Detection of Wound Pathogen
by an Intelligent Electronic Nose

Fengchun Tian[1], Xuntao Xu[1,2,*], Yue Shen[3],
Jia Yan[1], Qinghua He[3], Jianwei Ma[1] and Tao Liu[1]

[1]College of Communication Engineering, Chongqing University, Chongqing, 400044, China
[2]Mianyang Vocational and Technical College, MianYang, SiChuan, 621000, China
[3]Department of Orthopedic and Traumatic Surgery, Center for War Wound and Trauma of PLA,
Institute of Surgery Research, Daping Hospital,
Third Military Medical University, Chongqing, 400042, China

A new method of detecting wound pathogens based on an electronic nose was proposed and realized. A gas sensor array consisting of six metal oxide gas sensors and one electrochemical gas sensor was used to identify seven species of pathogens common in wound infection. By selecting the wavelet transform coefficients preferentially with a scatter matrix and using the mean of the selected coefficients as the feature, the identification accuracies of the probabilistic neural network classifier for the seven species of pathogens all reached 100%. The new feature extraction method showed high performance in the rejection of gas sensor drift. Theoretical analysis and experimental results indicate that this method can be used to accurately identify the common pathogens present in wound infection and can be further used in the real-time detection of wound infection.

1.   Introduction

Wound healing is affected by several factors, among which infection is one of the most common hindrances to wound healing. A rapid and timely identification of the bacterial type and the phase of wound infection will help physicians choose the appropriate treatment, enabling rapid wound healing. Morphological examinations using a microscope and growth test under certain conditions are usually the traditional identification methods for wound infection. However, these methods take at least more than 48 h, lack a unique authoritative criterion for the detection of certain bacteria, and require a series of sample preprocesses before the test, such as preconcentration, selective enrichment, and smear microscopy examination.

---

*Corresponding author: e-mail: myxxtao@126.com

Different bacteria, with different bacterial enzymes, have different capacities of decomposition, producing different metabolites. Moreover, because bacterial growth has four phases, namely, lag phase, log phase, stationary phase, and death phase, the volatile compound concentration of a bacterium also differs at different growth phases. Accordingly, the type and growth phase of bacteria in wound infection can be monitored by examining the volatile compound concentration around the wound.[1] There is a well-established and growing interest in the detection and identification of microorganisms by measuring the concentrations of volatile organic compounds (VOCs) they release. *In vitro* bacterial culture studies using gas chromatography/mass spectrometry (GC-MS) have identified a large number of metabolic volatile compounds including fatty acids, aliphatic alcohols, ketones, dimethyl polysulfides, and alkenes. The mean concentration of each VOC analyte is listed in Table 1 for each species and uninoculated controls after 6 h of incubation in BacT/ALERTRFA-medium-containing bottles.[2]

In medicine, biosensors are used in the assessment of drug or biomolecule concentrations in blood, serum or other corporal fluid *in vivo* or *in vitro* measurements, and data are used in diagnostic or on-line monitoring.[3] A biosensor consists of an immobilized biologic molecule (enzymes, cellules or antibodies) next to a transducer, which transforms a chemical signal into an electric signal or into other kinds of output as an optical, acoustic and heat signal when an analyte reaches it.[4] The principal advantage of enzyme biosensors is related to their high specificity and catalytic activity. However, they have the disadvantages of low stability and low possibility of reuse. An electronic nose (enose) is composed of an array of gas sensors and the corresponding pattern recognition algorithm.[5] It can imitate the olfaction system of humans and mammals and recognize odorant gases. Compared with traditional test methods, the enose has the characteristics of being noninvasive, convenient, and highly efficient, and of functioning in real time and is potentially superior in the detection of wound infection.

Table 1
Mean bacterial VOC concentrations at 6 h in aerobic medium (ppb, vol/vol).

	Control medium+blood	*P. aeruginosa*	*S. pneumoniae*	*E. coli*	*S. aureus*	*N. meningitidis*
Acetaldehyde	890	1100	5300	11000	2400	350
Acetic acid	6900	16000	460	5400	1200	870
Ethanol	1200	1500	3000	21000	800	770
Acetone	3600	6600	5100	6100	3500	7900
Ammonia	1000	1100	370	500	1800	1200
H_2S	10	40	20	4100	60	ND
Methanethiol	ND	ND	60	750	180	ND
Dimethylsulfide	690	860	4300	9600	2200	320
Dimethyldisulfide	200	330	180	430	280	360

ND: analyte not detected.

2.   Methods and Materials

2.1   *Enose for wound pathogen detection*

The detection system for wound pathogens based on the intelligent enose system comprises a wound headspace gas sampling unit, a sensor array, a signal conditioning circuit, and a data acquisition and processing unit.   The gas sensors in the array, consisting of six metal oxide gas sensors and one electrochemical gas sensor, are selected on the basis of their sensitivity to the common volatile product in wound infection.   More details on these gas sensors are given in Table 2.   To enhance the system's ability to restrain environmental interference, a temperature sensor (LM35DZ), a humidity sensor (HIH4000), and a pressure sensor (SM5552) are added into the sensor array, to simultaneously collect data of ambient parameters.   The response signals in the headspace gas of the wound obtained using the sensor array pass through the conditioning circuit and are collected and preserved in the data acquisition card.   The pattern recognition software examines the bacterial type and infection degree of the wound on the basis of the different response modes of the sensor array.

2.2   *Probabilistic neural network classifier for the enose*

The probabilistic neural network (PNN) classifier is widely used in enose owing to its fast training rate and strong ability to add new samples.   The architecture of a typical PNN consists of four layers: input layer, pattern layer, summation layer, and output layer.[6] The input layer receives training samples and distributes the input to the neurons in the pattern layer. The pattern layer calculates the probabilities of input samples belonging to a certain class. The summation layer of neurons compute the maximum likelihood of pattern X being classified in class i. Finally, the output layer obtains the estimated class of pattern X,

$$o(X) = \arg \max_{1 \leq i \leq m} \left\{ \sum_{j=1}^{N_i} \omega_{ij} \frac{1}{(2\pi\sigma^2)^{p/2}} \exp \left[-\frac{\| X - X_{ij} \|^2}{2\sigma^2} \right] \right\}, \quad i \in \{1,2,...,m\} \quad (1)$$

Table 2
Datasheet of gas sensor.

Sensor	Typical analyte
TGS2600	Air contaminants, methane, carbon monoxide, iso-butane, ethanol, hydrogen
TGS2602	VOCs and odorous gases, ammonia, H_2S, toluene, ethanol
TGS2620	Vapors of organic solvents, combustible gases, methane, CO, iso-butane, hydrogen, ethanol
TGS822	Methane, carbon monoxide, iso-butane, *n*-hexane, benzene, ethanol, acetone
QS-01	Hydrogen, CO, methane, iso-butane, ethanol, ammonia
GSBT11	VOCs, HC, smoke, organic compounds
4ETO	Ethylene oxide, ethanol, methyl-ethyl-ketone, toluene, carbon monoxide

Here, p is the dimension of the pattern vector X, m is the total number of classes in training samples, σ is the smoothing parameter, X_{ij} is the neuron vector, ω_{ij} denotes the mixture weight, and N_i denotes the total number of samples in class i.

As is known, the Bayes classifier can be expressed as

$$X \in \omega_i, \text{ if } \omega_i = \arg \max_{j=1}^{M} p(\omega_j/X).$$ (2)

Namely, the input sample X belongs to the class with the maximum a posteriori (MAP) probability.

Here, the a posteriori probability is $p(\omega_i/X) = \dfrac{p(X/\omega_i)\,p(\omega_i)}{p(X)}$. For the same input sample, $p(X)$ is a common term, so the Bayes classifier can be expressed as

$$X \in \omega_i, \text{ if } \omega_i = \arg \max_{j=1}^{M} (p(X/\omega_j)(\omega_j)).$$ (3)

With sufficient samples, eq. (4) can be used to approximate the conditional probability density of the class ω_i,[7] namely,

$$p(x/\omega_i) \approx \hat{p}(x/\omega_i) = \frac{1}{(2\pi)^{p/2}\sigma^p} \times \frac{1}{N_i} \sum_{j=1}^{N_i} \exp\left(-\frac{\|X - c_j^i\|^2}{\sigma^2}\right).$$ (4)

Here, c_j^i is the sample of class ω_i, p is the dimension of the input sample, and N_i is the number of samples of the ith pattern.

With sufficient samples, the a priori probability density of class ω_i can be approximated using

$$\hat{p}(\omega_i) = N_i/N,$$ (5)

where N is the total number of samples.

Thus, the Bayes classifier based on a priori probability and conditional probability density can be expressed as the Bayes classifier based on the Gaussian kernel function

$$X \in \omega_i, \text{ if } \omega_i = \arg \max_{k=1}^{M} \left(\frac{1}{(2\pi)^{p/2}\sigma^p} \times \frac{1}{N} \sum_{j=1}^{N_k} \exp\left(-\frac{\|X - c_j^k\|^2}{2\sigma^2}\right)\right).$$ (6)

A comparison of eqs. (1) and (6) shows that, under the condition that the a priori probabilities of all classes are equal, the PNN classifier is equivalent to the Bayes classifier based on the Gaussian kernel function, and can achieve the optimal classification with minimum error.

2.3   *Feature extraction method based on wavelet transform for the enose*

The amplitude of the response signal of the gas sensor correlates with the gas concentration. To eliminate the effect of the difference between pathogen concentrations of the same class on the qualitative identification, the original signals should be normalized using

$$S(i,j) = (S'(i,j) - \min_{\substack{i=1}}^{M} \min_{\substack{j=1}}^{N_i}(S'(i,j))) / (\max_{\substack{i=1}}^{M} \max_{\substack{j=1}}^{N_i}(S'(i,j))) - \min_{\substack{i=1}}^{M} \min_{\substack{j=1}}^{N_i}(S'(i,j))), \qquad (7)$$

where $S'(i,j)$ is the original signal, $i = 1, 2, \ldots, M$ is the class number, $j = 1, 2, \ldots$, and N_i is the number of samples in each class.

The feature extraction method in the enose includes the following: basic feature extraction method based on the original curve,[8,9] which uses the maximum value, maximum slope, and minimum slope of the original curve as features; the feature extraction method based on the fitted curve,[10] which fits the original curve to various models and takes the model parameters as features; the feature extraction method based on a transform domain,[11,12] which maps the original signals into a new space and then extracts features in the transformed domain. Considering the anti-interference and antidrift abilities, a one-dimensional (1-D) discrete wavelet transformation of the pretreated response signal was carried out. Some wavelet transform coefficients were selected preferentially with their mean used as a major feature. This method was compared with the identification effect method using the maximum of the response signal as a feature.

The discrete wavelet coefficient of an arbitrary signal $S(t) \in L^2(R)$ can be described as[13]

$$C_{j,k} = \int_{-\infty}^{\infty} S(t) \, \overline{\psi_{i,k}(t)} \, dt. \qquad (8)$$

Here, $\psi(t)$ is the wavelet function.

$$\psi_{j,k}(t) = 2^{-j/2} \psi \left(\frac{t - k \cdot 2^j}{2} \right) = 2^{-j/2} \psi(2^{-j}t - k). \qquad (9)$$

To satisfy the orthogonal requirement of feature extraction, we use the Daubechies compactly supported orthogonal asymmetric wavelet. The response signal of gas sensors is affected by various noises that have been proved to be of higher-frequency components.[14] Moreover, because of the short-term drift and long-term drift of gas sensors, some low-frequency drifting components are blended with the real signals, so the wavelet decomposition level cannot be too high or too low. With a very low decomposition level, many high-frequency components will be retained in the subband coefficients, which is not conducive to restraining noises. With a very high decomposition level, the low-frequency bands may enter the region of the sensor drift,

which is ineffective in restraining drift. With the above considerations, DB6 (with a decomposition level of 6) is selected, and the wavelet decomposition coefficients of the signals are taken as initial features.

To reduce feature dimension and increase identification accuracy, the optimal feature subset was constructed on the basis of the scatter matrix.

Let $c_j^{(i)}$ be the discrete wavelet transform coefficient for wound pathogen sample j in class i with $\{c_j^{(i)}, j = 1, 2,\dots, N, i = 1, 2, \dots, n\}$. Define the intraclass scatter matrix as

$$s_w = \sum_{i=1}^{n}\sum_{k=1}^{N_i} (c_k^{(i)} - m^{(i)})(c_k^{(i)} - m^{(i)})^T, \tag{10}$$

and the interclass scatter matrix as

$$s_b = \sum_{i=1}^{n} N_i(m^{(i)} - m_0)(m^{(i)} - m_0)^T, \tag{11}$$

Here, $m^{(i)} = \dfrac{1}{N_i}\sum_{k=1}^{N_i} c_k^{(i)}$ is the distribution center of the pathogen samples in class i, and $m_0 = \dfrac{1}{\sum_{i=1}^{n} N_i}\sum_{i=1}^{n} N_i m^{(i)}$ is the distribution center of all pathogen samples.

The intraclass scatter matrix indicates the repeatability of the feature for the same class samples, and the interclass scatter indicates the ability of the feature to distinguish samples from different classes. The constructed feature subset should have a small intraclass scatter and a large interclass scatter. Fisher distance is defined as

$$J = tr(s_b/s_w), \tag{12}$$

where tr denotes the trace of the matrix. The larger the J, the larger the interclass distance or the smaller the intraclass distance, the more accurate will be the classification. Calculate the Fisher distance of each wavelet decomposition coefficient for the initial features. Select the 31st–40th of the approximating coefficients as the optimal features. Send the mean of selected coefficients into the PNN classifier. Figure 1 presents the box plots that show the selected features and maximum features of the six classes of pathogen samples. The box plot for each class contains information on the mean, quartile value, and outliers of all samples in this class, revealing the intraclass distances of a class. Moreover, the distribution across the six box plots indicates the interclass distances between different classes. Evidently, for samples in the same class, the degree of scatter of the selected feature is smaller than that of the maximum feature, and for samples in different classes, the discrimination ability of the selected feature is stronger than that of the maximum feature.
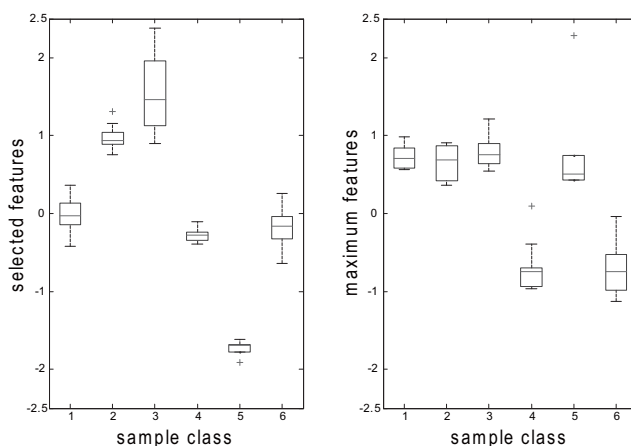
Fig. 1. Box plots for the selected features (left) and maximum features (right).

3.   Results and Discussion

3.1   *Detection of single pathogen*

Giacometti *et al*. studied 676 patients who underwent surgical treatment over a six-year period.[15] All the patients showed signs and symptoms indicating surgical wound infection. One thousand and sixty bacterial strains were isolated from 614 individuals. In particular, a single agent was identified in 271 patients, multiple agents were found in 343 patients, and no agent was identified in 62 patients. A high preponderance of aerobic bacteria was observed. Among the common pathogens were *S. aureus* (191 patients, 28.2%), *P. aeruginosa* (170 patients, 25.2%), *E. coli* (53 patients, 7.8%), *S. epidermidis* (48 patients, 7.1%), and *E. faecalis* (38 patients, 5.6%). Other studies also revealed that the common pathogens responsible for wound infection include *P. aeruginosa*, *E. coli*, *A. baumannii*, *S. aureus*, *S. epidermidis*, *K. pneumoniae*, and *S. pyogenes*.[16–18] These seven species of bacteria used in our test were purchased from the Chinese National Institute for the Control of Pharmaceutical and Biological Products, and the National Center for Medical Culture Collection. After three successive generations of subculture, the purchased bacteria became stable, and were then inoculated into agar slants. The headspace gas of bacteria in each slant was imported into the enose for the test. The test period for each sample comprises a baseline step (3 min), an adsorption step (3 min), and a desorption step (4 min). Ten minutes after each test, the next test can be started. The zero gas and sample gas pass through the enose at a rate of 50 ml/min, with a sampling frequency of 110 Hz. Each of the seven species of pathogens was cultured in two test tubes, and each tube was measured five times. We obtained 70 measurements. A typical response curve of a gas sensor to the seven species of bacteria and a response curve of the gas sensor array to one species of bacteria are shown in Figs. 2(a) and 2(b), respectively.
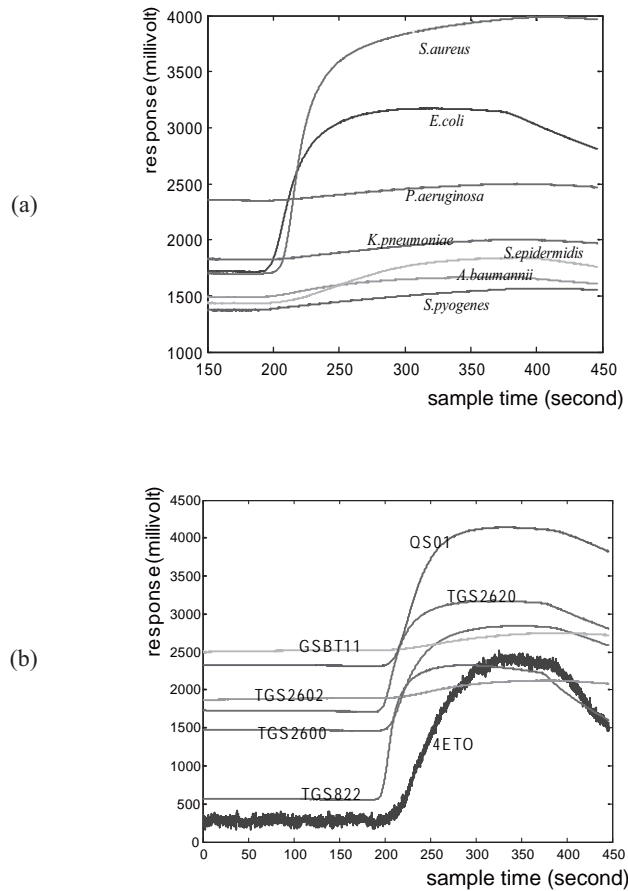
Fig. 2.   (a) Response curves of a gas sensor (TGS2620) to seven species of bacteria. (b) Response curves of gas sensors array to one species bacteria (*E. coli*).

Half of the data was randomly selected as the training set and the other half as the testing set. The identification result can be expressed as

$$result = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ \cdots & x_{ij} & \cdots & \cdots \\ x_{m1} & x_{m2} & \cdots & x_{mm} \end{bmatrix},$$

(13)

where x_{ij} denotes the number of samples of class i determined to belong to class j (the judgment is correct when $i = j$, and incorrect otherwise), and m is the total number of classes in the test set.

The identification accuracy, sensibility, and specificity of samples in class i are defined using eqs. (14)–(18). The identification results of PNN with the maximum feature and those with the selected feature are given in Tables 3 and 4, respectively.

$$Accuracy(i) = \left[\sum_{i,j=1}^{m} x_{ij} - \left(\sum_{\substack{j=1\\j\neq i}}^{m} x_{ji} + \sum_{\substack{j=1\\j\neq i}}^{m} x_{ij}\right)\right] \Big/ \sum_{i,j=1}^{m} x_{ij} \times 100\% \tag{14}$$

$$Sensitivity(i) = x_{ij} \Big/ \sum_{j=1}^{m} x_{ij} \times 100\% \tag{15}$$

$$Specificity(i) = \left[1 - \sum_{j=1,j\neq i}^{m} x_{ji} \Big/ \left(\sum_{i,j=1}^{m} x_{ij} - \sum_{j=1}^{m} x_{ij}\right)\right] \times 100\% \tag{16}$$

Table 3
Results of single-pathogen detection for seven species (with the maximum feature).

	1	2	3	4	5	6	7	Accurcy	Sensitivity	Specificity
1	5	0	0	0	0	0	0	100%	100%	100%
2	0	5	0	0	0	0	0	88.6%	100%	86.7%
3	0	2	3	0	0	0	0	94.3%	60%	100%
4	0	1	0	4	0	0	0	97.1%	80%	100%
5	0	0	0	0	5	0	0	100%	100%	100%
6	0	1	0	0	0	4	0	97.1%	80%	100%
7	0	0	0	0	0	0	5	100%	100%	100%

1, *P. aeruginosa*; 2, *E. coli*; 3, *S. aureus*; 4, *S. epidermidis*;
5, *A. baumannii*; 6, *K. pneumoniae*; 7, *S. pyogenes*;

Table 4
Results of single-pathogen detection for seven species (with the selected feature).

	1	2	3	4	5	6	7	Accurcy	Sensitivity	Specificity
1	5	0	0	0	0	0	0	100%	100%	100%
2	0	5	0	0	0	0	0	100%	100%	100%
3	0	0	5	0	0	0	0	100%	100%	100%
4	0	0	0	5	0	0	0	100%	100%	100%
5	0	0	0	0	5	0	0	100%	100%	100%
6	0	0	0	0	0	5	0	100%	100%	100%
7	0	0	0	0	0	0	5	100%	100%	100%

1–7 are the same as in Table 1.

3.2   *Detection of mixed pathogens*

More than one species of pathogens may exist in an actual wound infection.[15] The detection ability of the enose for mixed pathogens was further demonstrated in our work. We classified *S. epidermidis*, *S. pyogenes* and *A. baumannii* into one group, and *S. aureus*, *K. pneumoniae* and *P. aeruginosa* into the other group. The headspace gas of each pathogen was blended with that of the other species in the same group, and we obtained 6 samples, each of which contained two different species of pathogens mixed together. The samples were then guided into the enose for the test. Each sample was measured 7 times. We obtained 42 measurements altogether. We randomly selected 24 of these 42 measurements as training samples and the other 18 as test samples. Table 5 shows the identification results of PNN.

3.3   *Antidrift ability of enose*

The enose usually classifies the test samples according to the modes of the training samples. However, because of the short-term drift and long-term drift of the gas sensors, the training and test samples measured at different times and under different conditions will be affected by various factors, which often results in misjudgment. The drift of the gas sensors is mainly caused by changes in the temperature and humidity of the test environment, the presence of trace components, and the aging of the gas sensors. Each feature extraction method should take into consideration its ability to prevent short-term drift and long-term drift. With this in mind, the test samples in § 3.1 were affected by interfering signals of short-term drift and long-term drift using eqs. (17) and (18), respectively.[19]

$$S_{\text{short-term}}(t) = s(t) + kdt \qquad (17)$$

Here, t is the sample time, d is the drift constant (1, 5, 10, 20 and 50%, respectively), k is the mean slope of the original signal, and $k = (\max(s)-\min(s))/\text{length}(s)$.

$$S_{\text{long-term}} = (s - \bar{s})(1 - ad) + (1 + d)\bar{s} \qquad (18)$$

Here, s is the original signal, $\bar{s}$ is its mean, d is the drift constant (1, 5, 10, 20 and 50%, respectively), and $\alpha \in [0,1)$ is a constant.

Table 5
Results for detection of mixed pathogens.

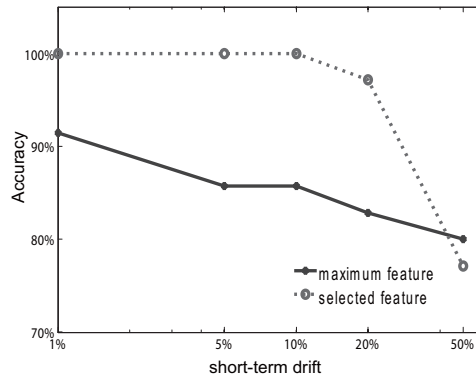|	Accuracy	Sensitivity	Specificity
S. epidermidis + S. pyogenes	94.4%	100%	93.3%
A. baumannii + S. epidermidis	100%	100%	100%
A. baumannii + S. pyogenes	100%	100%	100%
S. aureus + K. pneumoniae	100%	100%	100%
S. aureus + P. aeruginosa	100%	100%	100%
K. pneumoniae + P. aeruginosa	94.4%	66.7%	100%

Fig. 3. Ability to restrain the short-term drift for different features.
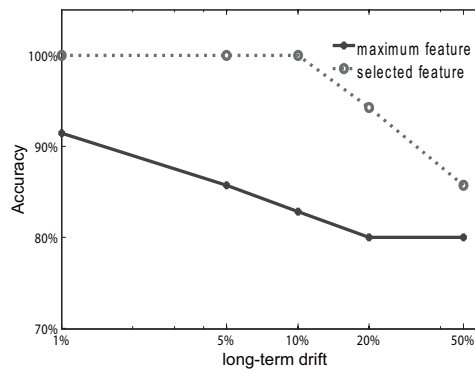

Fig. 4. Ability to restrain the long-term drift for different features.

By using the PNN classifier, the classification accuracies for the maximum feature and selected features are shown in Figs. 3 and 4, respectively. Under the influence of a strong drift, classification accuracy decreases, but the anti-drift ability of the selected feature is obviously stronger than that of the maximum feature.

4.   Conclusion

The traditional identification method for wound bacteria is laborious and time-consuming, which prevents wound infection from being diagnosed immediately. A rapid method of identifying pathogens is proposed and realized in this study. By using the gas sensor array, the rapid identification was achieved by analyzing the headspace gas of wound pathogens. Theoretical analysis and experimental results show that the newly

constructed intelligent detection system for wound pathogens, in combination with the new feature extraction method, can accurately identify the seven common pathogens present in wound infection. The identification results of single and mixed components reached 100% and 94.4%, respectively. Moreover, with a strong anti-drift ability, the system can be further used in the real-time detection of wound infection.

Acknowledgments

References

1 J. Carlsson: Appl. Environ. Microbio. **25** (1973) 287.
2 R. A. Allardyce, A. L. Hill and D. R. Murdoch: Diagn. Microbiol. Infect. Dis. **55** (2006) 255.
3 J. Liu and G. Li: Sens. Actuators, B **65** (2000) 26.
4 D. L. Garcia-Gonzalez and R. Aparicio: Grasas Y Aceites-Sevilla. **53** (2002) 96.
5 E. L. Hines, E. Llobet and J. W. Gardner: Circuits Devices Syst. **6** (1999) 297.
6 D. F. Specht: Neural Networks **3** (1990) 109.
7 E. Parzen: Ann. Math. Stat. **33** (1962) 1065.
8 S. Roussel, G. Forsberg, V. Steinmetz, P. Grenier and V. Bellon-Maurel: J. Food Eng. **37** (1998) 207.
9 E. Llobet, J. Brezmes, X. Vilanova, J. E. Sueiras and X. Correig: Sens. Actuators, B **41** (1997) 13.
10 L. Carmel, S. Levy, D. Lancet and D. Harel: Sens. Actuators, B **93** (2003) 67.
11 C. Distante, M. Leo, P. Siciliano and K. C. Persaud: Sens. Actuators, B **87** (2002) 274.
12 A. Leone, C. Distante, N. Ancona, K. C. Persaud, E. Stella and P. Siciliano: Sens. Actuators, B **105** (2005) 378.
13 I. Daubechies and B. J. Bates: J. Acoustic. Soc. Am. **93** (1993) 1671.
14 F. C. Tian, S. X. Yang and K. Dong: Sensors **5** (2005) 85.
15 A. Giacometti, O. Cirioni, A. M. Schimizzi, M. S. Del Prete, F. Barchiesi, M. M. D'Errico, E. Petrelli and G. Scalise: J. Clin. Microbiol. **38** (2000) 918.
16 A. Setkus, A. J. Galdikas, A. Kancleris, A. Olekas, D. Senulien, V. Strazdien, R. Rimdeika and R. Bagdonas: Sens. Actuators, B **115** (2006) 412.
17 S. L. Gorbach, J. W. Mayhew, J. G. Bartlett, H. Thadepalli and A. B. Onderdonk: J. Clin. Invest. **57** (1976) 478.
18 T. D. Gibson, O. Prosser, J. N. Hulbert, R. W. Marshall, P. Corcoran, P. Lowery, E. A. Ruck-Keene and S. Heron: Sens. Actuators, B **44** (1997) 413.
19 E. Llobet, J. Brezmes, R. Ionescu, X. Vilanova, S. Al-Khalifa, J. W. Gardner, N. Barsan and X. Correig: Sens. Actuators, B **83** (2002) 238.