http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 11 (7): 850-858, 2012
ISSN 1812-5638 / DOI: 10.3923/itj.2012.850.858

Background Interference Elimination in Wound Infection Detection by Electronic Nose Based on Reference Vector-based Independent Component Analysis

[1]Fengchun Tian, [1]Jia Yan, [1]Shan Xu, [1]Jingwei Feng, [2]Qinghua He, [2]Yue Shen and [1]Pengfei Jia
[1]College of Communication Engineering, Chongqing University, Chongqing, 400030, China
[2]State Key Laboratory of Trauma, Burns and Combined Injury, Institute of Surgery Research, Daping Hospital, Third Military Medical University, Chongqing, 400042, China

Abstract: Background interference is serious and widespread problem in wound infection detection by electronic nose (ENose). When mice are used as experimental subjects, the background interference, i.e., the odor of the mice themselves, is very strong and useful information is often buried in it. A new method of eliminating the background interference and detecting wound infection, based on an ENose in cooperation with reference vector-based Independent Component Analysis (ICA) denoising algorithm is proposed. It employs ICA to decompose each signal of the sensor array and extract the independent components and then discriminates the useful sources and Background interference through the Correlation with the reference Vector. The independent components of which the background interference had been eliminated are used as the inputs of Radial Basis Function (RBF) network for discrimination. The result shows that this method is effective and practical for background interference elimination in the detection of wound infection by ENose.

Key words: Electronic nose, wound infection, background interference elimination, independent component analysis, reference vector

INTRODUCTION

An electronic nose (ENose), which is composed of an array of gas sensors as well as the corresponding pattern recognition algorithm, is able to imitate the olfaction system of humans and mammals and is used for the recognition of gas and odor (Hines *et al*., 1999; Bicego *et al*., 2002; Ciosek and Wroblewski, 2006). It plays a constantly growing role in the identification and quantification of odor (Abd El-Aziz, 2011) and has become a powerful tool to detect of vapor chemicals in disease diagnostics (Turner and Magan, 2004; Gardner *et al*., 2000) such as for lung diseases (Anh *et al*., 2005; Di Natale *et al*., 2003; Phillips *et al*., 2003) diabetes (Yu *et al*., 2005; Ryabtsev *et al*., 1999), urinary tract disease (Lin *et al*., 2001; Pavlou *et al*., 2002).

Wounds are a big public hazard in the world and the incidence of two severe complications of wounds, infection and Multiple Organ Failure (MOF), is a serious concern for doctors. Rapid and timely monitoring of traumatic inflammation, especially identification of type of wound infection and infection levels of bacteria, is conducive to guiding the doctor's diagnosis and treatment. It takes long time for the traditional bacteriological diagnosis on wound infection and therefore the treatment is often delayed. The defense reaction of posttraumatic bodies contains the inflammatory response resulted from injured and broken cells, which is a sterile inflammatory process and inflammatory response caused by bacteria when secondary infection appears. Although, the two inflammatory responses are different in nature, it is difficult to distinguish clinically in early stage. It is dependent on the experience of physicians to determine whether the wound is infected before obtaining the evidence of bacteriology and immunology detection. And the two most important ways for physicians to diagnose are to observe the characters of excreta of the wound and to smell the wound odor. The bacterial infected wounds have a variety of special smell before the obvious excreta appear. If we can discriminate the special smell of different bacterial wounds, it will be conducive for rapid and timely diagnosis of wound infection. However, the smell ability of human is not sensitive enough to discriminate the odor. The ENose has much better sensitivity compared to that of human being. So the use of ENose technology to achieve a faster and easier diagnosis of wound infection is feasible. Previous works have demonstrated that it is feasible to use ENose to detect bacteria including investigation of bacterial Volatile Organic Compounds (VOCs) from cultures and also from swabs taken from wound infected patients (Gibson *et al*., 1997; Dutta *et al*., 2002; Setkus *et al*., 2006; Persaud *et al*., 2008; Thomas *et al*., 2010; Byun *et al*., 2010).

Corresponding Author: Fengchun Tian, College of Communication Engineering, Chongqing University, Chongqing, 400030, China

Independent Component Analysis (ICA) is a multivariate statistical method which is used to extract hidden components of signals, only given observed signals that are assumed to be linear mixtures of some unknown sources. The purpose of ICA is to decompose original signals into several mutually independent components, using statistics of order greater than two from the probability densities of the signals. In this process, we don't know other priori knowledge except that the sources are statistically independent. ICA developed accompanied by Blind Source Separation (BSS) problem and it is a way to solve the BSS.

In a complex background, the received signals are often mixtures of different information sources. Usually, it may appear that, we cannot obtain effective discrimination among various measurement samples by directly putting the original signals to the pattern classifier. So, in order for good prediction or classification. It is necessary to make suitable data preprocessing to obtain useful information from the data. ICA is based on the statistics independence between the information sources. Compared with the traditional filtering and cumulative average methods, it hardly damages the useful details of the signals in removing the noise; its denoising performance is usually better than the traditional filtering method. Compared with the traditional method of signal separation based on feature analysis, such as Principal Component Analysis (PCA) (Khan *et al.*, 2004) and Singular Value Decomposition (SVD), ICA is an analysis method based on the higher order statistical properties and in many applications analysis of higher order statistical properties more conforms to reality. Recently, there are many successful applications of ICA, including applications to biomedical signal processing (Ikeda and Toyama, 2000; Wubbeler *et al.*, 2000), image processing (Hyvarinen and Kosko, 2001; Luo and Boutell, 2005; Soltanizadeh and Shokouhi, 2008; Wen and Miao, 2012) and financial data analysis (Oja *et al.*, 2000). ICA approach has also successfully been used in electronic nose data (Di Natale *et al.*, 2002; Kermit and Tomic, 2003; Balasubramanian *et al.*, 2008; He *et al.*, 2008). However, they did not solve the problem that how to discriminate the useful independent source and the noise. Especially in the application of wound infection detection by ENose, ICA has not been used to eliminate the strong background interference.

INDEPENDENT COMPONENT ANALYSIS

A key point in multivariate data analysis is to find suitable representations of data. A suitable representation means a certain desirable feature of the data becomes more accessible in the further analysis using a certain transform of the original data.

Basic linear ICA model can be described as (Comon, 1994; Hyvarinen and Oja, 2000):

$$x = As \tag{1}$$

where, $s = [s_1, s_2, \ldots, s_N]^T$ is N-dimensional unknown source signa S_i ($i = 1, \ldots, N$) are mutually independent; $x = [x_1, x_2, \ldots, x_M]^T$ is the M-dimensional observed signal; A is an unknown M×N mixing matrix containing the coefficients of the mixing system. The observed signal x is linear combinations of the unknown source signal s. This statistical model is called a basic ICA model (Hyvarinen *et al.*, 2001). ICA will estimate the matrix A and source signals s in the same time, only knowing the observed signals x. It means we will find a linear transform or unmixing system B, which satisfies:

$$y = Bx \tag{2}$$

where, $y = [y_1, y_2, \ldots, y_N]^T$ and y_i ($i = 1, \ldots, N$) are as independent as possible and B is an estimate of A^{-1}. By doing this, approximations of the original source signals before the mixing are recovered. Because mixing matrix A and source signals are unknown, the solution process of ICA algorithm is not a process to find the inverse matrix of A and the separated signal y is only the approximation of the source signal. So, there are two uncertainties: (1) there is a certain proportional relationship between the separated signals and the source signals in amplitude and (2) We cannot determine the order of the separated signals. However, in many applications, most information of signals is represented in the signal waveform rather than the signal amplitude and the order. In addition, although we do not know much about the source signals, in some cases, we can discriminate them according to actual situation after separating the independent source signals. So, it is acceptable for the two uncertainties in ICA.

As mentioned before, ICA can be applied only if the source components are independent of each other. Mathematically, statistical independence is defined by the joint probability density. Provided that two random variables m_1 and m_2 are independent if and only if their joint probability function density can be decomposed following formula (Hyvarinen and Oja, 2000; Hyvarinen *et al.*, 2001):

$$p(m_1, m_2) = p_1(m_1)\, p_2(m_2) \tag{3}$$

where, $p_1(m_1)$ and $p_2(m_2)$ are the marginal probability density functions of m_1 and (m_2), respectively. This definition can be extended to the case of n random variables, in which the joint probability density is a product of marginal probability densities of n random variables. Unlike PCA which only uses the second-order statistics of the signals and describes the data in the orthogonal constraint, ICA finds out the independent hidden information, removes higher order redundant information among the components and extracts the independent sources through analyzing the higher order statistical correlation of multidimensional observed data.

In electronic nose signal analysis, a received signal from a sensor is usually a result of a combination of different gases in varying proportions. It means that the response of a sensor depends on the joint influence of a mixture of gases. We can think that the measured signal of a sensor is usually a weighted superposition of several independent components. So, the ICA is an effective decomposition method to analyze the ENose signals. In addition, the studies of physiology of olfaction suggest that the preprocessing of cognitive and perceptive information for human has a feature of removing the redundant. And ICA also shows similar characteristics in this aspect because the mutual information between individual components is the least. Therefore, if using ICA instead of PCA to ENose systems should be able to obtain better treatment results. Here we introduced the wound infection detection ENose denoising algorithm based ICA.

DENOISING ALGORITHM BASED ICA

For wound infection detection ENose, the experiment's environment and mouse body odors produced a strong background interference of sensor signals and affected the system's recognition accuracy seriously. This paper proposed a reference vector-based ICA denoising algorithm to eliminate the strong background interference in experiments.

ICA model with noise: The following formula can be considered as ICA model with noise (Hyvarinen *et al.*, 2001):

$$x = A(s+n) \tag{4}$$

where, $s = [s_1,s_2, ...,s_N]^T$ is N-dimensional unknown source signal $s_i = (i = 1, ..., N)$ are mutually independent; $x = [x_1,x_2, ...,x_M]^T$ is the M-dimensional observed signal; $n = [n_1,n_2, ...,n_N]^T$ is the N-dimensional additive noise; A is an unknown M×N mixing matrix containing the coefficients of the mixing system. The observed signal x is linear combinations of the unknown source signals. In the basic ICA model, we often assume that there are no interfering signals. But this is an ideal situation and in fact, interference signals are present and often cannot be ignored.

ICA model with noise considers the noise as an independent component of source signals and one independent component is the noise in the matrix after the ICA transformation. We can separate the mutually independent target signal and noise by performing ICA of the observed signals. However, due to the uncertainty of the ICA, the order and magnitude of the independent components after the ICA transform are uncertain. So the desired signals and noise cannot be directly distinguished in the output the ICA and it is necessary to discriminate the transformed components. So, we proposed reference vector-based ICA denoising algorithm.

Reference vector-based ICA denoising algorithm: First, we obtain independent components by performing ICA transform of the observed signals. Then we define a reference vector and distinguish the useful signals and noise by comparing the correlations between each independent component and the reference vector.

We propose two methods to define reference vector. The first one is to take the average of each observed signals as the reference vector. Because the observed signals consist mainly of the useful signals rather than noise, so that the independent component which has smallest correlation with the reference vector is regarded as the noise. Therefore, we calculate correlation coefficients between each independent component and the reference vector, respectively and distinguish the useful signals and noise by comparing the correlation coefficients. Then, we remove the noise component and use the useful signals for pattern recognition.

For our ENose data, there are fifteen-dimensional observed signals $x_1,x_2, ...,x_{15}$ and we can obtain three-dimensional independent components $y_1,y_2, ...,y_{15}$ by ICA transform. The specific algorithm is as follows: (1) Generate the reference vector $x_0 = [x_1+x_2+.....+x_{15}]/15$; (2) Calculate the correlation coefficients between $y_1,y_2, ...,y_{15}$ and the reference vector x_0, respectively and obtain CORR (y_1, x_0),..., CORR (y_2, x_0),...,CORR (y_{15}, x_0); (3) Compare the absolute values of the fifteen correlation coefficients and the independent components with greater absolute value of correlation coefficients with reference vector x_0 are identified as the responses of sensors on useful source signals, while the independent component with smallest absolute value of correlation coefficient with reference vector x_0 is considered as the response on the background interference; here, suppose that absolute value of CORR (y_{15}, x_0) is the smallest and y_{15} is the response on the background interference; (4) Remove y_{15} and obtained a fourteen-dimensional data $y_1, y_2,.....,y_{14}$ to used as inputs of classifier which are the responses of sensors on useful source signals and so, we eliminate the effect of background interference; (5) Finally, in order to make the response data be in the same order of magnitude, we normalize the remaining independent components $y_1, y_2,.....,y_{14}$ using the formula:

$$yi = \frac{y_i - y_{min}}{y_{max} - y_{min}} i = 1,....,y_{14}$$

Here, y_{max} and y_{min} are the maximum and minimum components among $y_1, y_2;.....,y_{14}$, y_i' is the normalized data of y_i = (I = 1, ..., 14) and put into the neural network classifier.

The second method to define reference vector is to directly take the measured background interference as a reference vector. Considering that the strong background interference during experiment mainly comes from the mouse body odors, we can take the sensor response of mice without wound as the reference vector. Separated independent component which is the most relevant to the reference vector is considered as the interference. The algorithm steps are similar to the method described previously and the step (4) is changed to remove the independent component which has the largest absolute value of correlation coefficient with reference vector.

For our study, there are 80 samples and after extracting the maximum of sensor response as feature we obtain an 80*15 raw data matrix. Each line represents one sample and each column represents the output of one sensor. After ICA transform is carried out for the 80*15 raw data matrix, we obtain 80*15 independent component matrix. Each column represents one independent component. Remove one dimension background interference according to reference vector-based denoising algorithm and receive 80*14 data matrix. Then we use cross validation method that divide 80 samples into k parts and k-1 parts are used for training the RBF network classifier and the rest one part is used for test.

Neural network classifier: After eliminating background interference by reference vector-based ICA denoising algorithm, we put the sensor responses into RBF neural network classifier to distinguish the wound infection types. RBF neural network is a feedforward back propagation neural network, consists of three layers: input layer, hidden layer and output layer. The network topology structure is shown in Fig. 1.

The hidden layer consists of a set of radial basis functions as the network activation function. The radial

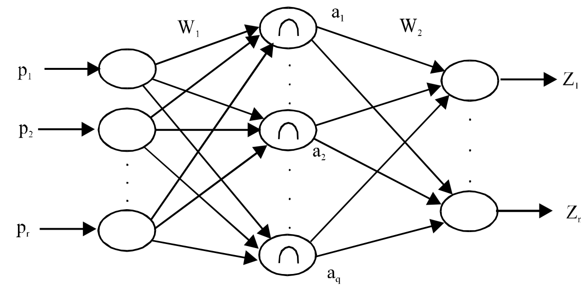

Fig. 1: Topology structure of RBF network

basis function is a Gaussian-type function. The hidden layer node calculates the Euclidean distance between the center of radial basis functions and the network input vector and then uses the result as the input of the radial basis functions. The output of radial basis functions is:

$$a_i = \exp(-\|P-C_i\|.\ b)^2 = \exp(-0.8326^2\|P-C_i\|^2/\sigma^2) \quad (5)$$

where, p is the r-dimensional input vector, C_i is the center vector of the i the hidden node, b is the bias of radial basis layer, σ is a real constant known as spread factor of radial basis function which expresses the kernel size, $\|.\|^2$ is the 2-norm of vectors. The outputs of the n nodes of output layer are linear combinations of the outputs of hidden layer nodes as:

$$z_j = b_j + \sum_{i=1}^{q} W2_{i,j} \times a_i \quad (6)$$

where, b_j is the bias of output layer, $W2_{i,j}$ is the weight between the ith hidden node and the jth output node.

RBF method is an interpolation technique in high-dimensional space, which constitutes a hidden layer space using a radial basis function as the basis function. The hidden layer transforms low-dimensional input data into high-dimensional space, making the linearly inseparable low-dimensional problem be linearly separable in high-dimensional space. For each training sample, it only corrects a small number of weights and bias. The RBF network is quite suitable for implementing multi-class and high-dimensional classification problems with advantages of convergence rate, probability of reaching global points, local sensitivity, etc., (Qasem and Shamsuddin, 2010; Mahi and Izabatene, 2011).

Gas sensor array: In this study, an array of fourteen metal oxide sensors and one electrochemical sensor was used according to the metabolites of the infected pathogens The sensors and there prime sensitive characteristics are shown in Table 1. To enhance the

Table 1: Pathogens in wound infection and their metabolites

Pathogens	Metabolites
Pseudomonas aeruginosa	Pyruvate, 2-nonanone, 2-undecanone, 2-aminoacetophenone, 1-undecene, dimethyldisulfide, esters, 2-heptanone, dimethyltrisulfide, sulphur compounds, butanol, 2-butanone, isopentanol, isobutanol, isopentyl acetate, toluene, methyl ketones
Escherichia coli	Ethanol, decanol, dodecanol, octanol, 1-propanol, indole, methyl ketones, lactic acid, succinic acid, formic acid, butanediol, dimethyldisulfide, dimethyltrisulfide, acetaldehyde, acetic acid, aminoacetophenone, pentanols, formaldehyde, hydrogen sulfide, methanethiol
Staphylococcus	Isobutanol, isopentyl acetate, 1-undecene, methyl ketones, ammonia, ethanol, 2,5-dimethylpyrazine isoamylamine, trimethylamine, *aureus* 2-methylamine, acetic acid, isopentanol, aminoacetophenone, formaldehyde

Table 2: Sensitive characteristics of gas sensors

Sensors	Sensitive characteristics
TGS800	Methane, carbon monoxide, isobutane, hydrogen, ethanol
TGS813	Methane, propane, ethanol, isobutane, hydrogen, carbon monoxide,
TGS816,	Combustible gases, methane, propane, butane, carbon monoxide, hydrogen, ethanol, isobutane,
TGS822	Organic solvent vapors, methane, carbon monoxide, isobutane, n-hexane, benzene, ethanol, acetone
TGS825	Hydrogen sulfide
TGS826	Ammonia, ethanol, isobutane, hydrogen
TGS2600	Gaseous air contaminants, methane, carbon monoxide, isobutane, ethanol, hydrogen
TGS2602	VOCs, odorous gases, ammonia, hydrogen sulfide, toluene, ethanol
TGS2620	Vapors of organic solvents, combustible gases, methane, carbon monoxide, isobutane, hydrogen, ethanol
WSP2111	Benzene, toluene, ethanol, hydrogen, formaldehyde, acetone
MQ135	Ammonia, benzene series material, acetone, carbon monoxide, ethanol, smoke
MQ138	Alcohols, aldehydes, ketones, aromatics
QS-01	VOCs, hydrogen, carbon monoxide, methane, isobutane, ethanol, ammonia
SP3S-AQ2	VOCs, methane, isobutane, carbon monoxide, hydrogen, ethanol
AQ	Carbon monoxide, methanol, ethanol, isopropanol, formaldehyde, acetaldehyde, sulfur dioxide, hydrogen, hydrogen sulfide, phenol, dimethyl ether, ethylene

ability of restraining environmental interference, a temperature sensor (LM35DZ), a humidity sensor (HIH4000) and a pressure sensor (SMI5552) are added into the sensor array; these sensors provide readings of the ambient information. The response signals of the sensor array obtained from the odor of the wound are first conditioned through a conditioning circuit and then sampled and saved in a computer via a 14-bit data acquisition card.

Sample preparation and measurement: There are five kinds of mice, wounded and uninfected, infected with *Pseudomonas aeruginosa*, *Escherichia coli*, *Staphylococcus aureus*, respectively, as well as no wounds and infection (used as background). The metabolites of the three pathogens are shown in Table 2. Each mouse has one wound in its hind leg. The mice for experiment are provided by Animal Experiment Center of Third Military Medical University. Each mouse is put into a big glass bottle with a rubber stopper. Two holes are made in the rubber stopper with two thin glass tubes inserted, respectively. One glass tube is fixed above the wound as close as possible. The VOCs of the mouse wound flow out of the bottle through the glass tube and clean air flow into the bottle through another glass tube. The dynamic headspace method is adopted during all the experiments and the process is as follows. The first stage is the baseline stage, in which the sensors are exposed to clean air for three minutes. The second stage is the response stage, which consists of a five minutes exposure of sensor to the headspace air of the wound conveyed into the sensor chamber by a pump. The third stage is the

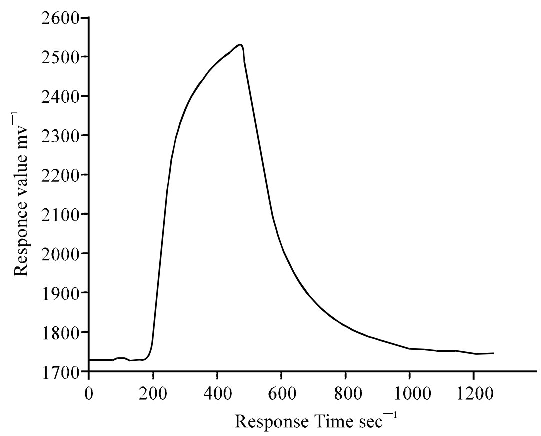

Fig. 2: Response of QS-01 on one *Staphylococcus aureus* infected mouse

recovery stage; the sensors are exposed to clean air again for fifteen minutes. At the end of each experiment, prior to the next experiment, a ten minutes purging of the sensor chamber using high-purity nitrogen is performed. The gas flow is controlled by a gas flow rate control system, which contains a rotor flowmeter, a pressure retaining valve, a steady flow valve and a needle valve. The flow rate is kept at 50 mL min^{-1}. Eighty experiments for all four kinds of mice in the same conditions are made and so 80 samples are collected. Figure 2-3 show the QS-01 sensor response on one *Staphylococcus aureus* infected mouse and background, respectively. The body odor of mouse itself is so strong that the useful wound information is buried in this background. From Fig. 2-3, we

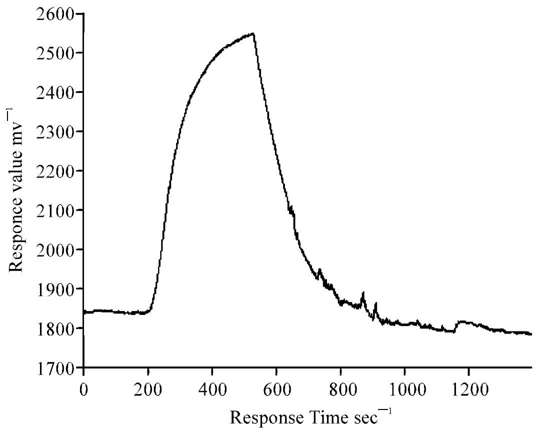

Fig. 3: Response of QS-01 on background

can see that these two signals are quite similar, the background is not a simple additive noise which is added to the signal of wounds, it also contains signal of some source which is different from the wounds odor.

RESULTS AND DISCUSSION

In order to evaluate the generalization performance, a standard statistical generalization error estimation method-cross validation, largely used for the purposes of classification, were used (Bishop, 1995).

All analysis of the data was implemented in MATLAB R2008b. Due to very fast convergence speed, the fast fixed-point iterative algorithm based on approximate negentropy, which was called Fast ICA algorithm, was used to conduct the ICA and the Fast ICA package was provided by Hyvarinen (1999). The Fast ICA algorithm was run using the symmetric approach, where in all the independent components were separated simultaneously. The non-linearity function g (u) = tanh (u) was used. Other non-linearity function was also tested but did not improve ICA performance significantly.

To demonstrate the validity of the proposed methods, the PCA, ICA and adaptive noise canceller with Normalized Least-Mean-Square (NLMS) algorithm were used as contrasts. So eight different methods are compared:

- The maximum response values of 15 sensors are used as features
- Extracting 15 principal components for the maximum response values of 15 sensors by PCA and the 15 principal components are used as features
- Extracting 3 principal components for the maximum response values of 15 sensors by PCA with cumulative variance contribution above 95% and the 3 principal components are used as features
- Extracting 5 principal components for the maximum response values of 15 sensors by PCA with cumulative variance contribution above 99% and the 5 principal components are used as features
- The maximum response values of 15 sensors which are eliminated background using adaptive noise canceller with NLMS algorithm are used as features the number of coefficients, step size and leakage factor in NLMS is 11, 0.02 and 0.6, respectively
- Extracting 15 independent components for the maximum response values of 15 sensors by ICA and the 15 independent components are used as features
- Eliminating background interference and dimension reduction for the maximum response values of 15 sensors by reference vector-based ICA denoising algorithm, wherein reference vector is the average of each observed signals
- Eliminating background interference and dimension reduction for the maximum response values of 15 sensors by reference vector-based ICA denoising algorithm, wherein reference vector is the response of background interference

Table 3: Comparison among four different methods

Method	Dimension	Kernel size	Classification rate (%)
Original response	15	4.9	85
PCA	15	4.2	92.50
PCA with cumulative variance contribution above 95%	03	4.3	62.50
PCA with cumulative variance Contribution above 99%	05	0.5	65
adaptive noise canceller	15	9.2	87.50
ICA	15	1.9	93.75
Reference vector-based ICA 1	14	2.6	95.0
Reference vector-based ICA 2	14	6.9	96.25

The plot of the classification rate with the different k value in k-fold cross validation is shown in Fig. 4. It is obvious that the performance changes with the varying k value. For the same k value, we also can obtain different results because of dividing data into different k parts. For a particular division, we can obtain a special classification rate. A leave-one-out (LOO) procedure test, which can be viewed as an extreme form of k-fold cross validation in which k is equal to the number of examples, was performed to validate model robustness. In our case, k is equal to 80. In LOO cross validation, all sample data of each class are used for training except one, which is left

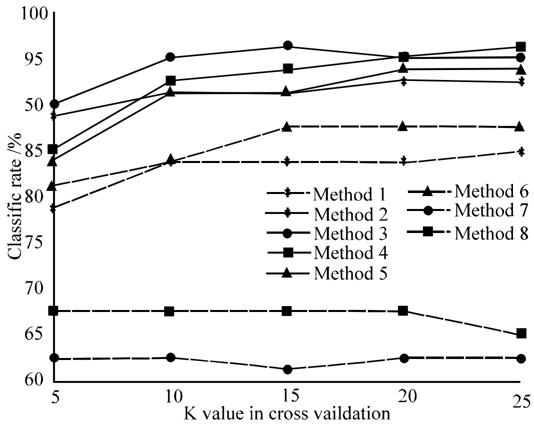

Fig. 4: Classification rate with different k value

for testing. The classification results using LOO and the parameters values of RBF neural network classifier of each method are given in Table 3.

From Table 3 it can be noticed that, decomposition of the raw data by PCA can remove the correlation between variables and improve the classifying performance. After dimension reduction, the classification rates reduced greatly, though the cumulative variance contribution is above 95 and 99%, respectively. It means that the principal components with high variance contribution do not necessarily benefit to classification. Eliminating background using adaptive noise canceller with NLMS algorithm can receive better classification result than the original response method and obtain 87.5% classification rate. It means that adaptive noise canceller can eliminate background interference to a certain extent. Extracting independent components of the raw data without dimension reduction operation by ICA has higher classification results than the former methods. It shows that, for wound infection detection ENose data, the ICA can decompose and DeNoise the original data better, extract the hidden information more beneficial to classification in the original data. However, only using ICA decomposition, we cannot discriminate independent useful source signals and background interference. So, we introduced two reference vector-based ICA denoising algorithms and the two proposed methods can achieve 95 and 96.25% correct recognition rate, respectively, which are much better than the other denoising methods.

Preprocessing the data of gas sensor array by ICA can realize redundancy eliminating and decorrelation of samples effectively and find independent useful source signals and noise. Using the two proposed methods, especially directly taking the measured background noise as a reference vector, we can discriminate the source signals and noise, eliminate the background interference and improve pattern recognition accuracy greatly.

CONCLUSION

In this study, a new method of detecting wound infection, based on an electronic nose (ENose) and reference vector-based ICA denoising algorithm, is proposed. The two proposed reference vector-based ICA denoising algorithms achieve 95 and 96.25% classification accuracy for mouse wound infection detection, respectively. They are better than that of PCA, PCA with cumulative variance contribution above 95%, PCA with cumulative variance contribution above 99%, adaptive noise canceller and the traditional ICA, which obtain 92.5, 62.5, 65, 87.5 and 93.75% classification accuracy, respectively. The results show that reference vector-based ICA denoising algorithm, especially directly taking the measured background interference as a reference vector, can discriminate the source signals and the interference effectively, eliminate the background interference and improve pattern recognition accuracy greatly for wound infection detection based ENose. Traditional PCA and adaptive noise canceller methods are not effective to process the complex ENose signals of wound infection detection. In summary, this method is a useful tool for classification for the case of strong background interference and an increasingly accessible technology for real time, accurate and fast detection of wound infection.

ACKNOWLEDGMENTS

This study is supported by China Postdoctoral Science Foundation funded project ((Project No: 20090461445), the Natural Science Foundation Project of CQ CSTC, 2009, Peoples Republic of China (CSTC, 2009BA2021), the Clinical Research Fund of Third Military Medical University (2007XG077), Innovative Talent Training Project, the Third Stage of “211 Project”, Chongqing University (Project No: S-09102) and the Fundamental Research Funds for the Central Universities (Project No. CDJXS10160001).

Fengchun Tian and Jia Yan developed the concept and all the work is under their direction. Jia Yan, Jingwei Feng and Pengfei Jia developed and designed experiments, accomplished hardware design. Lianqing Fu and Shan Xu accomplished software design.

Qinghua He and Yue Shen developed the concept, provided experiment samples and medical guidance.

REFERENCES

Abd El-Aziz, S.E., 2011. Control strategies of stored product pests. J. Entomol., 8: 101-122.

Anh, D.T.V., W. Olthuis and P. Bergveld, 2005. A hydrogen peroxide sensor for exhaled breath measurement. Sens. Actuators B: Chem., 111: 494-499.

Balasubramanian, S., S. Panigrahi, C.M. Logue, C. Doetkott, M. Marchello and J.S. Sherwood, 2008. Independent component analysis-processed electronic nose data for predicting *Salmonella typhimurium* populations in contaminated beef. Food Control, 19: 236-246.

Bicego, M., G. Tessari, G. Tecchilli and M. Bettinelli, 2002. A comparative analysis of basic pattern recognition techniques for the development of small size electronic nose. Sens. Actuators B: Chem., 85: 137-144.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, Oxford, UK., ISBN: 9780198538646.

Byun, H.G., K.C. Persaud and A.M. Pisanelli, 2010. Wound-state monitoring for burn patients using E-nose/SPME system. ETRI J., 32: 440-446.

Ciosek, P. and W. Wroblewski, 2006. The analysis of sensor array data with various pattern recognition techniques. Sens. Actuators B: Chem., 114: 85-93.

Comon, P., 1994. Independent component analysis: A new concept. Signal Process., 36: 287-314.

Di Natale, C., A. Macagnano, E. Martinelli, R. Paolesse and G. D`Arcangelo *et al*., 2003. Lung cancer identification by the analysis of breath by means of an array of non-selective gas sensors. Biosens. Bioelectron., 18: 1209-1218.

Di Natale, C., E. Martinelli and A. D'Amico, 2002. Counteraction of environmental disturbances of electronic nose data by independent component analysis. Sens. Actuators B, 82: 158-165.

Dutta, R., E.L. Hines, J.W. Gardner and P. Boilot, 2002. Bacteria classification using Cyranose 320 electronic nose. Biomed. Eng., Vol. 1. 10.1186/1475-925X-1-4.

Gardner, J.W., H.W. Shin and E.L. Hines, 2000. An electronic nose system to diagnose illness. Sensors Actuators B, 70: 19-24.

Gibson, T.D., O. Prosser, J.N. Hulbert, R.W. Marshall, P. Corcoran *et al*., 1997. Detection and simultaneous identification of microorganisms from headspace samples using an electronic nose. Sens. Actuators B, 44: 413-422.

He, X., S. Wei and R. Wang, 2008. Independent component analysis and neural network applied on electronic nose system. Proceedings of the International conference on Bioinformatics and Biomedical Engineering, May 16-18, 2008, Shanghai, China, pp: 490-493.

Hines, E.L., E. Llobet and J.W. Gardner, 1999. Electronic noses: A review of signal processing techniques. Proc. IEE Circuits Devices Syst., 146: 297-310.

Hyvarinen, A. and B. Kosko, 2001. Image Denoising by Sparse Code Shrinkage. Wiley-IEEE Press, New York, ISBN: 9780470544976, pp: 554-568.

Hyvarinen, A. and E. Oja, 2000. Independent component analysis: Algorithms and applications. Neural Networks, 13: 411-430.

Hyvarinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Networks, 10: 626-634.

Hyvarinen, A., J. Karhunen and E. Oja, 2001. Independent Component Analysis. Wiley, New York, ISBN: 9780471405405.

Ikeda, S. and K. Toyama, 2000. Independent component analysis for noisy data-MEG data analysis. Neural Networks, 13: 1063-1074.

Kermit, M. and O. Tomic, 2003. Independent component analysis applied on gas sensor array measurement data. IEEE Sens. J., 3: 218-228.

Khan, M.A.U., M.K. Khan, M.A. Khan, M.T. Ibrahim, M.K. Ahmed and J.A. Baig, 2004. Principal component analysis of directional images for face recognition. Inform. Technol. J., 3: 290-295.

Lin, Y.J., H.R. Guo, Y.H. Chang, M.T. Kao, H.H. Wang and R.I. Hong, 2001. Application of the electronic nose for uremia diagnosis. Sens. Actuators B: Chem., 76: 177-180.

Luo, J. and M. Boutell, 2005. Natural scene classification using overcomplete ICA. Pattern Recognit., 38: 1507-1519.

Mahi, H. and H.F. Izabatene, 2011. Segmentation of satellite imagery using RBF neural network and genetic algorithm. Asian J. Applied Sci., 4: 186-194.

Oja, E., K. Kiviluoto and S. Malaroiu, 2000. Independent component analysis for financial time series. Proceedings of the IEEE Adaptive Systems for Signal Processing, Communications and Control Symposium AS-SPCC., October 1-4, 2000, Lake Louise, AB Canada, pp: 111-116.

Pavlou, A.K., N. Magan, C. McNulty, J. Jones, D. Sharp, J. Brown and A.P. Turner, 2002. Use of an electronic nose system for diagnoses of urinary tract infections. Biosens. Bioelectron., 17: 893-899.

Persaud, K.C., A.M. Pisanelli, A. Bailey, K. Dunn and M. Falasconi *et al*., 2008. Identification of wound infection by limited set of volatile products. Proceedings of the IEEE International Conference Sensors, October 26-29, 2008, Lecce Italy, pp: 1375-1377.

Phillips, M., R.N. Cataneo, A.R.C. Cummin, A.J. Gagliardi and K. Gleeson *et al*., 2003. Detection of lung cancer with volatile markers in the breath. Chest, 123: 2115-2123.

Qasem, S.N. and S.M. Shamsuddin, 2010. Generalization improvement of radial basis function network based on multi-objective particle swarm optimization. J. Artif. Intell., 3: 1-16.

Ryabtsev, S.V., A.V. Shaposhnick, A.N. Lukin and E.P. Domashevskaya, 1999. Application of semiconductor gas sensors for medical diagnostics. Sens. Actuators B: Chem., 59: 26-29.

Setkus, A., A.J. Galdikas, Z.A. Kancleris, A. Olekas and D. Senuliene *et al*., 2006. Featuring of bacterial contamination of wounds by dynamic response of SnO2 gas sensor array. Sens. Actuators B, 115: 412-420.

Soltanizadeh, H. and S.B. Shokouhi, 2008. Feature extraction and classification of objects in the rosette pattern using component analysis and neural network. J. Applied Sci., 8: 4088-4096.

Thomas, A.N., S. Riazanskaia, W. Cheung, Y. Xu and R. Goodacre *et al*., 2010. Novel noninvasive identification of biomarkers by analytical profiling of chronic wounds using volatile organic compounds. Wound Rep. Reg., 18: 391-400.

Turner, A.P.F. and N. Magan, 2004. Electronic noses and disease diagnostics. Nat. Rev. Microbiol., 2: 161-166.

Wen, X.L. and M.X. Miao, 2012. Research and simulation on blind separation algorithm in mixing image based on ICA. Inform. Technol. J., 11: 456-460.

Wubbeler, G., A. Ziehe, B. Mackert, M. Muller, K.R.L. Trahms and C. Curio, 2000. Independent component analysis of noninvasively recorded cortical magnetic DC-fields in humans. IEEE Trans. Biol. Med. Eng., 47: 564-599.

Yu, J.B., H.G. Byun, M.S. So and J.S. Huh, 2005. Analysis of diabetic patient`s breath with conducting polymer sensor array. Sens. Actuators B: Chem., 108: 305-308.