Sensors & Actuators: B. Chemical 298 (2019) 126926

Contents lists available at ScienceDirect

Sensors and Actuators B: Chemical

journal homepage: www.elsevier.com/locate/snb

Local warning integrated with global feature based on dynamic spectra for FAIMS data analysis in detection of clinical wound infection

Tong Sun[a,c,1], FengChun Tian[a], YuTian Bi[b,1], XiaoZheng Zhong[b], Jiao He[a], TaiCong Yang[a], QingShan Guo[b], Ying Lei[b], YanYi Lu[b], Lin Zeng[b], QingHua He[b,*]

[a] School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China
[b] State Key Laboratory of Trauma, Burns and Combined Injury, Daping Hospital, Surgery Institute of the Third Military Medical University, Chongqing 400042, China
[c] School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China

ARTICLE INFO

Keywords:
Wound infection
Machine learning
Escherichia coli
Field Asymmetric Ion Mobility Spectrometry (FAIMS)

ABSTRACT

Infections have long been a thorny problem that severely threatened public health and resulted in tremendous economic losses worldwide. Current detection methods for wound infection do not fully meet the requirements of preventing and treating this disease. Therefore, people are looking for better alternatives, wherein FAIMS (Field Asymmetric Ion Mobility Spectrometry) technology, by virtue of its high sensitivity, rapid response and noninvasive operation, is a promising candidate. This paper aims to investigate the possibility of FAIMS technology in detecting wound infections quickly and accurately. For this purpose, we gathered an odor dataset of clinical wound samples with the employment of a FAIMS instrument, the Lonestar (Owlstone, UK) analyzer. To enhance detectability, we proposed a novel algorithm framework, i.e., Local Warning integrated with Global Feature (LWGF), which is verified on distinguishment between twenty patients with single or mixed infection of *Escherichia coli* (*E. coli*) and six wounded patients without infection. Experimental results showed that the LWGF successfully identified the patients with the best average AUC of 0.98, and the best recognition rate of 96.15%, which are much higher than the conventional methods.

1. Introduction

Trauma has become the leading cause of morbidity and mortality worldwide, which severely threatens public health and the global economy [1]. About one in ten mortalities is caused by a traumatic injury, which gives rise to more than 5.8 million deaths per year [2]. Along with this, it is hard to avoid microbial infections, which occur in not only trauma but also a variety of injuries such as diabetic foot and cancers [3,4]. Various infections have been significantly threatening human health [5–8]. Though the majority of infections are found through visual inspection, however it cannot determine the exact agent, nor the infection if it is at an early stage. Therefore, rapid and effective detection of infectious agents is a necessary prerequisite for subsequent treatment. However prevalent clinical detection methods for infections (e.g., microbiological culture, serological diagnosis, molecular biological testing) usually suffer from shortcomings such as expensive cost, time-consuming, bad portability, invasive injury, requirement of specialists and materials, which cannot meet current medical demands, especially for mass wounded diagnosis occurred in war and disaster. Statistics show that empirical antibiotic treatment was often administered [9], which is untargeted, resulting in a three-fold increase in mortality compared to targeted antibiotic treatment [10]. Therefore, people have been striving for new and better alternatives to meet the demands of infection prevention, diagnosis, and monitoring.

As bacteria produce specific volatile organic compounds (VOCs) and/or gases, which can be utilized for the detection of bacterial infection [11]. Different infected and uninfected wounds generate unique metabolites including known chemicals such as alcohol, aldehydes, acids, esters, ketones and unknowns, part of which results in infection-specific VOCs and/or gases [12]. Therefore it is able to determine wound infections with employment of gas phase detection technology such as gas chromatography/mass spectrometer (GC/MS), electronic nose (E-nose) and ion mobility spectrometry (IMS). GC/MS has been widely adopted in the chemical determination of volatile bacterial products [13–21]. While it is not suitable for mass screening of wound infection due to its high cost of use and time, and bad portability, etc. E-nose, which mimics the biological olfactory system, has also been used for bacteria discrimination for decades [22–29]. While E-nose usually

* Corresponding author.
E-mail address: qinghuahe_paper@126.com (Q. He).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.snb.2019.126926
Received 25 June 2019; Received in revised form 30 July 2019; Accepted 31 July 2019
Available online 31 July 2019

suffers from poor-sensitivity (ppm level), drift, aging, etc., among which the poor-sensitivity is an obvious obstacle for infection detection, hence people tried to elevate detection limit of E-nose by improving performance of the hardware and algorithm. Persaud K.C. and Byun et al. [22,23] used an electronic nose incorporating an automated solid-phase micro extraction (SPME) desorption system to achieve wound-state monitoring; Xun-Tao et al. [24] used a gas condensation unit to improve the limit of E-nose. Guo et al. [25] attempted to solve this problem in the aspect of feature extraction and model optimization. Besides of the two technology, IMS has also been used in bacteria detection for more than twenty years [30–36]. Because IMS possesses the characteristics of high sensitivity, quickness of response, and availability for miniaturization, it has also become an option for wound infection detection. As an evolution of IMS, FAIMS [37] technology inherits the merit of IMS and can detect and quantify a wide range of chemical compounds directly from a flow of gas or from the headspace of liquid or solid samples.

This study aims to combine FAIMS technology and machine learning to investigate the potential of which as a rapid screening approach for clinical wound infection via wound odors. FAIMS and other similar technologies have been extensively used in the areas of food and health care, and the researchers have developed and applied various algorithms to fully exploit the potential of the technologies. When detecting, FAIMS generates a kind of map called dispersion field (DF) spectrum, where each element is the current intensity, and the vertical dimension is the DF intensity (varies from 0% to 100% of its max) and the horizontal dimension is the compensation voltage (CV, usually from -6 V to 6 V), so each row of the spectrum is a response curve at a DF intensity. To our knowledge, current ways to process the spectrum are based on the employment of several stable spectra, where either the entire spectrum (global), or a specific area or only one DF curve (local) is used to extract the feature. For example, Sinha et al. [38–40] assessed the applicability of FAIMS towards detection of early infections in the stored agricultural products such as potatoes and onions. They used a specific area (such as 47%–77% of DF and -1.30 V to -0.90 V of CV) of two or three stable spectra which were selected from the collected spectra, and then extracted the max response value of each DF curve in the area as the feature. Kontunen et al. [41] detected the surgical smoke of various tissue types by differential ion mobility spectrometry (DMS) to find the possible biomarkers for cancer or bacterial infection, where one spectrum was collected and the whole was flattened to a vector as the feature. Rutolo et al. detected soft rot disease (commonly for potatoes) by FAIMS, and collected two spectra for a sample. They only used one DF curve of each spectrum for analysis because they believed that the more DF curves yielded no improved outcome when undertaking data analysis. Arasaradnam et al. [42,43] detected inflammatory bowel disease and irritable bowel syndrome by the Lonestar (Owlstone, UK), respectively, where one stable spectrum was collected for each sample and the whole was used to extract the feature. In our opinion, perhaps using a few stable spectra is not a big problem if the sample is stable enough. However, if the sample is unstable, such as in the presence of sample degradation, then merely using the stable spectra is not the proper way to make analysis, because key information may hide in those unstable spectra, and so it is when either use a whole or specific part of one spectrum.

The traditional spectra or concentration profiles have been widely used for data analysis in chemometrics, which are effective and reliable. And for the other hand, the image processing algorithms are also widely used in this field. We consider that the two methods belong to different technical routes. Nevertheless, in our opinion, they share the same essence: measurement, statistics, inference, and verification. In this paper, we employ the image processing algorithm to solve the odor measurement problem. Specifically, we proposed a novel algorithm framework for the DF spectra analysis, i.e., Local Warning Integrated with Global Feature (LWGF), which exploits both the local and global information based on a series of dynamic spectra.

2. Material and methods

2.1. Ethics approval

This work was conducted in accordance with the ethical principles of Good Clinical Practice and the Declaration of Helsinki. The Medical Ethics Committee of Daping Hospital of The Third Military University (Chongqing, China) approved the protocol before the commencement of the study based on informed consent of all patients involved in the study.

2.2. Collection of wound infection samples

Human wound samples were obtained from the trauma center and intensive care unit of Daping Hospital of The Third Military University. Wound samples were collected during surgery or postoperative care, where the types include excision of necrotic muscle or tissue, secretion, drainage fluid. Collection processes neither disturbed the scheduled treatment plan of the patients nor added extra pain and burden to them. The collected sample was divided into at least two identical specimens, where the number of the specimen depended the quantity of the sample (in most cases, we acquired three specimens). One specimen was immediately sent to the clinical laboratory of the hospital to conduct official infection test. The rest one(s) was analyzed by the Lonestar, where most of them were firstly pre-cultured in mercaptoacetate broth for 12–20 h and a few were directly tested. As a preliminary study, our primary goal is to investigate the feasibility of FAIMS technology on the detection of wound infection. Therefore, we planned to culture most of the clinical samples at initial stage of this study, because it is the routine process involved in many traditional bacteria tests.

2.3. Brief introduction of FAIMS technology and the acquisition of the spectra

The detection process of FAIMS technology can be divided into ionization, separation and detection [37]. When volatile sample is introduced via a carrier gas, it is firstly ionized into either positive or negative ions which continuously flow into the electrode channel, where a high intensity radio frequency electric field ED (dispersion field) is applied (Lonestar using a Ni-63 source). As the fact that charged ions exhibit different mobility under high intensity electric field, these ions pushing by a flow advance on different trajectories. Handled by a DC voltage EC (compensation voltage), the detector simultaneously measures the current intensity of positive and negative ions which just pass the channel without being annihilated. With parallel control of ED and EC, two dispersion field pictures (positive and negative characteristic spectra) are generated. The typical sensitivity of FAIMS is believed to be lower than ppm level in practice, which depends on the detection condition, such as the carrier gas and operating temperature. FAIMS can detect a wide range of chemicals including aromatic amines, amines, phosphorous compounds, sulfoxides, ketones, esters, alkenes, alcohols, aromatics, water, and etc.

The data acquisition set up consists of the Lonestar portable analyzer, ATLAS^{TM} sample kit and a zero gas generator (Leman Instrument, France). The sampling process was conducted under common set [42,43], but we used a 10 ml glass vial to place our sample into the sampler. We repeated scanning a specimen for about 30 spectra (one spectra includes a positive one and a negative one), and the scanning stopped when the experiment operator found the spectra turned to be too weak by visual inspection. With the resolution setting of 51 × 512, one spectra can be obtained within 3 min. Due to the difference in specimen quantity of each patient, so the numbers of collected spectra of each patient were not the same. According to the official test results, the patients were categorized into three groups: single infection group, mixed infection group, and control group (uninfected with open wound) and we obtained a clinical wound sample database of 60

Table 1
Subject characteristics of the clinical wound odor database.

Index	Infection type	No. of subjects	Range of age	Mean of age	No. of male	No. of female
1	**Control group**	**6**	**24-71**	**54.5**	**3**	**3**
	Single Infection Group (13 groups)	**32**	**16-68**	**42.3**	**26**	**6**
2	***Escherichia coli***	10	21-65	45.3	9	1
3	*Acinetobacter baumannii*	1	28	28	1	0
4	*Klebsiella pneumoniae*	3	17-55	29.7	3	0
5	*Citrobacter freundii*	3	49-65	59.7	1	2
6	*Staphylococcus aureus*	3	41-54	49.7	3	0
7	*Pseudomonas aeruginosa*	3	33-45	37.3	3	0
8	*Enterococcus hirae*	1	16	16	0	1
9	*Bacillus cereus*	2	28-32	30	1	1
10	*Enterococcus durans*	1	16	16	0	1
11	*Proteus mirabilis*	1	51	51	1	0
12	*Cupriavidus paucula*	1	60	60	1	0
13	*Corynebacterium striatum*	1	68	68	1	0
14	*Enterobacter cloacae*	2	56-63	59.5	2	0
	Mixed infection group (18 groups)	**22**	**13-63**	**32.2**	**20**	**2**
15	*Acinetobacter baumannii, Burkholderia cepacia*	1	42	42	1	0
16	***Acinetobacter baumannii, Escherichia coli***	3	13-44	33.7	3	0
17	*Acinetobacter baumannii, Enterococcus faecium*	1	20	20	1	0
18	*Acinetobacter baumannii, Pseudomonas aeruginosa*	1	27	27	1	0
19	***Aeromonas hydrophila, Escherichia coli***	1	31	31	0	1
20	*Alcaligenes xylosoxidans, Pseudomonas aeruginosa*	1	27	27	1	0
21	***Escherichia coli, Enterococcus faecalis***	1	13	13	1	0
22	***Escherichia coli, Proteus mirabilis***	1	34	34	1	0
23	***Escherichia coli, Staphylococcus aureus***	1	15	15	1	0
24	*Enterobacter cloacae, Klebsiella pneumoniae,*	1	28	28	1	0
25	*Enterobacter cloacae, Pseudomonas aeruginosa,*	1	13	13	1	0
26	*Pseudomonas aeruginosa, Klebsiella pneumoniae*	3	17-48	36.7	3	0
27	*Pseudomonas aeruginosa, Enterobacter aerogenes*	1	63	63	1	0
28	*Pseudomonas aeruginosa, Staphylococcus aureus*	1	40	40	1	0
29	*Staphylococcus aureus, Corynebacterium striatum*	1	45	45	1	0
30	***Escherichia coli, Acinetobacter baumannii, Pseudomonas aeruginosa***	1	27	27	1	0
31	***Escherichia coli, Enterococcus faecium, Pseudomonas aeruginosa***	1	36	36	0	1
32	***Escherichia coli, Enterococcus faecalis, Aeromonas hydrophila***	1	48	48	1	0

patients shown in Table 1.

2.4. Machine learning task and validation protocol

A machine learning task requires sufficient independent identically distributed (i.i.d.) instances to train a model which is used to predict the new instance. As in our database, the frequency of *E. coli*-related infection was much higher than that of the others, therefore we set the task to determine whether a test patient was infected with *E. coli* (single or mixed infection). Specifically, we included the control group as the negative category (6 patients), and the *E. coli*-related groups as the positive category (20 patients), which are marked with underline in Table 2. The reason to include the mixed-infection groups of *E. coli* was to increase the number of validation patients. For the validation protocol, we adopted the cross-validation set of hold one out and real independent test, i.e., for each test, one patient's spectra were predicted and validated by the model trained by others'. In our formulation, all the original spectra (dynamic spectra) rather than one or two stable spectra were adopted to learn the model.

2.5. LWGF algorithm framework

2.5.1. Data pre-processing and global processing module

The basic strategy of the LWGF is to jointly exploit the information of different spacial levels, where data analysis are parallel processed on each level and then the outputs of which are integrated to make final decision. Herein, we gave a paradigm based on two spacial levels, i.e., the global level (whole spectrum) and local level (a block). As shown in Fig. 1, the training process contains 4 modules. In module 1, the original positive and negative spectra (each with size of 51*512) subtracted the mean value of the red box area on the top left corner to remove background noise. Then, we connected the two the green boxes (each with size of 36*256) end to end, because the response of significance would not occur beyond the area, so we obtained one spectrum with size of 72*256 after pre-processing.

In module 2, the common image feature (e.g., LBP) F_I was extracted from the whole spectra, and then it was used to train a machine learning model where its parameter feature was furtherly extracted to form the global parametric feature F_G. In our application, the Least Squares Support Vector Machines (LS-SVM) model was trained on the image feature F_I. After that, we extracted the latent variable as the feature of one spectra, where the latent variable was a real value parameter indicating the distance to hyperplane of the SVM model. However, since the spectra amounts of each patient were different, so their feature dimension were different as well. There are some methods to unify the feature dimension such as down sampling and histogram statistic, nevertheless just using the mean value F_G of the latent variables of a patient could solve the problem well.

2.5.2. Local warning module

The local warning of module 3 (shown in the subfigure of Fig. 1) started with an image partition, where the whole spectra was divided into blocks with small size, e.g., of 4*4 in our case, and so we had 1152 blocks. Secondly, we computed the mean response value (MRV) on a block by

$$\mathrm{MRV}(b) = \frac{1}{N_b} \sum_{i \in A_b} \sum_{j \in A_b} I(i,j) \tag{1}$$

where b is the index of block, $I(i, j)$ the response intensity of pixel (i, j), A_b the area, N_b is the number of the pixels. Thirdly, we excluded the block with very low absolute value of mean MRV. In our case, the block with the value lower than 0.03 was eliminated (about 750 blocks in our case), where the distribution of the mean MRV on each block is shown in Fig. 2. Then, the feature data on each block were clustered. DBSCAN

Table 2
Specific information of the negative and the positive category.

Subject index	Category	Group index	Age	Sex	Injury	Sample type	No. of spectra	Infection level of *E. coli*
N1	Negative	1 (uninfected)	46	Male	Crush injury	Exudate culture	61	–
N2	Negative	1 (uninfected)	24	Male	Explosive wound	Cortex	70	–
N3	Negative	1 (uninfected)	47	Male	Open fracture	Tissue culture	39	–
N4	Negative	1 (uninfected)	68	Female	Hip infection	Secretion	25	–
N5	Negative	1 (uninfected)	71	Female	Bedsore	Tissue culture	46	–
N6	Negative	1 (uninfected)	71	Female	Bedsore	Pyogenic fluids culture	43	–
P1	Positive	2 (single *E. coli*)	53	Male	Right calf chronic infection	Pyogenic fluids culture	27	+ + + +
P2	Positive	2 (single *E. coli*)	45	Male	Left thigh infection	Pyogenic fluids culture	29	+ +
P3	Positive	2 (single *E. coli*)	45	Male	Left thigh infection	Drainage fluid culture	25	+ +
P4	Positive	2 (single *E. coli*)	65	Female	Hip crush injury	Drainage fluid culture	62	+ +
P5	Positive	2 (single *E. coli*)	21	Male	Left leg crush injury	Pyogenic fluids culture	43	+
P6	Positive	2 (single *E. coli*)	39	Male	Right arm open fracture	Pyogenic fluids culture	82	+
P7	Positive	2 (single *E. coli*)	39	Male	Right arm open fracture	Pyogenic fluids culture	29	+ +
P8	Positive	2 (single *E. coli*)	61	Male	Rectal tumor	Pyogenic fluids culture	34	+ + +
P9	Positive	2 (single *E. coli*)	23	Male	Appendicitis	Pyogenic fluids culture	36	+ + +
P10	Positive	2 (single *E. coli*)	62	Male	Rectal cancer	Drainage fluid culture	34	+ + +
P11	Positive	16 (mixed *E. coli*)	44	Male	Ulcer of the left foot	Pyogenic fluids culture	121	+
P12	Positive	16 (mixed *E. coli*)	27	Male	Pelvic fractures	Pyogenic fluids	85	+ + +
P13	Positive	16 (mixed *E. coli*)	13	Male	Pelvic fractures	Necrotic muscle	71	+ + +
P14	Positive	19 (mixed *E. coli*)	31	Female	Open fracture of left tibiofibula	Tissue culture	77	+ +
P15	Positive	21 (mixed *E. coli*)	13	Male	Pelvic fractures	Necrotic muscle	66	+
P16	Positive	22 (mixed *E. coli*)	34	Male	Rectal cancer	Pyogenic fluids culture	118	+
P17	Positive	23 (mixed *E. coli*)	15	Male	High falling	Pyogenic fluids and necrotic tissue culture	68	+ +
P18	Positive	30 (mixed *E. coli*)	27	Male	Pelvic fractures	Necrotic muscle and pyogenic blood	74	+ + +
P19	Positive	31 (mixed *E. coli*)	36	Female	Pelvic fractures	Necrotic muscle	43	+ +
P20	Positive	32 (mixed *E. coli*)	48	Male	Pelvic fractures	Necrotic muscle	59	+ + +

*Infection level are official results from clinical laboratory department.

[44] is a popular clustering technology without need to specify the number of clusters which is suitable for our task. However if set the Eps (epsilon: the threshold of cluster distance of DBSCAN) to be a hard value in the usual way, we found the clustering effect would be very bad because the intervals of the MRV feature sequence varied greatly on different blocks and patients. To deal with this problem, we presented a method (listed on Table 3) to determine soft Eps adaptively, where the Grubbs's test [45] was iteratively used to delete the outlying interval. As shown in Fig. 3, where one row represents a sequence of the MRV feature of a patient, if we roughly set a hard value for the Eps such as 0.1 or 0.3, it could not cluster well on both of the two blocks, but the soft Eps method is able to find good cluster on the blocks simultaneously.

Next, for each patient, we were to find the valid block which was able to identify the patient based on the corresponding cluster. Here, we proposed a method named Interactive Discrimination based on

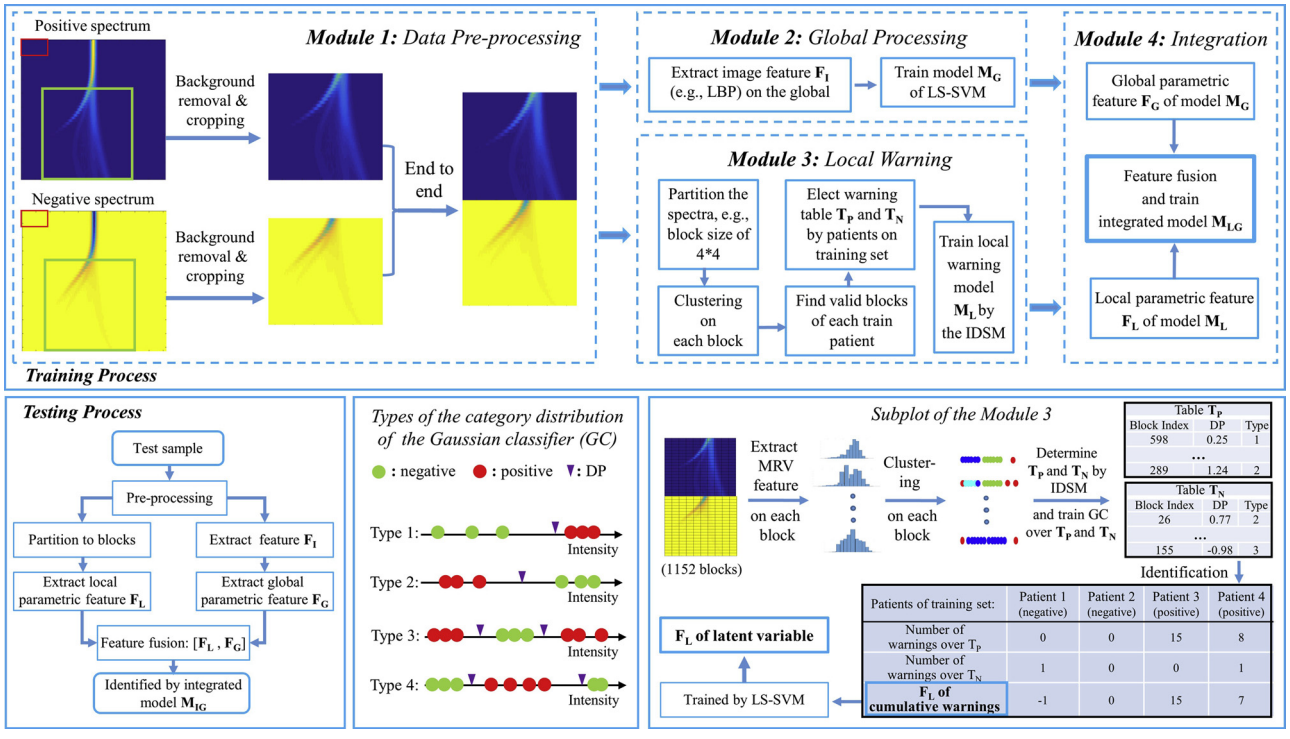

Fig. 1. Overview diagram of the proposed LWGF.

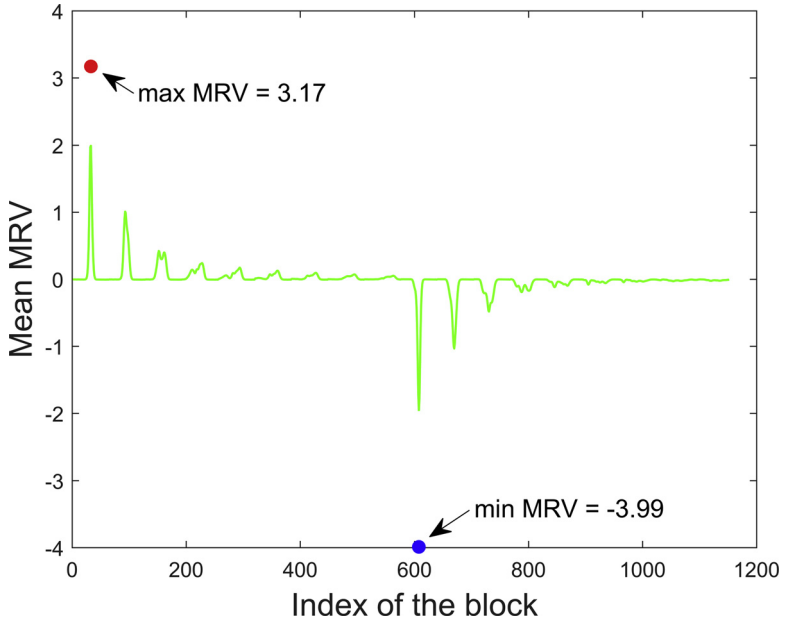

Fig. 2. Distribution of the MRV over blocks.

Table 3
Algorithm flow to determine the soft Eps of DBSCAN based on Grubbs's test.

Input:
A sequence S_N of the MRV feature to be clustered, where the N is the length of S.
Output:
The soft Eps.
Procedure:
1) Sort the sequence S in ascending order.
2) Compute the intervals (denoted by ITV) between every two adjacent points by $ITV(i) = S(i+1) - S(i),\ i = 1,\ 2,\ ...,\ N-1$.
3) Iteratively employ the Grubbs's test to delete the outliers of the ITV until no more outliers are found.
4) Compute the mean value of the remaining ITV, which is denoted by eps.
5) The soft Eps is determined by $Eps = 3 \times eps$.

*As the Grubbs's test is based on the hypothesis of norm distribution, which is also accepted in our algorithm, so according to the Pauta criterion, the three times of the eps is a proper value which is able to cluster the points of the same class with probability of 0.9973.

Similarity Measurement (IDSM), which consists 6 steps below.

(1) Choose the negative category as the base class (B class) and the other one as the target class (T class).
(2) Compute the Jaccard index I_J of the B class on each block.

For two interval E and F, the Jaccard index is defined as

$$J(E, F) = \frac{|E \cap F|}{|E \cup F|}. \tag{2}$$

Based on Eq. (2), we defined a Jaccard similarity function $I_J(\bullet)$ to measure the similarity of the B class on a block. Assume there are N patients of the B class, and patient i has $C_{i,b}$ valid clusters on block b where each cluster interval is denoted as $D_j^{(i,b)}$, $j = 1, 2, \ldots, C_{i,b}$. The $I_J(b)$ is defined as

$$I_J(b) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{C_{i,b}}\sum_{j=1}^{C_{i,b}} J\left(D_j^{(i,b)}, \bigcup_{k=1,k\neq i}^{N}\bigcup_{j=1}^{C_{k,b}} D_j^{(k,b)}\right). \tag{3}$$

Fig. 4 gives us an intuitive understanding of the function, where one row represents a set of the MRV feature of the B-class patient. The I_J is zero on block A, where all the valid clusters of the patients are disjoint

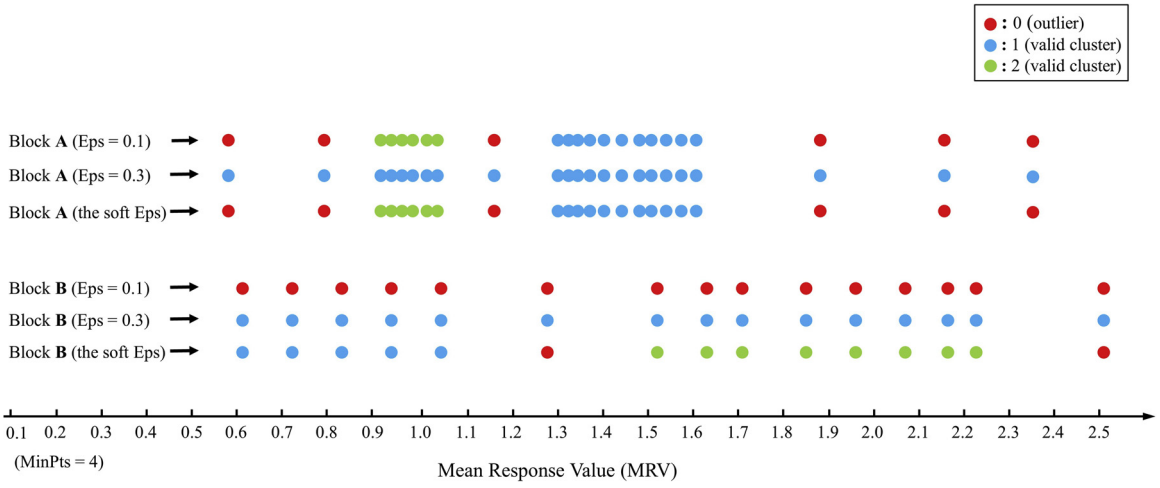

Fig. 3. Comparison of the soft Eps and the hard Eps.

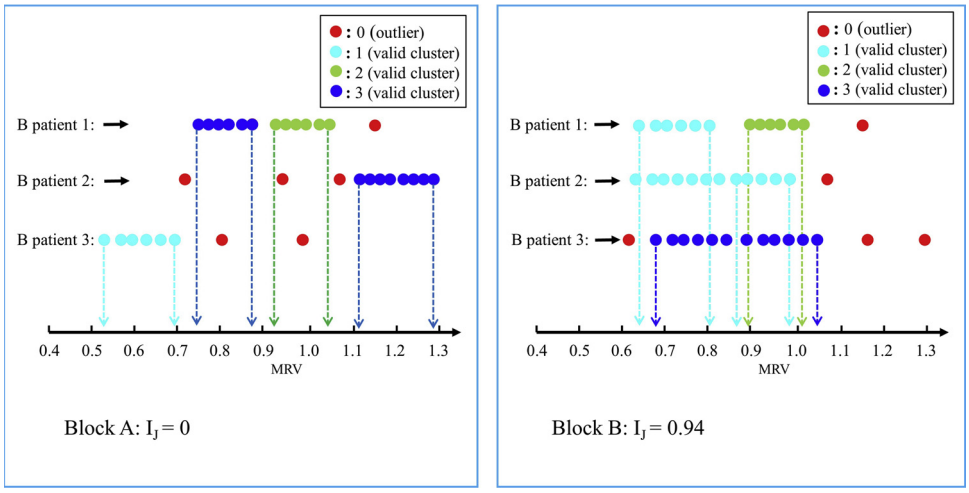

Fig. 4. Schematic diagram of the similarity measurement.

meaning they are totally dissimilar. While on block B, the I_J is very high which conforms to the great similarity of the patients.

(3) Set a threshold value TH_J, such as 0.75, and exclude the blocks with I_J lower than the value.
(4) For each T class patient, find the warning indexes.

As shown in Fig. 5, region B is the union of the valid cluster interval of all the B class patients. The region T1 and T2 is the valid cluster interval of one T class patient, respectively, and the distance 1 and 2 are the corresponding boundary distance to region B. We convert the boundary distance into relative distance according to

$$Rdist = Bdist/(\sigma(T_{cluster}) + \sigma(B_{region})), \quad (4)$$

where *Rdist* is the relative distance of one cluster of the T class patient, *Bdist* the boundary distance, $T_{cluster}$ and B_{region} denotes the feature points in the interval of the T cluster and the region B, respectively, and $\sigma(\bullet)$ computes the standard deviation of the feature points. Then sort all the T clusters in descending order according to corresponding *Rdist*, and exclude those clusters with *Rdist* lower than threshold TH_{Rdist} (such as 3). The larger *Rdist* means the higher probability that the new observations of the two class could not invade each other (a new observation locates nearer to the opposite class). Select the first few (such as the first 3) clusters as the warning indexes for each T patient, which consist a table T_P for the positive category.

(5) Exchange the B class and T class, then repeat step (1)-(4), so obtain a table T_N of the negative category.
(6) Train the local warming model M_L.

Based on T_N and T_P, we computed the number of warnings of each train patient over the tables, where a Gaussian Classifier (GC, detail mathematical derivation was given in the appendix) was presented to identify the warning. Take Fig. 5 for example, as cluster T 1 was a warning cluster of T_P, use the GC to determine a decision point (DP) between the region T 1 and region B. When a new valid cluster of the test patient fall in the left side of the DP_1, it will be identified as a warning of T_P. Finally, the T_P warnings minus the T_N warnings getting one local parametric feature F_L called cumulative warnings (F_{L_C}), and the other local feature F_{L_L} is the latent variable yielded by LS-SVM with input of the F_{L_C}. It should be noted that we only take four types of the category distribution into account in the GC, which are shown in Fig. 1.

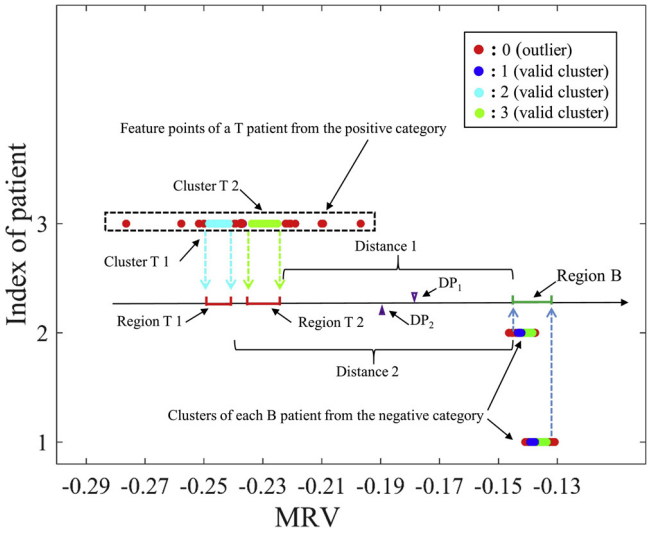

Fig. 5. Schematic diagram of the IDSM method.

2.5.3. Integration

Concatenating the global feature F_G with the local feature F_{L_C} or F_{L_L}, we obtained two kinds of fused feature to train the final integrated model MLG, respectively.

3. Results and discussion

3.1. Formulation of the data analysis

Data analysis is formulated in three aspects. 1) Based on global feature F_I, we investigated the performance difference between the employment of the dynamic spectra (dynamic way) and the stable spectra (stable way); 2) We demonstrated the performance of the proposed LWGF methods based on the dynamic way in ROC, AUC, and accuracy; 3) We validated the model stability of the LWGF with respect to the two free parameters TH_J and TH_{Rdist}.

Three prevalent feature extraction techniques, i.e., gray level co-occurrence matrix (GLCM) [46], local binary pattern (LBP) [47,48], multiscale wavelet energies (MWE) [49]. Here we present a short introduction of the techniques. The GLCM is a texture analysis method, and it calculates the statistics, such as entropy, homogeneity, energy, of a probability of the occurrence of one pixel with a gray level *i* and another pixel with a gray level *j*, where the latter pixel *j* lies on the position of distance *d* and direction *θ* off the pixel *i*. The LBP is also a powerful means of texture analysis, and it has been extended into many versions. The basic version of LBP is to calculate the respective value of the surrounding pixels to the center pixel by thresholding in a 3 × 3-neighborhood. In this way, the 8 points can generate an 8-bit binary

Table 4
Parameters of the algorithm components.

LBP	GLCM	MWE	LS-SVM	RF	KNN	ELM
8 bits; uniform LBP	distance = 1; direction angle: 0°, 45°, 90°, 135°	db4	RBF kernel; grid search; one Vs one	"ntree" = 500; "mtry" = 3	K = 5	Activation Function = tribas; Hidden Neurons = 130; C = 1024;

Table 5
Average recognition rates (10-times) for the test spectra with employment of dynamic or stable spectra.

Classifier	LS-SVM			RF			ELM			KNN		
Feature	LBP	GLCM	MWE	LBP	GLCM	MWE	LBP	GLCM	MWE	LBP	GLCM	MWE
Stable-1	76.15	74.62	70.00	69.23	73.08	65.38	66.54	**79.23**	70.77	73.08	**80.77**	**76.92**
Stable-3	73.33	70.64	71.15	75.64	**85.90**	**74.36**	58.85	65.77	65.26	69.23	75.64	73.08
Dynamic	**84.34**	**77.59**	**74.30**	**78.28**	76.96	73.14	**72.61**	74.11	**73.48**	**78.72**	77.79	73.68

*Stable-1 and Stable-3 denotes the stable method with one stable spectrum or three used for a patient.
**The recognition rate is the number of correct predictions for the spectra of one test patient.

number called LBP code which is used to reflect the texture information for this area. The MWE is to calculate the energies of Wavelet coefficients at different scales, where the Wavelet coefficients are derived from the inner product of signal and Wavelet basis function.

And four popular classifier, i.e., LS-SVM, extreme learning machine (ELM), random forest (RF), K-Nearest Neighbor (KNN), were adopted to be the basic algorithm components. All the programs involved in this paper were written in Matlab (2016b) on Windows, and the parameters of the components were listed in Table 4.

3.2. Stable way or dynamic way

We listed the average recognition rates for the test spectra with employment of dynamic or stable spectra in Table 5. It seems that the dynamic way is better than the stable way as it won seven of the twelve. Because of the different principal of the classifiers, the abilities of which to utilize the information and resist the noise of the spectra are different as well. Nevertheless, as long as we can effectively extract the information hiding in the dynamic spectra and input it to the prediction model, a good results will be expected.

The final aim of the detection is to identify the patient. So, with reusing the LS-SVM, we obtained the results for the patients, where the recognition rates of the global, local, and the integration methods (LWGF based on the global feature F_{G_LBP} in dynamic way) are listed in Table 6. It clearly shows that the dynamic way is much better in accuracy than the corresponding stable way, which probably because the diversity of the feature in dynamic way is richer than that in the stable way due to more spectra were used. If only use one stable spectrum for the test patient, then it is equal to that for the test spectrum. Therefore, we believe that using several stable spectra is detrimental to the identification for patients.

3.3. Performance and model stability of the proposed LWGF

Besides, we draw the ROCs of the stable, dynamic, and the LWGF methods in one prediction in Fig. 6. It is confirmed that both of the two LWGF methods achieve the expected results. Specifically, the second LWGF method ($F_{L\_L} + F_{G\_LBP}$) reaches the average recognition rate of 96.15% and average AUC of 0.98 (10-times, $TH_J = 0.70$, $TH_{Rdist} = 3.0$), which has tremendous advantage than the global methods. So, it indicates that the local features have been effectively integrated with the global feature by the LWGF algorithm. With respect to the two free parameters, the average recognition rates and AUCs are listed in Tables 7 and 8, respectively, which demonstrates the LWGF methods are stable and able to sustain the advantage to the global methods in the range.

3.4. Degradation of the clinical wound sample

The degradation is a notable issue which indicates the cancellation process of the clinical wound sample. We discover that the odor of the fresh sample usually disappeared after collecting 30 spectra. However it is not a continuously weakening process, but an overall declining with slight fluctuations. In our work, although the ATLAS^{TM} sample kit is able to heat the sample to a fixed temperature with rapid and full stir, which cannot make the sample odor release continuously and uniformly as expected. Obviously, it is a tricky problem because the dynamic change would increase the difficulty of detection. However, from another angle, it is also the demonstration for the reasonableness of the technical route of clustering adopted in the LWGF algorithm.

4. Conclusions

In this paper, we proposed the integration algorithm LWGF which was validated in the detection of clinical infection via wound sample odor. As a preliminary study, this work showed that FAIMS technology has the potential for rapid screening of wound infection. We hold the

Table 6
Average recognition rates (10-times) for the test patient of the global feature, local warning feature, and the proposed LWGF.

Global feature F_{G_MWE} (stable-3)	Global feature F_{G_MWE} (dynamic)	Global feature F_{G_GLCM} (stable-3)	Global feature F_{G_GLCM} (dynamic)	Global feature F_{G_LBP} (stable-3)	Global feature F_{G_LBP} (dynamic)	Local feature F_{L_C}	Local feature F_{L_L}	LWGF F_{L_C} + F_{G_LBP}	LWGF F_{L_L} + F_{G_LBP}
73.85/4.73	**76.92/0**	68.46/5.38	**80.38/3.82**	73.85/3.53	**90.77/$_{5.50}$**	84.62/$_0$ ($TH_J = 0.7$, $TH_{Rdist} = 3.2$)	84.62/$_0$ ($TH_J = 0.7$, $TH_{Rdist} = 3.2$)	**95**/$_{1.86}$ ($TH_J = 0.7$, $TH_{Rdist} = 3.2$)	**96.15**/$_0$ ($TH_J = 0.7$, $TH_{Rdist} = 3.2$)

*The subscript after the slash indicates the standard deviation.

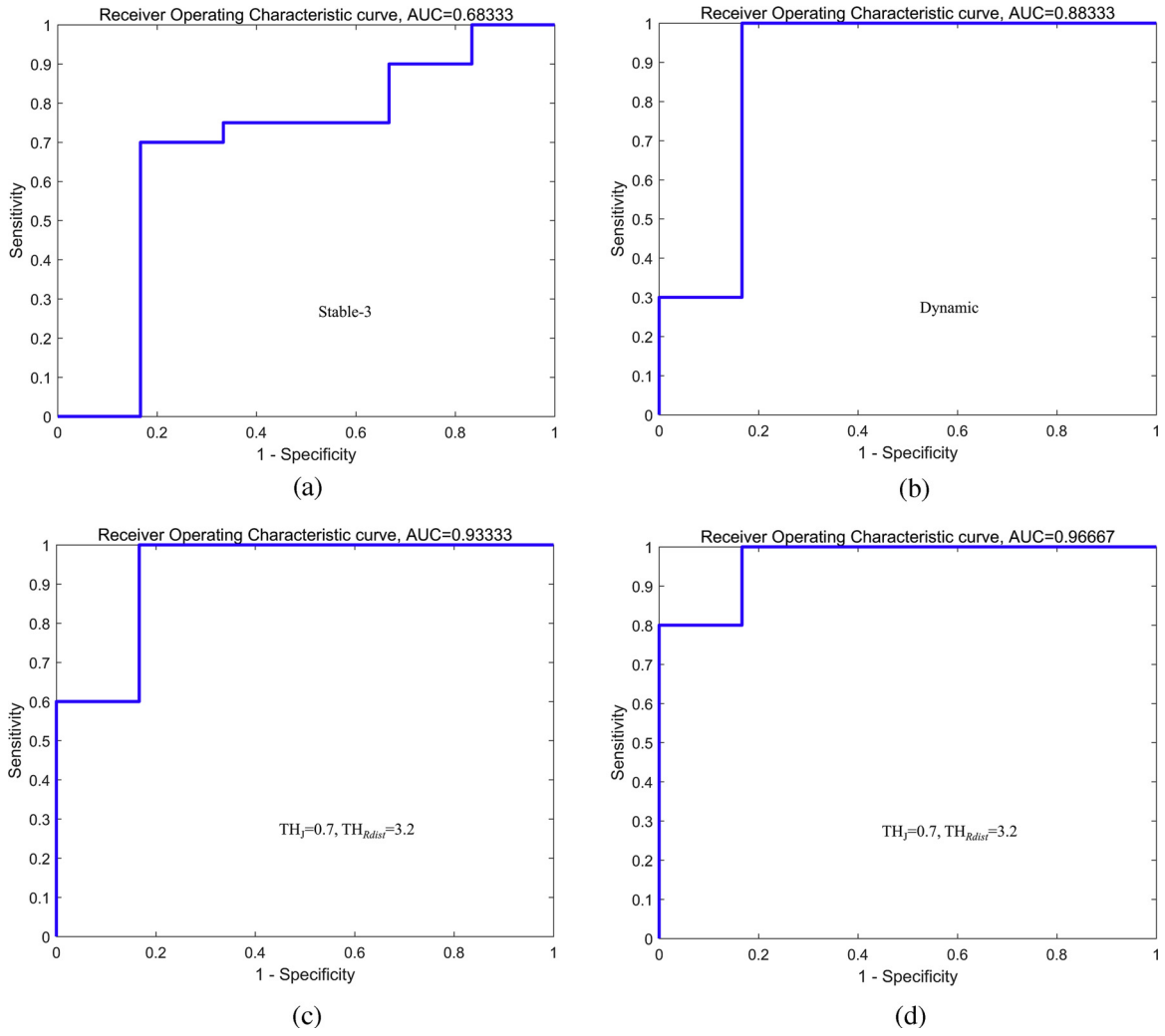

Fig. 6. (a) ROC of the F_G_LBP with three stable spectra. (b) ROC of the F_G_LBP with dynamic spectra. (c) ROC of the LWGF with F_L_C. (d) ROC of the LWGF with F_L_D.

opinion that effective use of the available information is the foundation to accomplish the detection task. So we choose to use the dynamic spectra rather than the stable spectra to construct the recognition model. However, it also brings the difficulty that how to extract and exploit the effective ones from the enormous stockpile of information, where the interference can coexist. Therefore, we proposed the LWGF algorithm to screen each block of the spectra to select the valid ones, and integrate the effective information of global and local space together to recognize the odor.

Table 7
Average recognition rates (10-times) for the test patient of the proposed LWGF with respect to the free parameters TH_J and TH_{Rdist}.

(a). Average recognition rates of the F_{L_C} fused with the global feature F_{G_LBP}

TH_J	TH_{Rdist}					
	3.0	3.2	3.4	3.6	3.8	4.0
0.70	$94.62/_{2.69}$	$95.00/_{1.86}$	$94.23/_{2.72}$	$95.00/_{1.86}$	$93.46/_{2.60}$	$93.46/_{2.60}$
0.75	$93.85/_{3.72}$	$95.00/_{1.86}$	$95.00/_{1.86}$	$95.00/_{1.86}$	$94.62/_{2.69}$	$93.85/_{3.24}$
0.80	$94.23/_{2.03}$	$93.85/_{2.69}$	$93.85/_{2.69}$	$93.46/_{3.65}$	$94.62/_{1.99}$	$93.46/_{3.65}$

(b). Average recognition rates of the F_{L_L} fused with the global feature F_{G_LBP}

TH_J	TH_{Rdist}					
	3.0	3.2	3.4	3.6	3.8	4.0
0.70	$91.15/_{1.86}$	$96.15/_{0}$	$88.47/_{0}$	$92.31/_{1.81}$	$88.46/_{0}$	$91.92/_{1.22}$
0.75	$91.54/_{1.62}$	$94.62/_{2.69}$	$93.08/_{2.43}$	$89.23/_{3.97}$	$88.85/_{2.84}$	$91.15/_{2.60}$
0.80	$95.00/_{1.86}$	$94.62/_{1.99}$	$95.38/_{1.62}$	$94.62/_{2.69}$	$95.38/_{1.62}$	$95.00/_{1.86}$

*The average recognition rates of the global method (LBP + LS-SVM) is $90.77/_{5.50}$.

Table 8
Average AUCs (10-times) of the proposed LWGF with respect to the free parameters TH_J and TH_{Rdist}.

(a). Average AUCs of the F_{L_C} fused with the global feature F_{G_LBP}

TH_J	TH_{Rdist}					
	3.0	3.2	3.4	3.6	3.8	4.0
0.70	$0.93/_{0.0147}$	$0.93/_{0.0088}$	$0.93/_{0.0136}$	$0.91/_{0.0197}$	$0.92/_{0.0137}$	$0.91/_{0.0163}$
0.75	$0.91/_{0.0152}$	$0.92/_{0.0142}$	$0.92/_{0.0161}$	$0.91/_{0.0136}$	$0.91/_{0.0177}$	$0.91/_{0.0179}$
0.80	$0.91/_{0.020}$	$0.92/_{0.0144}$	$0.91/_{0.0201}$	$0.91/_{0.0227}$	$0.91/_{0.0141}$	$0.91/_{0.0189}$

(b). Average AUCs of the F_{L_L} fused with the global feature F_{G_LBP}

TH_J	TH_{Rdist}					
	3.0	3.2	3.4	3.6	3.8	4.0
0.70	$0.98/_{0.0035}$	$0.97/_{0.0111}$	$0.95/_{0.0090}$	$0.95/_{0.0026}$	$0.96/_{0.0070}$	$0.96/_{0.0131}$
0.75	$0.92/_{0.0158}$	$0.92/_{0.0168}$	$0.91/_{0.0157}$	$0.91/_{0.0196}$	$0.90/_{0.\ 0137}$	$0.90/_{0.0151}$
0.80	$0.93/_{0.0153}$	$0.92/_{0.0157}$	$0.93/_{0.0149}$	$0.94/_{0.0142}$	$0.92/_{0.0126}$	$0.90/_{0.0194}$

*The average AUC of the global method (LBP + LS-SVM) is $0.88/_{0.0181}$.

Authors' contribution

Tong Sun: designed and implemented the proposed machine learning algorithm, analyzed the clinical FAIMS data, wrote the manuscript. QingHua He: Leader of the project [grant number: 2014DFA31560], acquisition of the clinical samples and the odor spectra, review of the manuscript. FengChun Tian: review of the manuscript. YanYi Lu, acquisition of the clinical samples and the odor spectra. YuTian Bi, Xiaozheng Zhong, Qingshan Guo, Ying Lei, Lin Zeng: took part in the acquisition of the clinical samples. Jiao He: took part in the algorithm design and data analysis, review of the data, code, and algorithm results. TaiCong Yang: took part in the data analysis.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by National International Science and Technology Cooperation Project of China [grant number: 2014DFA31560] and China-Canada Medical Intelligent E-nose Center of Chongqing [grant number: cstc2013gjhz10003].

Appendices

Let random variable V_i denotes the response intensity value of the i-th pixel on a spectrum, and $V_i = \sum_{r=1}^{N_i} C_{i,r}$, where N_i is the number of chemical species appearing at the i-th pixel position, and $C_{i,r}$ is a random variable representing the current intensity of the r-th chemical species at the position. For an area S on the spectrum, we define the average current intensity $\overline{V_S} = (\sum_{i\in S} V_i)/M_S$, where M_S is the pixel number of S. It is reasonable to assume that $\overline{V_S}$ obey the *Gaussian* distribution, according to the central limit theorem.

For two categories of the valid cluster points X and Y, calculate the expectation μ_X, μ_Y, and standard deviation σ_X,σ_Y of the groups, respectively. The probability density function curve of the two categories and the decision point (DP) p are shown in Fig. A1. For a test sample x, the decision rule is:

$$\begin{cases} x \in \text{category} X \text{if x} \le p \\ x \in \text{category} Y \text{if x} > p \end{cases}, \text{when} \mu_X \le \mu_Y,$$

$$\begin{cases} x \in \text{category} X \text{if x} \ge p \\ x \in \text{category} Y \text{if x} < p \end{cases}, \text{when} \mu_X > \mu_Y.$$

Therefore, the total misclassification probability ER of the x is formulated as

$$ER = ER_X + ER_Y = \int_{-\infty}^{p} (\sqrt{2\pi}\sigma_X)^{-1} \exp(-(x-\mu_X)^2/(2\sigma_X^2))dx + \int_{p}^{+\infty} (\sqrt{2\pi}\sigma_Y)^{-1} \exp(-(x-\mu_Y)^2/(2\sigma_Y^2))dx, \quad \text{(A1)}$$

where ER_X and ER_Y is the misclassification probability of the negative and positive category, respectively. Then, we aim to determine an optimal decision point p to minimize the misclassification probabilities. Let $\partial ER/\partial p = 0$, we obtain

$$(\sqrt{2\pi}\sigma_X)^{-1} \exp(-(x-\mu_X)^2/(2\sigma_X^2)) - (\sqrt{2\pi}\sigma_Y)^{-1} \exp(-(x-\mu_Y)^2/(2\sigma_Y^2)) = 0 \quad \text{(A2)}$$

$$\Leftrightarrow (\sigma_X^2 - \sigma_Y^2)p^2 + 2(\mu_X\sigma_Y^2 - \mu_Y\sigma_X^2)p - \mu_X^2\sigma_Y^2 + \mu_Y^2\sigma_X^2 + 2\sigma_X^2\sigma_Y^2 \ln(\sigma_Y/\sigma_X) = 0. \quad \text{(A3)}$$

It means that p meets the quadratic equation Eq. (A-3) which yields two solutions

$$p_1 = \{-2(\mu_X\sigma_Y^2 - \mu_Y\sigma_X^2) + [(\mu_X\sigma_Y^2 - \mu_Y\sigma_X^2)^2 - 4(\sigma_X^2 - \sigma_Y^2)(-\mu_X^2\sigma_Y^2 + \mu_Y^2\sigma_X^2 + 2\sigma_X^2\sigma_Y^2 \ln(\sigma_Y/\sigma_X)]^{1/2}\}/[2(\sigma_X^2 - \sigma_Y^2)], \quad \text{(A4)}$$

$$p_2 = \{-2(\mu_X\sigma_Y^2 - \mu_Y\sigma_X^2) - [(\mu_X\sigma_Y^2 - \mu_Y\sigma_X^2)^2 - 4(\sigma_X^2 - \sigma_Y^2)(-\mu_X^2\sigma_Y^2 + \mu_Y^2\sigma_X^2 + 2\sigma_X^2\sigma_Y^2 \ln(\sigma_Y/\sigma_X)]^{1/2}\}/[2(\sigma_X^2 - \sigma_Y^2)]. \quad \text{(A5)}$$

Then one of the solutions between μ_X and μ_Y is accepted, and the other one should be rejected.

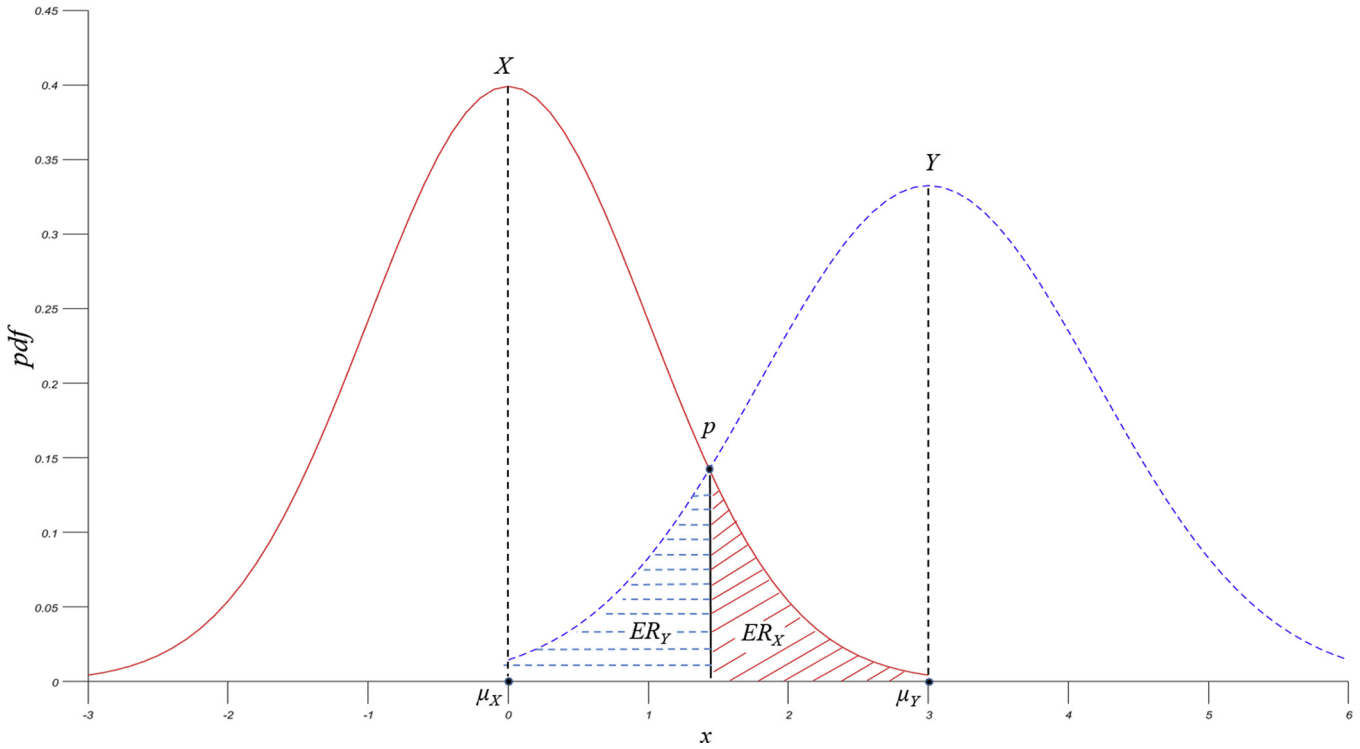

Fig. A1. Schematic diagram of the minimus error probility.

References

[1] C.K. Sen, G.M. Gordillo, S. Roy, R. Kirsner, L. Lambert, T.K. Hunt, et al., Human skin wounds: a major and snowballing threat to public health and the economy, Wound Repair Regen. 17 (2010) 763–771.

[2] W.H. Organization, Injuries and Violence: the Facts 16 Geneva World Health Organization, 2010.

[3] M. Plummer, C.D. Martel, J. Vignat, J. Ferlay, F. Bray, S. Franceschi, Global burden of cancers attributable to infections in 2012: a synthetic analysis, Lancet Glob. Health 4 (2016) e609–e616.

[4] N. Yusuf, A. Zakaria, M.I. Omar, A.Y.M. Shakaff, M.J. Masnan, L.M. Kamarudin, et al., In-vitro diagnosis of single and poly microbial species targeted for diabetic foot infection using e-nose technology, BMC Bioinformatics 16 (2015) 158.

[5] J. Rose, T.G. Weiser, P. Hider, L. Wilson, R.L. Gruen, S.W. Bickler, Estimated need for surgery worldwide based on prevalence of diseases: a modelling strategy for the WHO Global Health Estimate, Lancet Glob. Health 3 (2015) S13–S20.

[6] S.I. Berríostorres, C.A. Umscheid, D.W. Bratzler, B. Leas, E.C. Stone, R.R. Kelz, et al., Centers for disease control and prevention guideline for the prevention of surgical site infection, JAMA Surg. 152 (2017) (2017) 784.

[7] R.H. Deurenberg, E. Bathoorn, M.A. Chlebowicz, N. Couto, M. Ferdous, S. Garcíacobos, et al., Application of next generation sequencing in clinical microbiology and infection prevention, J. Biotechnol. 250 (2017) 16–24.

[8] D.J. Leaper, C.E. Edmiston, World Health Organization: global guidelines for the prevention of surgical site infection, J. Hosp. Infect. 95 (2017) 135–136.

[9] T. Saviauk, J.P. Kiiski, M.K. Nieminen, N.N. Tamminen, A.N. Roine, P.S. Kumpulainen, et al., Electronic nose in the detection of wound infection bacteria from bacterial cultures: a proof-of-principle study, Eur. Surg. Res. 59 (2018) 1–11.

[10] Lieuwe D.J. Bos, Peter J. Sterk, Marcus J. Schultz, Volatile metabolites of pathogens: a systematic review, PLoS Pathog. 9 (5) (2013) e1003311(2013-5-9), 9.

[11] J. Chen, H. Shi, Q. Wang, J.N. Tang, Categories, detection techniques and applications of bacterial volatile metabolites, J. Microbiol. (2015).

[12] A. Giacometti, O. Cirioni, A.M. Schimizzi, M.S.D. Prete, F. Barchiesi, M.M. D'Errico, et al., Epidemiology and microbiology of surgical wound infections, J. Clin. Microbiol. 38 (2000) 918–922.

[13] S. Alugupalli, B. Olsson, L. Larsson, Detection of 2-eicosanol by gas chromatography-mass spectrometry in sputa from patients with pulmonary mycobacterial infections, J. Clin. Microbiol. 31 (1993) 1575.

[14] A.N. Thomas, Novel noninvasive identification of biomarkers by analytical profiling of chronic wounds using volatile organic compounds, Wound Repair Regen. 18 (2010) 391–400.

[15] C.M. Timm, E.P. Lloyd, A. Egan, R. Mariner, D. Karig, Direct growth of bacteria in headspace vials allows for screening of volatiles by gas chromatography mass spectrometry, Front. Microbiol. 9 (2018) 491.

[16] F. David, B. Tienpont, P. Sandra, Chemotaxonomy of bacteria by comprehensive GC and GC-MS in electron impact and chemical ionisation mode, J. Sep. Sci. 31 (2015) 3395–3403.

[17] E. Tait, J.D. Perry, S.P. Stanforth, J.R. Dean, Identification of volatile organic compounds produced by bacteria using HS-SPME-GC-MS, J. Chromatogr. Sci. 52 (2014) 363–373.

[18] W. Filipiak, Molecular analysis of volatile metabolites released specifically by staphylococcus aureus and pseudomonas aeruginosa, BMC Microbiol. 12 (2012) 113.

[19] P. George, T. Erica, H.C. William, T. Michelle, E. Jason, G. Alan, Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: analysis by solid-phase microextraction and gas chromatography-mass spectrometry, J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 877 (2009) 2011–2018.

[20] A. Scottthomas, Novel Diagnostics: Progress Toward a Breath Test for Invasive, (2011).

[21] A.J. Scott-Thomas, M. Syhre, P.K. Pattemore, M. Epton, R. Laing, J. Pearson, et al., 2-Aminoacetophenone as a potential breath biomarker for Pseudomonas aeruginosa in the cystic fibrosis lung, BMC Pulm. Med. 10 (2010) 56-.

[22] H.G. Byun, K.C. Persaud, A.M. Pisanelli, Wound-state monitoring for burn patients using E-Nose/SPME system, ETRI J. 32 (3) (2010) 440–446 (2010-06-08), 32.

[23] K.C. Persaud, A.M. Pisanelli, A. Bailey, Woundmonitor: Monitoring Volatiles to Detect Infection, (2008).

[24] X.U. Xun-Tao, F.C. Tian, Y. Jia, M.A. Jian-Wei, Rapid detection of wound pathogen by Enose with a gas condensation unit, Chin. J. Sens. Actuators 22 (2009) 303–306.

[25] X. Guo, C. Peng, S. Zhang, J. Yan, S. Duan, L. Wang, et al., A novel feature extraction approach using window function capturing and QPSO-SVM for enhancing electronic nose performance, Sens. Basel 15 (2015) 15198–15217.

[26] R. Dutta, E.L. Hines, J.W. Gardner, P. Boilot, Bacteria classification using Cyranose 320 electronic nose, Biomed. Eng. Online 1 (2002) 4.

[27] N. Yusuf, M.I. Omar, A. Zakaria, A.I. Jeffree, R. Thriumani, A.A. Abdullah, et al., Evaluation of E-nose technology for detection of the causative bacteria in different culture media on diabetic foot infection, Biomed. Eng. Sci. (2014).

[28] A.D. Wilson, B. Manuela, Advances in electronic-nose technologies developed for biomedical applications, Sens. Basel 11 (2011) 1105.

[29] F. Reinhard, A.H.J. Kolk, B. Conrad, B. Patricia, P.R. Klatser, A.C. Woodman, Prospects for clinical application of electronic-nose technology to early detection of Mycobacterium tuberculosis in culture and sputum, J. Clin. Microbiol. 44 (2006) 2039–2045.

[30] I.A. Ratiu, V. Bocos-Bintintan, A. Patrut, V.H. Moll, M. Turner, C.L.P. Thomas, Discrimination of bacteria by rapid sensing their metabolic volatiles using an aspiration-type ion mobility spectrometer (a-IMS) and gas chromatography-mass spectrometry GC-MS, Anal. Chim. Acta 982 (2017) 209–217.

[31] M. Sasidhar, B. Lars Mathias, S. Andreas, B. J?Rg Ingo, Detection of volatile metabolites of Escherichia coli by multi capillary column coupled ion mobility spectrometry, Anal. Bioanal. Chem. 394 (2009) 791.

[32] M. Schrader, M. Jürgens, R. Hess, P. Schulz-Knappe, M. Raida, W.G. Forssmann, Ion mobility spectrometry for microbial volatile organic compounds: a new identification tool for human pathogenic bacteria, Appl. Microbiol. Biotechnol. 93 (2012)

2603–2614.
[33] J.A. Schultz, E.K. Lewis, A.S. Woods, S.N. Jackson, K.K. Murray, J.M. Hayes, et al., Matrix assisted laser desorption ionization ion mobility time-of-flight mass spectrometry of bacteria, ACS Symp. 1065 (2011) 143–160.
[34] A.P. Snyder, W.M. Maswadeh, A. Tripathi, J.P. Dworzanski, Detection of gram-negative Erwinia herbicola outdoor aerosols with pyrolysis - gas chromatography/ion-mobility spectrometry, Field Anal. Chem. Technol. 4 (2015) 111–126.
[35] J.D. Sobela, A. Lorber, Diagnosing vaginal infections through measurement of biogenic amines by ion mobility spectrometry, Eur. J. Obstet. Gynecol. Reprod. Biol. 163 (2012) 81–84.
[36] N.J.C. Strachan, F.J. Nicholson, I.D. Ogden, An automated sampling system using ion mobility spectrometry for the rapid detection of bacteria, Anal. Chim. Acta (1995).
[37] J.A. Covington, M.P. van der Schee, A.S.L. Edge, B. Boyle, R.S. Savage, R.P. Arasaradnam, The application of FAIMS gas analysis in medical diagnostics, Analyst 140 (2015) 6775–6781.
[38] R. Sinha, L.R. Khot, B.K. Schroeder, FAIMS based sensing of Burkholderia cepacia caused sour skin in onions under bulk storage condition, J. Food Meas. Charact. 11 (2017) 1578–1585.
[39] R. Sinha, L.R. Khot, B.K. Schroeder, Y. Si, Rapid and non–destructive detection of Pectobacterium carotovorum causing soft rot in stored potatoes through volatile biomarkers sensing, Crop. Prot. 93 (2017) 122–131.
[40] R. Sinha, L.R. Khot, B.K. Schroeder, S. Sankaran, FAIMS based volatile fingerprinting for real-time postharvest storage infections detection in stored potatoes and onions, Postharvest Biol. Technol. 135 (2018) 83–92.
[41] A. Kontunen, M. Karjalainen, J. Lekkala, A. Roine, N. Oksala, Tissue identification in a porcine model by differential ion mobility spectrometry analysis of surgical smoke, Ann. Biomed. Eng. 46 (2018) 1091–1100.
[42] R.P. Arasaradnam, M. McFarlane, E. Daulton, E. Westenbrink, N. O'Connell, S. Wurie, et al., Non-invasive distinction of non-alcoholic fatty liver disease using urinary volatile organic compound analysis: early results, J. Gastrointestin. Liver Dis. 24 (2015) 197–201.
[43] R.P. Arasaradnam, E. Westenbrink, M.J. McFarlane, R. Harbord, S. Chambers, N. O'Connell, et al., Differentiating coeliac disease from irritable bowel syndrome by urinary volatile organic compound analysis – a pilot study, PLoS One 9 (2014) e107312.
[44] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases With Noise, (1996).
[45] F.E. Grubbs, Procedures for detecting outlying observationsin samples, Technometrics 11 (1969) 21.
[46] C.C. Gao, X.W. Hui, GLCM-based texture feature extraction, Comput. Syst. Appl. (2010).
[47] T. Ahonen, J. Matas, C. He, M. Pietikäinen, Scia, Oslo, Norway, JuneRotation Invariant Image Description With Local Binary Pattern Histogram Fourier Features, Image Analysis, Scandinavian Conference2009, Rotation Invariant Image Description With Local Binary Pattern Histogram Fourier Features, Image Analysis, Scandinavian Conference (2009).
[48] S. Wan, H.C. Lee, X. Huang, T. Xu, T. Xu, X. Zeng, et al., Integrated local binary pattern texture features for classification of breast tissue imaged by optical coherence microscopy, Med. Image Anal. 38 (2017) 104–116.
[49] L.N. Sharma, S. Dandapat, A. Mahanta, Multiscale wavelet energies and relative energy based denoising of ECG signal, IEEE International Conference on Communication Control & Computing Technologies, (2010).

Tong Sun received his M.S. degree in computer technology from HUST (Huazhong University of Science and Technology) of China in 2008. He is currently a Ph.D. student in electronic circuit and system from School of Microelectronics and Communication Engineering, Chongqing University, with main research interests in machine olfactory, ion mobility spectrometry, computer vision, statistical learning, machine learning and deep learning.

FengChun Tian received his Ph.D. degree in 1997 in electrical engineering from Chongqing University. He is currently a professor with the College of Communication Engineering of Chongqing University. His research interests include Electronic nose technology, artificial olfactory systems, pattern recognition, chemical sensors, signal/image processing, wavelet, and computational intelligence. In 2006 and 2007, he was recognized as a part-time Professor of GUELPH University, Canada.

YuTian Bi an associate Professor in the Respiratory Department at Daping hospital of Third Military Medical University of China. He received B.S. and M.S. degrees from the Third Military Medical University of China in 1988 and 1997, respectively. His major research interests lie in the pathogenesis of infection of respiratory system.

Xiaozheng Zhong received his Bachelor degree of medicine in 1994 in clinical medicine from North Sichuan Medical College. He is currently an attending physician of traumatology department of Daping Hospital.

He Jiao received her Bachelor degree in communication engineering in 2017 from Southwest University, China. She is currently pursuing a master's degree at the School of Microelectronics and Communication Engineering, Chongqing University. Her research interests are electronic nose and machine learning.

TaiCong Yang received his Bachelor degree in communication engineering in 2016 from Chongqing University, China. He is currently pursuing a doctor's degree at the School of Microelectronics and Communication Engineering, Chongqing University. His research interests are electronic nose and machine learning.

Qingshan Guo received his Doctor's degree from the Third Military Medical University in 2006 in Osteopathic Medicine. He is an associate professor and associate chief physician of traumatology department of Daping Hospital.

Ying Lei is a registered nurse and enterostomal therapist. She is the head nurse of trauma center in Daping hospital. She is a clinical assistant director nurse and specialized in wound treatment.

YanYi Lu received the M.S. degree in biomedical engineering from Chongqing Medical University, Chongqing, China, in 2015. She is currently a research associate in Daping Hospital, Surgery Institute of the Third Military Medical University, with main research interests in biomedical signal processing, electronic nose technology and artificial intelligence.

Lin Zeng graduated from Yuzhou University in 1989 with a bachelor's degree in chemical technology. She has been engaged in biomedical experiments. She is currently obtained the title of senior laboratory technician in Daping Hospital, Surgery Institute of the Third Military Medical University.

QingHua He received her Ph. D. degree in 2003 in biomedical engineering from Chongqing University. She is currently an associate research fellow and associate professor in Daping Hospital, Surgery Institute of the Third Military Medical University. Her research interests is biomedical signal processing, include medical electronic nose and brain-computer interface.