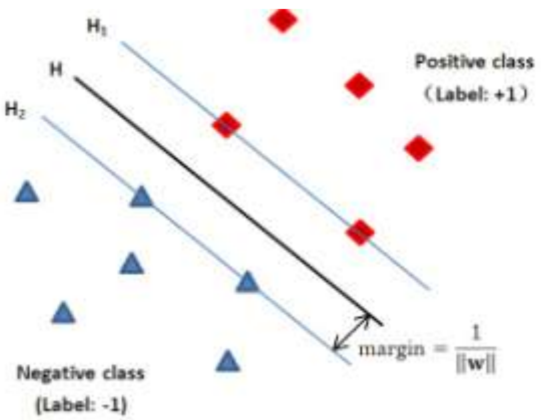

Figure 4 SVM separation of linearly separable two classes' data

The finding the separating hyper-plane can be formalized as Eq. (5).

$$\begin{aligned} &\min \frac{1}{2} < \mathbf{w}, \mathbf{w} > + C\sum_{i=1}^{n}\xi_i, \\ &\text{s.b.} < \mathbf{w}, \mathbf{x}_i > + b \ge +1 - \xi_i, \text{for } y_i = +1, \\ &\quad < \mathbf{w}, \mathbf{x}_i > + b \le -1 + \xi_i, \text{for } y_i = -1 \\ &\quad\quad \xi_i \ge 0, \forall i \end{aligned} \quad (5)$$

where ξ_i is called the slack variable, and *C* is the penalty factor. By using Lagrange multiplier techniques, Eq. (5) canbe changed to the following dual optimization problem:

$$\max \sum_{i=1}^{n} \alpha_i - \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j < \mathbf{x}_i, \mathbf{x}_j >,$$
$$\text{s.b. } \sum_{i=1}^{n} \alpha_i y_i = 0, \alpha_i = [0, C] \qquad (6)$$

Using Lagrange multipliers, the optimal desired weight vector of the discriminant hyper-plane is $w = \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i$. So the best discriminant hyper-plane can be derived as:

$$f(x) = \sum_{i=1}^{n} \alpha_i y_i < \mathbf{x}_i, \mathbf{x} > + b, \qquad (7)$$

where b is the bias of the discriminant hyper-plane.

The hyper-plane determined by Eq. (7) is linear and can solve the linearly separable classification problem. The nonlinear problem can be mapped to a new space by a nonlinear transformation with the help of the kernel function. Suppose $\Phi(\mathbf{x})$ is a map function, then the algorithm only depends on the data through dot products in high-dimensional feature space. Define k is such a kernel function as follow:

$$k(\mathbf{x}_i, \mathbf{x}_j) = < \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) > . \qquad (8)$$

In Eq. (8), the dot product in high-dimension space can be expressed as a kernel function. Similar to Eq. (7) in the linear problem, there is also a discriminant function for this nonlinear problem:

$$f(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i y_i k(\mathbf{x}_i, \mathbf{x}) + b. \qquad (9)$$

IV. RESULTS AND DISCUSSION

In this section, we will firstly research the influence of τ on classification results of E-nose, and to achieve this, we set the value of τ from 5 to 60. SVM is used to predict the classes of wound data, and its parameters are optimized by an enhanced quantum-behaved particle swarm (QPSO) [6]. Leave-one-out (LOO) technique is used to train and test the classifier. The best results of each considered dynamic moment are shown in Table 3.

TABLE I. CLASSIFICATION ACCURACY OF SVM BASED ON DIFFERENT DYNAMIC MOMENTS (%)

Feature	Training data set	Test data set
M2P	100	83.33
R2	100	96.43
M3P	100	85.71

Fig.5 and Fig.6 show the classification accuracy of training data set and test data set when the value of τ in dynamic moments is different. And it is clear that this parameter can strongly influence the performance of E-nose in distinguishing the classes of different wound infection.

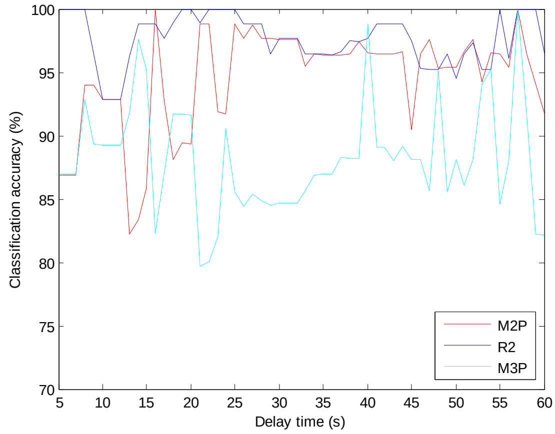

Figure 5 Classification accuracy of training data set when dynamic moments with different delay time (τ)

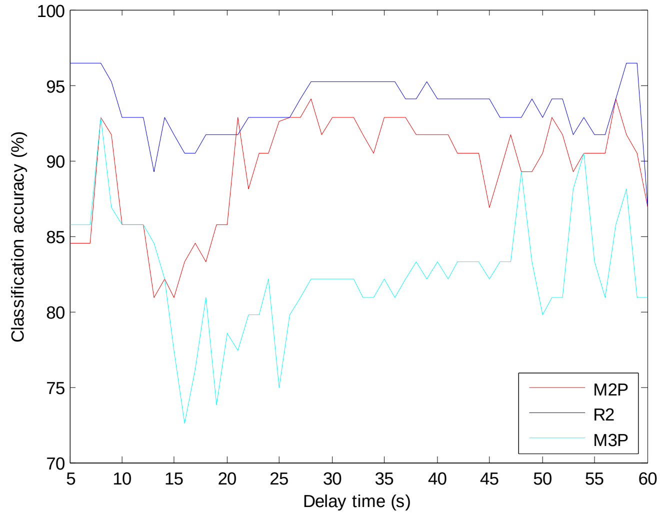

Figure 6 Classification accuracy of test data set when dynamic moments with different delay time (τ)

Besides three dynamic moments, other two traditional feature, maximum value of the steady-state response and maximum value of the first-order derivative are also used to construct the feature matrix of wound infection data.The classification results of these two feature are shown in Table 4.

TABLEIV. CLASSIFICATION ACCURACY OF SVM BASED ON

Feature	Training data set	Test data set
Maximum value of the steady-state response	100	91.67
Maximum value of the first-order derivative	97.56	94.05

DIFFERENT FEATURE EXTRACTION METHODS (%)

It can be found from Table 3 and 4, when E-nose is used to distinguish the classes of different wound infection, the accuracy of R2 is much higher than M2P, M3P and other two considered feature. The description ability of R2 is better than other two dynamic moments. To a certain extent, these results may prove that R2 contains more useful information which can help E-nose make a correct prediction.

V. CONCLUSION

It has been proved that E-nose is an effective way to detect the wound infection. As the first step of data processing, feature extraction has strongly influence on the results of E-nose in distinguishing the classes of different wound infection. The gas sensor array of E-nose can be seen as a dynamic system whose response temporally evolves following the concentration of the odors. phase space is employed to construct the feature matrix of wound infection data in this paper. Three different dynamic moments (M2P, R2 and M3P) are computed as the feature of wound infection data, the influence of time delay τ on classification accuracy is also researched. The data processing results prove that the classification accuracy of SVM is the highest when R2 is used as the feature among all considered feature.

ACKNOWLEDGMENT

REFERENCES

[1] H. Guohua, W. Lvye, M. Yanhong, Z. Lingxia, Study of grass carp (Ctenopharyngodon idellus) quality predictive model based on electronic nose, Sensors and Actuators B, 35 (2012) 301-308.

[2] E. Gobbi, M. Falasconi, G. Zambotti, V. Sberveglieri, A. Pulvirenti, G. Sberveglieri, Rapid diagnosis of Enterobacteriaceae in vegetable soups by a metal oxide sensor based electronic nose, Sensors and Actuators B, 207 Part B (2015) 1104-1113.

[3] E. A. Chapman, P. S. Thomas, E. Stone, C. Lewis, D. H. Yates, A breath test for malignant mesothelioma using an electronic nose, European Respiratory Journal, 40 (2012) 448-454.

[4] A. D'Amico, C. D. Natale, C. Falconi, E. Martinelli, R. Paolesse, G. Pennazza, M. Santonico, P. J. Sterk, U-BIOPRED study, Detection and identification of cancers by the electronic nose, Expert Opinion on Medical Diagnostics, 6 (2012) 175-185.

[5] P. Jia, F. Tian, Q. He, S. Fan, J. Liu, S. X. Yang. Feature Extraction of Wound Infection Data for Electronic Nose Based on a Novel Weighted KPCA. Sensors and Actuators B, 201 (2014), 555-566.

[6] P. Jia, F. Tian, S. Fan, Q. He, J. Feng and S. X. Yang, A novel sensor array and classifier optimization method of electronic nose based on enhanced quantum-behaved particle swarm optimization, Sensor Review, vol. 34/3, (2014), 304-311.

[7] F-C Tian., X-T Xu, Y Shen, J Yan, J-W Ma and T Liu, Detection of wound pathogen by an intelligent electronic nose, Sensors and Materials, vol. 21,(2009), 155-166.

[8] A. N. Thomas, S. Riazanskaia, W. Cheung, Y. Xu, R. Goodacre, CLP Thomas, M. S. Baguneid and A. Bayat, Novel noninvasive identification of biomarkers by analytical profiling of chronic wounds using volatile organic compounds, Wound Repair and Regeneration, vol. 18,(2010), 391-400.

[9] H. Kantz, T. Schreiber, Nonlinear time-series analysis, Cambridge University Press: Cambridge, 1997.

[10] E. Martinelli, C. Falconi, A. D'Amico, C. D. Natale, Feature extraction of chemical sensors in phase space, Sensors and Actuators B: Chemical, 95(2003) 132-139.

[11] F. Takens, Detecting strange attractors in turbulence, in: D. Rand, L. S. Young (Eds), Dynamic systems and turbulence, Warwick, 1980, 366-381.

[12] A. Fichera, C. Losenno, A. Pagano, Clustering of chaotic dynamics of a lean gas-turbine combustor, Appl. Energy, 69 (2001), 101-117.

[13] M. Annuziato, H. D. I. Abarbanel, Nonlinear dynamics for classification of multiple flow regimes, in:Proceeding of the International Conference on Soft Computing, SOCO, Genoa, 1999.

[14] V. Vapnik, The nature of statistical learning theory, Springer, New York, 1995.