Background Elimination in Wound Infection by Electronic Nose Based on Adaptive Noise Cancellation System and Genetic Algorithm

Yalin Yu
College of Electronic and Information Engineering
Southwest University
Chongqing, China
e-mail: ningling61@126.com

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

***Abstract*—**when electronic nose (E-nose) is used to diagnose the wound infection of mice, there exists strong background interference which reduces the accuracy of the E-nose. An adaptive noise cancellation system (ANCs) based on adaptive filters is introduced into denoising the wound infection data, and the genetic algorithm (GA) is used to reduce the redundancy of the Multi-channel data. The method using ANCs combined with GA has achieved good performance, and the accuracy of LRLS-GA is 100% with less calculated amount.

Keywords*-Genetic algorithm; Adaptive noise cancellation; Background elimination; Electronic nose; Wound infection*

*****

I. INTRODUCTION

Wound infection is very serious in clinics. Every year many stretalher cases die of the infection of wound. So it's important to detect inflammation reaction of wounds quickly, including the type of bacteria and the extent of the wound infection.

The type and growth phase of bacteria in wound infection can be monitored by examining the volatile organic compounds (VOCs) around the wound. So it's a feasible measure to use electronic nose (E-nose) to detect wound infection. The application of E-nose to identify microorganisms has received widespread attention. Compared with traditional methods, such as gas chromatography-mass spectrometry (GC-MS), the E-nose is noninvasive, convenient, and highly efficient. So it is potentially superior in the detection of wound infection.

There is a strong background interference in animal experiments in which the wound of mice is infected with bacteria so as to affect the discrimination of different types of infection. Most useful information is buried in the background, and the background is the worst impediment to obtain good discrimination results. So it is important to find a useful method to remove the background interference.

Adaptive noise cancellation system (ANCs) is a very effective method of removing strong background. The system can automatically adjust adaptive filter parameters to meet some criterion requirements so as to achieve the optimal filtering on condition that the statistical characteristics input signals are unknown. It can obtain higher extraction effect when good reference signal is provided.

In addition, there are 15 sensors in our E-nose system, and crossover response exits among different channels, increasing the redundancy of the sample information. To some extent, this will reduce the final recognition rate. Some measures must be employed to find the optimized sensor combination.

The fact that genetic algorithm (GA) is a guided search method that is capable of carrying out optimization in conjunction with any simulation model has allowed it to be used in a wide range of areas, such as engineering, robotics, classification, et al.

In this paper, four kinds of adaptive cancellation algorithms are used to remove the background among the VOCs information, and then a ANCs-GA method is introduced to the E-nose signal processing to achieve higher recognition rate of wound bacterial infection.

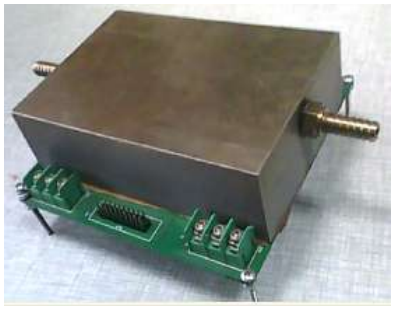
II. MATERIALS AND METHODS

A.Materials

The mice which are used as the subjects of wound infection are roughly the same in varieties, posture characteristics, and health condition before wound infection. The mice are provided by Third Military Medical University Article Affiliated Hospital. In the experiment the mice are divided into four groups. There is no bacteria in the first group mice wound, and the wounds of the other three components are respectively infected by pseudomonas aeruginosa, escherichia coli and golden glucose aureus.

B.Sensor array

In the construction of a gas sensor array, fourteen metal oxide sensors and one electrochemical sensor are selected. They are nine TGS sensors made in Japan (Figaro Engineering Inc.), TGS-826, TGS-813, TGS-825, TGS-800, TGS-816, TGS-2620, TGS-822,TGS-2602, TGS-2600, one XSC sensor made in China (New Creators Electronic Technology Co. Ltd), WSP-2111, two MQ sensors made in China (Winsen Electronics Technology Co. Ltd), MQ-138, MQ-135, one QS sensor made in Japan, QS-01, one FIS sensor made in Japan (FIS Inc.), SP3S-AQ2, and one electrochemical sensor made in England (Dart Sensors Ltd), AQ sensor. The gas sensor array is placed in a stainless steel test chamber. A 32-channel and 14-bit high precision data acquisition system (DAS) is employed for fifteen gas sensors, a humidity sensor, a temperature sensor and a pressure sensor. Heater voltage of each sensor is 5±0.05 V, and the circuit voltage is also 5±0.01 V. Sensor array is placed in a stainless steel chamber of which the volume is 240ml, experimental use 32 channel, 14 data acquisition system. A USB data acquisition card is employed in the experiment, and the frequency of sampling is 200Hz. The chamber and sensor array are shown in Figure 1.



(a) Chamber

(b) Array of sensors

Figure 1 Chamber and array of sensors

C.Experiment method

Dynamic headspace method is adopted in the experiments. Under the action of the pump, the gas is inhaled through the chamber. There are three stages in each experiment, in the first stage which lasts 3 minutes the pure air is accessed through the sensor array, then from the third minute to the eighth, the headspace VOCs is made to go through the chamber. The last 15 minutes again the pure air is accessed into the sensor array in the purpose of clearing the sensor. The time between two adjacent experiments is 5 minutes. The flow velocity of the sampling gas is controlled by float-type flowmeter of which the flow velocity is 50 ml/min. 14 bit USB data acquisition card is used in the experiments. The sample cycle is 0.1 second.

For each mouse, the sampling experiment proceed for five times, so each group of mice which are infected by the same bacteria can produce twenty samples. Meanwhile a healthy mouse is also sampled by the E-nose under the same experimental environment, which is used as the reference signal of the ANCs. The flowchart of the whole experiment is shown in Figure 2.

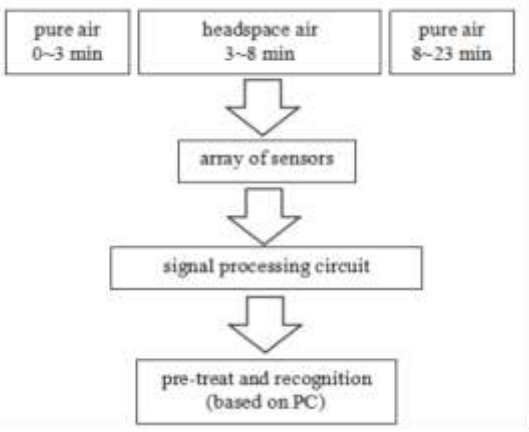


Figure 2 Flowchart of experiment

D. Theory of ANCs

As is shown in Figure 3, the signal D received from the source is the main channel signal, which consists of the useful signal S and the noise V. The input of reference channel is X which is uncorrelated with S. The relationship between V and X is linear correlation. It is necessary for ANCs to work that X is related to V while independent on S. Y Is the output of the adaptive filter while e = D − Y is error signal which guides the iteration of the filter in order to make Y approach V. According to the optimum criterion, e is approximately equal to D when Y mostly approaches V, thus the ANCs realizes the reconstruction of the useful signal S.

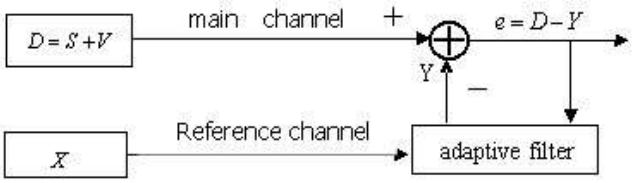


Figure 3 Block diagram of ANCs

E. Adaptive noise cancellation algorithms

The ANC algorithms used in this research are least mean square (LMS), normalized least mean square (NLMS), recursive least square (RLS) and lattice recursive least square (LRLS).LMS and RLS are traditional algorithms in which there are some shortcomings,so two corresponding improvements are introduced into the research,namely NLMS and LRLS.

1 LMS

Coefficient update formula (the step of filter is) :

$$W(n+1) = W(n) + 2\mu e(n)X(n-k), k = 1,2,..., N \qquad (1)$$

In which, error signal $e(n) = D(n) - Y(n)$; $X(n)$ is the reference channel's input of ANCs while $D(n)$ is the input of main channel; $Y(n) = W^T(n)X(n)$, is the output of the adaptive filter; μ is the step parameters of the adaptive filter which is used to control the stability and convergence rate. To ensure the stability of the adaptive process, μ needs to be $0 < \mu < 2/NP_{in}$, $P_{in} = E[X^2(n)]$.

LMS based on gradient search is a classic self-adaptive algorithm. During each iteration the calculated amount of addition and multiplication is only (2N+1), but its convergent performance is poor, and the gradient noise is magnified. In order to overcome the shortcoming of the algorithm, an improved LMS is introduced into the experiment.

2 NLMS

Coefficient update formula:

$$W(n+1) = W(n) + \mu_n e(n)X(n)/[P_x(n) + \alpha] \quad (2)$$

In which, $P_x(n) = X^T(n)X(n)$, is the power estimation of $X(n)$; $0 < \mu_n < 2$, α is a constant.

As an improved LMS,NLMS can effectively overcome the gradient noise amplification of traditional LMS algorithm. Its amount of calculation is as small as traditional LMS,and it has a good performance in stability and convergence.

3 RLS

Coefficient update formula:

$$W(n+1)=W(n)+g(n+1)e(n+1) \quad (3)$$

$$g(n+1)=P(n)X(n+1)/[\lambda+X^T(n+1)P(n)X(n)] \quad (4)$$

$$P(n+1)=\lambda^{-1}[P(n)-g(n+1)X^T(n+1)P(n) \quad (5)$$

In which, $g(n)$ is gain vector; $P(n)$ is the inverse matrixof $R_{XX}(n)$ which is the autocorrelation matrix of $X(n)$; λ which is a constant is the forget factor of RLS, and $0<\lambda<1$.

RLS is also based on transverse filter and its performance in stability and convergence is better than LMS. But as a recursive algorithm, the calculated amount of addition and multiplication is $2.5N^2+4N$ for each iteration. The process is complicated, which will limit its application to some extent. So an improved RLS algorithm which has less calculated amount is introduced into the experiment.

4 LRLS

Forward prediction error:

$$f_{i+1,T}=f_{i,T}-k^b_{i+1,T}b_{i,T-1} \quad (6)$$

Backward prediction error:

$$b_{i+1,T}=b_{i,T-1}-k^f_{i+1,T}f_{i,T} \quad (7)$$

And,forward reflection coefficient $k^f_{i+1,T}=\dfrac{\Delta_{i+1,T}}{\sigma^f_{i,T}}$, backward reflection coefficient $k^b_{i+1,T}=\dfrac{\Delta_{i+1,T}}{\sigma^b_{i,T-1}}$, partial autocorrelation function : $\Delta_{i+1,T}=\lambda\Delta_{i+1,T-1}+\dfrac{b_{i,T-1}f_{i,T}}{1-\gamma_{i-1,T-1}}$,

Likelihood variables: $\gamma_{i,T}=\gamma_{i-1,T}+\dfrac{b^2_{i,T}}{\sigma^b_{i,T}}$,forward prediction error method : $\sigma^f_{i,T}$, backward prediction error method : $\sigma^b_{i,T}$.

when $T\le P$,

$$\sigma^f_{i+1}=\sigma^f_{i,T}-k^b_{i+1,T}\Delta_{i+1,T} \quad (8)$$

$$\sigma^b_{i+1,T}=\sigma^b_{i,T-1}-k^f_{i+1,T}\Delta_{i+1,T} \quad (9)$$

Others,

$$\sigma^f_{i+1,T}=\lambda\sigma^f_{i+1,T-1}+\frac{f^2_{i+1,T}}{1-\gamma_{i,T-1}} \quad (10)$$

$$\sigma^b_{i+1,T}=\lambda\sigma^b_{i+1,T-1}+\frac{b^2_{i+!,T}}{1-\gamma_{i,T}} \quad (11)$$

LRLS is based on lattice filters. It constantly revises itself for each new input dataso as to achieve the least square accurately. The calculated amount of LRLS is 3~6 times than LMS's,and much smaller than RLS's, so the algorithm is also called as a fast-RLS. Meanwhile its speed of convergence is fast,and the numerical character of the coefficient is good.

F.GA

In this paper, GA uses real coding to represent the decision variable, as all the optimization problems considered have really valued decision variables. Tournament selection with a tournament size of two has been used. A crossover function is single-point, and an adaptive feasible mutation operator has been adopted.

III. EXPERIMENTS AND THE RESULTS

A.Experimental operation

The source signal and noise have been achieved by the method which is mentioned in Section 2.C. There are six groups during which the methods used to pre-treat the source signal is different. In the first group, the signal is unfiltered, while in Group 2, 3, 4, 5, ANCs with different algorithms is employed to remove the background of the source signal. In order to reduce dimensions of the data, feature extraction is introduced to the processing. And in this paper the maximum of the sensor's response curve is chosen as the eigenvalue, so a [80*15] feature matrix is obtained, in which the first twenty rows are the data of the wound with no infection, the data from Row 21 to Row 40 are from the wound infected with pseudomonas aeruginosa, the data from Row 41 to Row 60 are from the wound infected with escherichia coli and the last twenty data are from the wound infected with golden glucose aureus. Then the matrix is normalized, and the normalizing method is shown in Equation 1.

$$y'_{i,j}=\frac{y_{i,j}-y^{\min j}}{y^{\max j}-y^{\min j}} \quad (13)$$

In which, $y'_{i,j}$ is the normalized matrix of y , $y^{\min j}$ is the minimum of column j ,and $y^{\max j}$ is the maximum of column j .

Finally the data of eachwound which is infected by different bacteria is divided into two sets according to the line number. The row with even serial numbers is the train set of radial basis function (RBF) neural network,and the row with odd serial numbers is used to test the RBF. The flow diagram of the processing is shown as Figure 4.

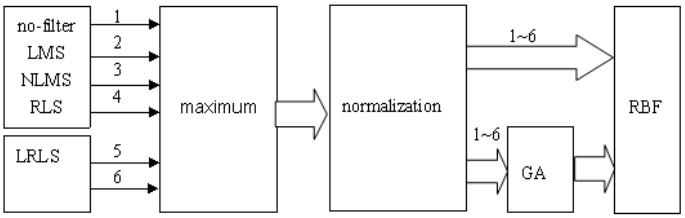


Figure 4Flow diagram of the processing

B.The results

In order to compare the filtering effect of different self-adaptive algorithms, each datum of wound infection adopt LMS, NLMS, RLS, and LRSL to reduce noise. The parameter of each self-adaptive algorithm employs partial factor analysis method in many experiments. The classification and identification for each algorithm is shown in Table 1.

Table 1 Classification results of four types wound infection based on RBF

	1	2	3	4
1	8	1	0	1
2	0	9	1	0
3	2	0	7	1
4	0	0	1	9

(a)

	1	2	3	4
1	10	0	0	0
2	2	8	0	0
3	0	2	8	0
4	0	1	1	8

(b)

	1	2	3	4
1	8	1	0	1
2	0	9	1	0
3	2	0	7	1
4	0	0	1	9

(c)

	1	2	3	4
1	8	1	0	1
2	0	9	1	0
3	2	0	7	1
4	0	0	1	9

(d)

Table 1 (a) corresponds to RBF classification results which are concluded only through feature extraction and normalizing method instead of filtering.

(b), (c) and (d) of Table 1 respectively correspond to RBF classification results which are concluded through NMS, RLS, LRLS algorithm, feature extraction and normalizing method.

The row vector and the column vector represent four wound types. $y_{i,j}$ is the amount of the result that the wound is classified into the j wound. And each algorithm gets the equation of the total recognition rate of the four wound types based on correct classification:

$accuracy = \frac{\sum x_{i,j}}{40}$, $\sum x_{i,j}$ is the sum of the four wound types' correct classification, 40 is the sample amount of test set. This algorithm is used to calculate the RBF neural network's recognition rate.

The parameter settings of table (a), (b), (c), (d) corresponds to the no-filter, NLMS, RLS, and LRLS of Table 2.

Table 2 shows the accuracy of the five different pre-treat algorithms corresponding to the four wound types.

Table 2 Recognition rate of RBF of different pre-treat method

method	order	step/forget factor	leakage	goal	spread	Recognition rate(%)
no-filter	-	-	-	-	-	82.5
no-filter+GA	-	-	-	-	-	90
LMS	-	-	-	-	-	-
NLMS	13	0.07	0.6	0.05	13	87.5
NLMS+GA	13	0.07	0.6	0.05	13	92.5
RLS	13	0.3	-	0.01	18	92.5
RLS-GA	13	0.3	-	0.01	18	100
LRLS	13	0.3	-	0.01	18	92.5
LRLS-GA	13	0.3	-	0.01	18	100

Recognition rate is the number of correctly classified samples of four bacterial species accounted for the total number of 40 samples, the calculating equation is shown in Table 1.

Leakage is a leakage factor. When it is less than 1, NLMS can reach the least error then iterate to improve the numerical performance of the algorithm.

Step factor corresponds to the LMS algorithm and the NLMS algorithm. Forget factor corresponds to the RLS algorithm and the LRLS algorithm.

Goal mse and spread are respectively RBF neural network's expected error and expansion coefficient.

It can be seen from (a) in Table 1 that without the recognition of the RBF sensor's removing the background the response data of the four types of wounds were not correctly classified cases. And the wrong number of the third kind of wound type even reaches 3. The table (b) corresponds to the classification and identification results after the NLMS algorithm managing. The right numbers are two more than the table (a). The first wound type doesn't has wrong classification. The table (c) and (d) correspond to the RLS algorithm and the LRLS algorithm, and the number of correct classification has reached 37, accounting for 92.5% of a total of 40 test samples, thus effectively improving the classification accuracy. The overall performance of the algorithms of in the classification of the four types of wounds are shown in Table 2. The third line of Table 2 is the LMS algorithm. When using the LMS algorithm to denoise the e-nose wound data, the cancellation is not convergent. After many repetitions and trying to choose different parameters settings, the cancellation is still divergent, which ruled out the possibility of chance and inappropriate parameter setting. Finally analyzed, that may be due to the too large change of the amplitude of the sensor response data. Nevertheless, the NLMS, RLS and LRLS algorithms are all stable but without divergence in the filtering process.

IV. CONCLUSION

As is shown in Table 2, the four ANCs algorithms all do well to improving the recognition accuracy of RBF except LMS, which is divergent because the amplitude of the initial data is large. And the method which combines ANCs with GAhas achieved higher recognition rate than the ones just use ANCs especially the LRLS-GA with 100%.

middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert

figures and tables after they are cited in the text. Use the abbreviation "Fig. 1", even at the beginning of a sentence.

ACKNOWLEDGMENT

The work is supported by Program for New Century Excellent Talents in University (No. [2013] 47), National Natural Science Foundation of China (Nos. 61372139, 61101233, 60972155), China Postdoctoral Science Foundation (No. 2016M602630) and Fundamental Research Funds for the Central Universities (No. XDJK2015C073).

REFERENCES

[1] Roussel S, Forsberg G, Steinmetz V, Grenier P, Bellon-Maurel V. Optimisation of electronic nose measurements. Part I : Methodology of output feature selection [J]. Journal of Food Engineering, 1998, 37: 207.

[2] Zhang S, Xie C, Zeng D, Zhang Q, Li H, Bi Z. A feature extraction method and a sampling system for fast recognition of flammable liquids with a portable E-nose [J]. Sensors and Actuators B, 2007, 124: 437.

[3] [3] Llobet E, Brezmes J, Vilanova X, Sueiras JE, Correig X. Qualitative and quantitative analysis of volatile organic compounds using transient and steady-state responses of a thick-film tin oxide gas sensor array [J]. Sensors and Actuatori B, 1997, 41: 13

[4] Feng J, Tian F, Yan J, He Q, Shen Y, Pan L. A background elimination method based on wavelet transform in wound infection detection by electronic nose [J]. Sensors & Actuators B Chemical, 2011, 157 (2): 395-400.

[5] Fabian C E, Narasimhan D, Buff E D. Multi-modal detection of surgical sponges and implements: US, US20060241396[P]. 2006.

[6] Yang L L. Signal detection and interference cancellation based on simplified matrix inversion for CDMA applications: US, US5872776 [P]. 1999.

[7] Roberts J B. Method of and apparatus for the elimination of the effects of internal interference in force measurement systems, including touch - input computer and related displays employing touch force location measurement techniques: US, US5563632 [P]. 1996.

[8] Berenz JJ, Mciver GW, Niesen JW, Dunbridge B, Shreve GA. Optimized human presence detection through elimination of background interference: US, US 6810135 B1 [P]. 2004.

[9] Macioszek J A, Robinson J M. Method for elimination of rheumatoid factor interference in diagnostic assays: US, US5698393 [P]. 1997.

[10] Hunt GC, Larus JR, Detreville JD, Jones MB, Chilimbi TA. Inter-process interference elimination: US, US 7694300 B2 [P]. 2010.