Sensors and Materials, Vol. 35, No. 1 (2023) 1–14
MYU Tokyo

S & M 3151

Smooth Discriminant Analysis Combined with an Electronic Nose System to Classify the Gas Information of Beer

Junjing Wang,[1*] Qingquan Bian,[1] and Manhua Wan[2]

[1]Henan Polytechnic Institute, Faculty of Electronic Information Engineering, Nanyang, Henan 473009, China
[2]Jiujiang Vocational and Technical College, School of Mechanical Engineering, Jiujiang, Jiangxi 332000, China

(Received October 18, 2022; accepted November 28, 2022)

Keywords: electronic nose, smooth discriminant analysis, pattern recognition, gas classification, beer

Even for the same brand of beer, beer quality may differ in different production batches. It is important to propose a fast and effective beer quality inspection technology to control the production quality of beer. In this paper, a smooth discriminant analysis (SDA) method combined with an electronic nose (e-nose) system is proposed to identify beer gas information in different production batches. A multi-pattern recognition method is combined to classify the gas information. Firstly, using the PEN3 e-nose system, different batches of beer gas information are obtained. Secondly, an SDA method is proposed, which strengthens the linearization between gas features and improves the processing effect of gas features. Thirdly, multi-pattern recognition methods are applied and combined with multiple feature dimensionality reduction methods to demonstrate the effectiveness of SDA. The results show that SDA achieves effective dimensionality reduction of different batches of beer gas features and obtains the best classification performance with the random forest (RF), with an accuracy of 97.70%, precision of 98.47%, and recall of 98.23%, thus achieving beer gas identification from different production batches.

1. Introduction

Beer is rich in various nutrients and water and contains essential amino acids for the human body. Beer also contains ten different vitamins and nutrients.[1,2] Beer has a low alcohol concentration and contains carbon dioxide gas. In the process of beer production, manufacturers strictly control the brewing process and various parameters to ensure the production of high-quality beer. However, consumers can sometimes detect differences in beer quality in different production batches, even for the same brand of beer. Therefore, a fast and effective detection method is required to effectively detect the production quality of beer, strictly control the beer production process, and provide an effective supervision method for producing high-quality beer.

Traditional beer quality testing methods are mainly based on physical and chemical index analysis and artificial sensory evaluation.[3,4] Physical and chemical indicators mainly include

*Corresponding author: e-mail: wjjing_0282@163.com
https://doi.org/10.18494/SAM4176

ISSN 0914-4935
https://myukk.org/

dissolved oxygen, turbidity, diethyl phthalide, and carbon dioxide content, which require precise detection conditions, are time-consuming, and cannot effectively represent the overall quality of beer.[5] Artificial sensory evaluation has the disadvantages of a long evaluation period, strong subjectivity, and poor repeatability.[5] Because the gas in beer directly affects people's sensory experience when drinking beer, it can reflect the overall quality of beer. The gas information of different beer brands is clearly different, and the gas information of different production batches is also different. In this paper, we apply this principle to realize the intelligent detection of the overall quality of beer based on the advanced sensing detection technology of the electronic nose (e-nose).

In the detection experiment of the e-nose system, the detection method of the dynamic headspace is the most common method, which has the advantages of high stability and high detection signal validity.[6,7] The data processing process consists of feature extraction, feature processing, and pattern recognition.[8] Feature extraction mainly includes time domain features, frequency domain features, and spatial domain features.[9] Feature processing mainly includes dimensionality reduction and selection.[10] Pattern recognition includes supervised, unsupervised, and semi-supervised learning methods.[11] The e-nose system sensor array has cross-sensitive response characteristics; thus, the extracted gas features are clearly correlated. If the original gas features are directly used for pattern recognition, it will directly affect the gas decision-making accuracy.

Feature processing methods are crucial to remove redundant gas features. Barriault *et al.* used principal component analysis (PCA) and linear discriminant analysis (LDA) to reduce the dimensionality of the leakage gas information.[12] Cho and Kurup proposed the decision tree method for the dimensionality reduction and showed the importance of variables.[13] Nozaki and Nakamoto implemented a computational method to predict the gas composition, and their predictive model utilized nonlinear dimensionality reduction of mass spectrometry data.[14] Wu *et al.* proposed a method for reducing the number of features by separately processing two different types of features, spatial and temporal.[15] Shin *et al.* utilized the relief feature selection technique to select suitable sensors from the sensor array.[16] Wijaya and Afianti proposed a feature selection method to eliminate redundant and irrelevant features, which can reduce the dimensionality as well as the complexity of the original problem.[17] Mishra *et al.* proposed a hybrid adaptive neuro-fuzzy jamming system model to optimize the sensor array.[18] Wijaya *et al.* proposed an integrated feature selection method based on information theory to reduce the dimension and complexity of gas features.[19] These feature processing methods realize the dimensionality reduction of gas information to a certain extent, and it can be seen that there is a clustering trend through visual analysis technology. However, in the case of high-dimensional and complex gas features, the dimensionality reduction effect cannot be optimized owing to the clear nonlinear relationship.

In the process of gas information classification, it is impossible to select effective gas features for gas classification only from the response value, which may easily lead to overfitting of the model. Since a sensor array has clear cross-sensitive response characteristics with different gas features correlated, we propose a feature dimension reduction method to reduce the dimension of the original gas features effectively. The optimal gas feature set can improve the classification

performance of the pattern recognition algorithm. The main contributions of this paper are as follows:

(1) The beer gas information of different production batches is obtained using the e-nose detection system, and a smooth discriminant analysis (SDA) method is proposed to reduce the dimensionality of gas features.

(2) By combining SDA with multi-pattern recognition technology, beer gas information of different production batches can be effectively identified, providing a fast and accurate analytical method for beer quality supervision.

In practical engineering applications, the main purpose of a beer producer is to ensure that the beer quality of each production batch should be the same or there should be no obvious difference. After obtaining different batches of beer gas information based on the e-nose system, SDA realizes the dimension reduction analysis of different batches of beer. By observing the dimension reduction and classification results, we can see the quality difference for different batches of beer, and each parameter in the production process should be strictly adjusted to control beer quality.

2. Materials and Methods

2.1 Beer sample and electronic nose system

The beer samples were taken from a brewery with production dates of January 19, February 20, March 19, April 18, and May 20, 2022. The samples were uniformly stored under an ambient temperature of 22 ± 0.5 °C and an ambient humidity of 40 ± 5 RH, avoiding light.

The commercialized PEN3 e-nose system was used to collect the beer gas information. Its main component comprises ten built-in metal oxide sensors, and the sensor array can detect cross-sensitive information of the gas to be measured. Figure 1 shows the structure of the PEN3

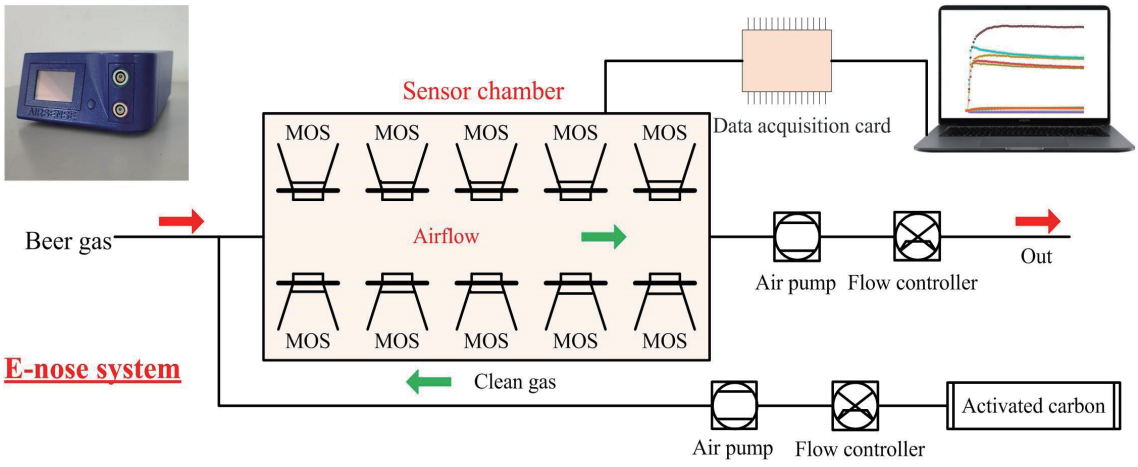

Fig. 1. (Color online) Structure of PEN3 e-nose system.

e-nose system. Table 1 shows the names, detection substances, and detection limits of the ten sensors. The response output value of the sensor is the ratio of the conductance (G) of the gas to be measured by the sensor to the conductance (G_0) of the clean air filtered by activated carbon, i.e., G/G_0. When the beer gas is in contact with the sensor array, the sensor's surface undergoes a redox reaction and an electrical signal is generated. Owing to the cross-sensitivity of the sensor, comprehensive detection information of the sample to be tested is obtained.

2.2 Experiment on gas detection

During the e-nose system detection experiment carried out in this study, the ambient temperature is 22 ± 0.5 ℃ and the ambient humidity is 40 ± 5 RH. The main experimental process is as follows:

(1) To ensure that the sensor array reaches the normal working temperature range (300–500 ℃) and to ensure that its output signal is stable, it is necessary to preheat the e-nose for 1 h.
(2) A sample of 10 ml of beer is placed in a 50 ml sample bottle and allowed to stand for 5 min to ensure sufficient headspace air.
(3) Clean air filtered by activated carbon is used to clear the sensor gas chamber and gas path for 60 s, where the gas flow rate is 300 ml/min.
(4) The dynamic headspace sampling method is used to obtain beer gas information, where the sampling time is 80 s, the sampling frequency is 1 Hz, and the gas flow rate is 300 ml/min. Figure 2 shows the response curve of the sensor array.

To obtain parallel samples, steps (2)–(4) are repeated, the gas information of 30 samples of beer from different production batches is collected, and a total of 150 samples are collected from five different production batches.

2.3 SDA method

LDA is a supervised dimensionality reduction method that gives the dimensionality-reduced data small intra-class differences and large inter-class differences, thereby improving the

Table 1
Main performance of gas sensors.

No.	Sensor	Main performance	Detectability (ppm)
1	W1C	Sensitive to aromatic compounds	10
2	W5S	Highly sensitive to nitrogen oxides, very sensitive with negative signal	1
3	W3C	Sensitive to aromatic compounds and ammonia	10
4	W6S	Sensitive to hydrogen	100
5	W5C	Sensitive to aromatic components and alkanes	1
6	W1S	Sensitive to methane, broad range	100
7	W1W	Sensitive to many terpenes, sulfur compounds, and organic sulfur compounds	1
8	W2S	Sensitive to ethanol, less sensitive to aromatic compounds	100
9	W2W	Sensitive to aromatic compounds, organic sulfur compounds	1
10	W3S	Detection of alkanes in high concentrations	10

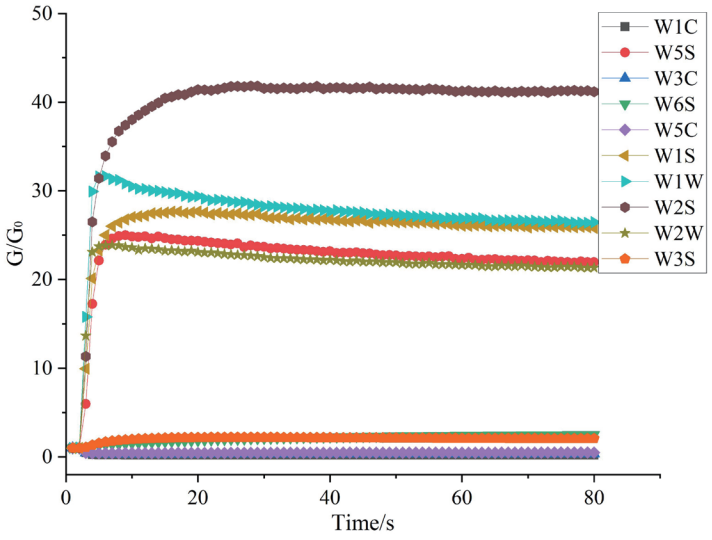

Fig. 2. (Color online) Response curve of the sensor array.

classification performance of pattern recognition algorithms. In the mathematical calculation process of LDA, the inter-class scattering matrix and the intra-class scattering matrix are denoted as A_b and A_w, respectively. First, $A_w^{-1}A_b$ is calculated, then the optimal feature matrix W in the projection direction of dimensionality reduction is obtained. Finally, the original sample is multiplied by the feature vector to obtain the feature matrix after dimensionality reduction.

Denoting the set of samples as $X_c=[x_c^1, x_c^2, \cdots, x_c^{n_c}]$, $c=1,\cdots,k$ represents samples, the cth class can be expressed as n_c, μ_c $\frac{}{n_c}\sum_{i=}^{n_c} x_c^i$ is the sample mean center of class c, and $\bar{x}=\frac{1}{k}\sum_{c=1}^{k}\frac{1}{n_c}\sum_{i=1}^{n_c}x_c^i$ is the mean of the dataset. In LDA, A_b and A_w are calculated as follows:

$$A_b=\sum_{c=1}^{k} n_c\left(W^T\mu_c-W^T\bar{x}\right)\left(W^T\mu_c-W^T\bar{x}\right)^T, \tag{1}$$

$$A_w=\sum_{c=1}^{k}\sum_{i}^{n_c}\left(W^Tx_c^i-W^Tu_c\right)\left(W^Tx_c^i-W^Tu_c\right)^T. \tag{2}$$

The inter-class and intra-class scattering matrices are obtained by directly calculating the centroids between the feature vectors. However, this calculation method cannot closely reflect the scattering relationship between samples and does not help find the best projection direction W in LDA. The radial basis function (RBF) is used to balance the linear relationship and the differences between features and has the function of smoothing data. Through the calculation using Eqs. (1) and (2), the RBF is introduced into the calculation formula of the scattering matrix

to achieve the information balance between each sample, which is conducive to finding the best projection direction in the process of dimensionality reduction.

In SDA, the values of the similarity matrices of the inter-class scattering matrix M and the intra-class scattering matrix N can respectively be calculated as

$$M^{ci} = \exp\left(\frac{-\left\|x_c^i - \overline{x}\right\|^2}{2g^2}\right), c = 1,\cdots,k; i = 1,\cdots,n_c \tag{3}$$

and

$$N_c^{ij} = \exp\left(\frac{-\left\|x_c^i - x_c^j\right\|^2}{2g^2}\right). i, j = 1,\cdots,n_c \tag{4}$$

Introducing the similarity matrix, the inter-class scattering matrix can be defined as

$$\begin{aligned} A_1(W) &= \frac{1}{n_c}\sum_{c=1}^{k}\sum_{i=1}^{n_c}\left[\left(W^T x_c^i - W^T\overline{x}\right)\left(W^T x_c^i - W^T\overline{x}\right)^T\right]M^{ci}, \\ &= \frac{1}{n_c}\sum_{c=1}^{k}\sum_{i=1}^{n_c}\left[W^T\left(x_c^i - \overline{x}\right)\left(x_c^i - \overline{x}\right)^T W\right]M^{ci}, \\ &= \frac{1}{n_c}\sum_{c=1}^{k}\sum_{i=1}^{n_c}W^T\left[\left(x_c^i - \overline{x}\right)\left(x_c^i - \overline{x}\right)^T M^{ci}\right]W, \\ &= W^T\left[\frac{1}{n_c}\sum_{c=1}^{k}\sum_{i=1}^{n_c}\left(x_c^i - \overline{x}\right)\left(x_c^i - \overline{x}\right)^T M^{ci}\right]W. \end{aligned} \tag{5}$$

Setting

$$L_1 = \frac{1}{n_c}\sum_{c=1}^{k}\sum_{i=1}^{n_c}\left(x_c^i - \overline{x}\right)\left(x_c^i - \overline{x}\right)^T M^{ci}, \tag{6}$$

we obtain

$$A_1(W) = W^T L_1 W. \tag{7}$$

The intra-class scattering matrix can be defined as

$$\begin{aligned} A_2(W) &= \sum_{c=1}^{k} \frac{1}{n_c^2} \sum_{i=1}^{n_c} \sum_{j=1}^{n_c} \left[\left(W^T x_c^i - W^T x_c^j \right) \left(W^T x_c^i - W^T x_c^j \right)^T \right] N_c^{ij}, \\ &= \sum_{c=1}^{k} \frac{1}{n_c^2} \sum_{i=1}^{n_c} \sum_{j=1}^{n_c} \left[W^T \left(x_c^i - x_c^j \right) \left(x_c^i - x_c^j \right)^T W \right] N_c^{ij}, \\ &= \sum_{c=1}^{k} \frac{1}{n_c^2} \sum_{i=1}^{n_c} \sum_{j=1}^{n_c} W^T \left[\left(x_c^i - x_c^j \right) \left(x_c^i - x_c^j \right)^T N_c^{ij} \right] W, \\ &= W^T \left[\sum_{c=1}^{k} \frac{1}{n_c^2} \sum_{i=1}^{n_c} \sum_{j=1}^{n_c} \left(x_c^i - x_c^j \right) \left(x_c^i - x_c^j \right)^T N_c^{ij} \right] W. \end{aligned} \tag{8}$$

Setting

$$L_2 = \sum_{c=1}^{k} \frac{1}{n_c^2} \sum_{i=1}^{n_c} \sum_{j=1}^{n_c} \left(x_c^i - x_c^j \right) \left(x_c^i - x_c^j \right)^T N_c^{ij}, \tag{9}$$

we obtain

$$A_2(W) = W^T L_2 W. \tag{10}$$

The optimization problem can be expressed as

$$\max A(W) = \max \frac{A_1(W)}{A_2(W)} = \max \frac{W^T L_1 W}{W^T L_2 W}. \tag{11}$$

The following eigenvalue problem can be solved to find the optimal projection direction:

$$L_1 \alpha = \lambda L_2 \alpha. \tag{12}$$

The optimization problem can be transformed into

$$\max \frac{\alpha_i^T L_1 \alpha_i}{\alpha_i^T L_2 \alpha_i}. \tag{13}$$

Because $L_1 \alpha = \lambda L_2 \alpha$, we have

$$L_1 \alpha_1 = \lambda_1 L_2 \alpha_1,\ L_1 \alpha_2 = \lambda_2 L_2 \alpha_2, \cdots, L_1 \alpha_2 = \lambda_i L_2 \alpha_i. \tag{14}$$

Finally, the optimization problem is transformed into

$$\max \frac{\alpha_i^T L_1 \alpha_i}{\alpha_i^T L_2 \alpha_i} = \max \frac{\lambda_i \alpha_i^T L_1 \alpha_i}{\alpha_i^T L_2 \alpha_i} = \lambda_i. \tag{15}$$

In the dimensionality reduction process, the cumulative contribution rate of the principal components determines the effectiveness of the dimensionality reduction results. In this work, we use the principal components whose cumulative feature contribution rate reaches more than 95% for pattern recognition.

2.4 Gas classification model

After obtaining the gas dimensionality reduction features, combining the multi-pattern recognition method is necessary to identify the gas features effectively, and then verify the effectiveness and universality of the feature dimensionality reduction method. In this work, we have selected four common pattern recognition algorithms in the process of gas classification.

Cortes and Vapnik first proposed the support vector machine (SVM) in 1995, which has many advantages in solving small samples, and nonlinear and high-dimensional pattern recognition. The principle of the SVM is to transform the input space into a high-dimensional feature space using a kernel function and find the optimal hyperplane in the new space. In this paper, the RBF is selected as the classification kernel function.[20] The genetic algorithm (GA) is applied to calculate the parameters affecting the classification performance of the SVM (the penalty factor c and the kernel function parameter g). In this study, the maximum number of generations in GA is 100, the population is 20, the search range of c is from 0 to 100, and that of g is from 0 to 1000. The fitness function is the highest accuracy rate of the training set under fivefold cross-validation.

The extreme learning machine (ELM) is a classification algorithm proposed by Hang of Nanyang Technological University.[21] It randomly generates the connection weights of the input layer and hidden layer and the threshold of the hidden-layer neurons. These parameters need not be adjusted during the training process.[21] Only the number of hidden-layer neurons must be set, and the optimal solution can be obtained. The ELM has the advantages of a high learning speed and good generalization performance. In this work, the number of hidden-layer neurons was 500.

The random forest (RF) is used in ensemble learning and is a method of integrating many decision trees into a forest, which is used to predict the final result.[22] The RF randomly selects k attributes from a sample set sampled by bagging to form a new dataset and then trains the decision tree. Compared with bagging, the RF introduces the perturbation of attributes. In this way, the difference between each trained subtree is as large as possible, the model after integration is difficult to overfit, and the generalization ability is enhanced. In practical classification tasks, the RF has excellent performance. In this work, k was 2.

Learning vector quantization (LVQ) is a neural network that combines competitive learning and mentored learning methods. LVQ includes input, competition, and output layers. In the

connection mode, the input and competition layers are fully connected, and the competition and output layers are partially connected.[23] An advantage of the LVQ neural network is its simple network structure, and it can complete very complex classification tasks by interacting with internal units. In this work, the number of hidden-layer neurons was 500, the learning rate was 0.01, and the LVQ network was trained for 1000 epochs.

The average of 20 calculation results was used as the final classification performance to eliminate the impact of randomness in the classification process. At the same time, to fully express the classification performance of the pattern recognition method, we used the classification accuracy, precision, and recall to represent the classification results.

3. Results and Discussion

3.1 Visualization of gas information

To compare the differences in the gas information of different production batches of beer, Fig. 3 shows a radar chart of the beer gas response information under different production batches at 60 s. It can be seen that the overall trend of the beer gas response information of different batches of the same brand is similar, and W2S, W1W, W1S, and W5S have greater gas responses than the other sensors. However, although the response values of the other sensors are

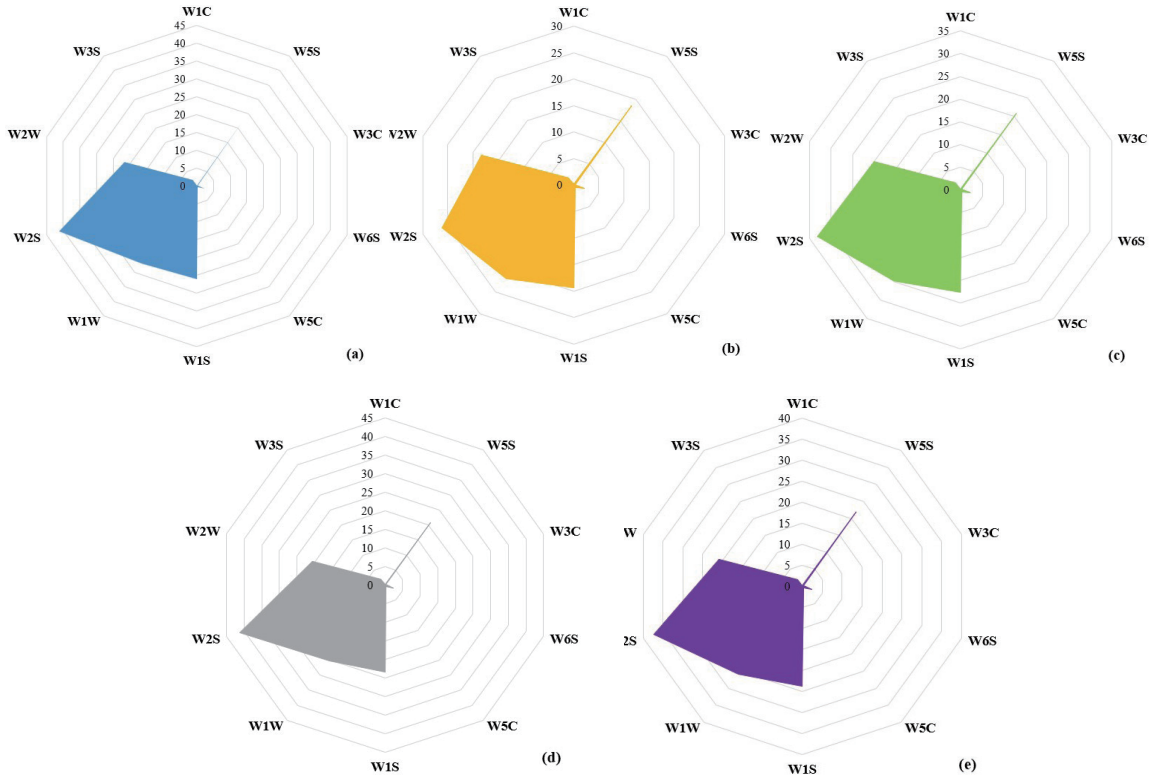

Fig. 3. (Color online) Radar chart of gas information. (a) First batch, (b) second batch, (c) third batch, (d) fourth batch, and (e) fifth batch.

small, the importance of a sensor cannot be determined by the magnitude of its response, which will easily lead to overfitting of the decision-making model. Therefore, it is important to comprehensively consider the gas features of each sensor response curve and combine them using effective feature processing methods to obtain the optimal feature set for realizing the effective decision-making of gas features.

3.2 Gas feature extraction

It can be seen from Fig. 2 that the detection signal of the beer gas detection feature is stable and that the detection curve increased rapidly then changed stably. In the early stage of gas detection, the dynamic headspace air quickly enters the sensor gas and undergoes a redox reaction with the sensor so that the detection signal rapidly increases in magnitude. With increasing detection time, the volatile gas consisting of the beer gas and the carrier gas reaches a dynamic balance, and the detection curve of the sensor tends to become stable. On the basis of the data characteristics of the beer gas information, we extract the maximum value, steady-state mean value, and integral value to represent the overall properties of the original detection signal. The maximum value represents the peak concentration of the gas entering the sensor gas chamber [Eq. (16)], the steady-state average value represents the state response of the volatile beer gas and the sensor gas reaching dynamic equilibrium [Eq. (17)], and the integral value represents the dynamic global response of the overall gas detection process [Eq. (18)]:

$$T_1 = \max\left|f(t)\right|, \tag{16}$$

$$T_2 = \frac{1}{10}\sum_{t=51}^{60} f(t), \tag{17}$$

$$T_3 = \int_1^{80} f(t)dt, \tag{18}$$

where $f(t)$ denotes the response value of the sensor at time t.

3.3 Dimensionality reduction visualization of gas features

To demonstrate the effectiveness of SDA in dimensionality reduction, we compare it with classical, unsupervised, and supervised dimensionality reduction methods. The comparison methods are PCA, kernel principal component analysis (KPCA), kernel entropy component analysis (KECA), and LDA. Figure 4 shows the results of dimensionality reduction for the different methods. It can be clearly seen that when PCA, KPCA, and KECA are used as unsupervised dimensionality reduction methods, although the beer gas features of different batches can be clustered to some extent, there is the disadvantageous feature of small differences between classes and large gaps within classes. Moreover, there is overlapping of the gas features of different batches of beer, which may not be conducive to gas identification. In the 2D

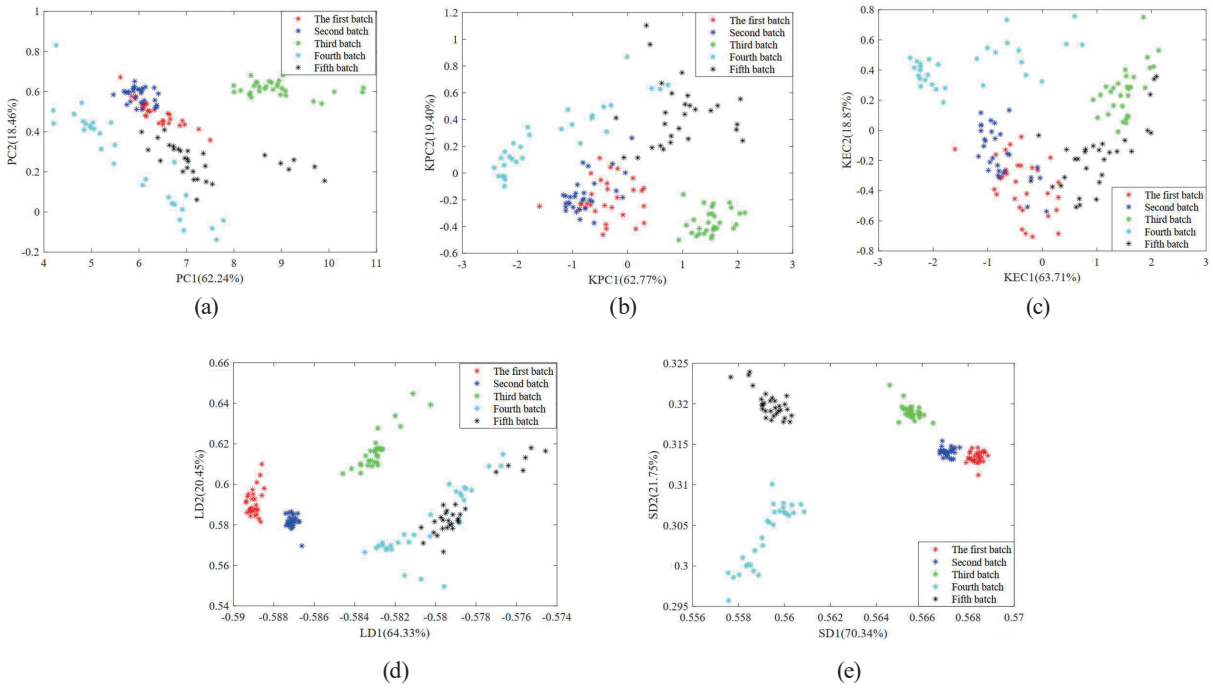

Fig. 4. (Color online) Dimensionality reduction results of gas features. (a) PCA, (b) KPCA, (c) KECA, (d) LDA, and (e) SDA.

dimensionality reduction results, the cumulative contribution rates of the first two principal components are 80.70% for the PCA dimensionality reduction method, 82.17% for KPCA, and 82.58% for KECA. Figures 4(d) and 4(e) respectively show the dimensionality reduction results of LDA and SDA. It can be seen that the supervised dimensionality reduction method makes the inter-class spacing significantly larger and the intra-class spacing smaller. Moreover, in the dimensionality reduction results of LDA, the beer gas information samples of the fourth and fifth batches overlap. The dimensionality reduction effect of SDA enables the beer gas information of different batches to be clearly clustered, and there is no overlap between samples in different categories, which is helpful for the intelligent identification of pattern recognition algorithms. In the 2D dimensionality reduction results, the cumulative variance contribution rates of the first two principal components are 84.78% for LDA and 92.09% for SDA. To obtain the optimal feature set for different dimensionality reduction methods, we employ the principal components with a cumulative variance contribution rate of more than 95% for pattern recognition. Finally, the optimal numbers of feature dimensions obtained by PCA, KPCA, KECA, LDA, and SDA are 12, 8, 7, 9, and 6, respectively.

3.4 Classification results

Figure 5 shows a histogram of the classification accuracy results of SVM, ELM, RF, and LVQ for different feature sets. It can be clearly seen that for the different classification algorithms, the feature sets obtained by the SDA dimensionality reduction method have the

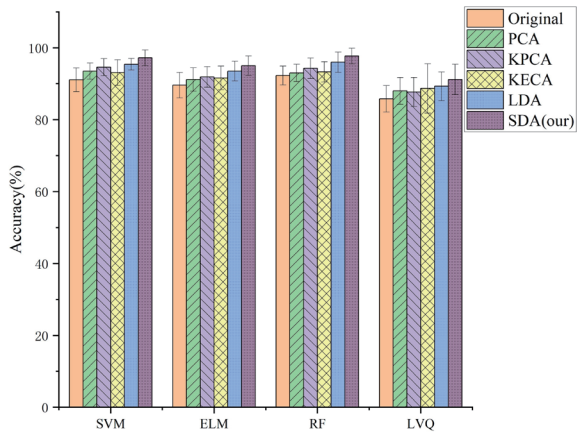

Fig. 5. (Color online) Histogram of classification accuracy of multi-pattern recognition algorithm under different feature sets.

Table 2
Classification accuracy of different feature dimensionality reduction methods combined with multi-pattern recognition algorithms (%).

Feature set	Dimension	SVM	ELM	RF	LVQ
Original	30	91.10 ± 3.28	89.60 ± 3.53	92.30 ± 2.62	85.80 ± 3.66
PCA	12	93.50 ± 2.24	91.20 ± 3.27	93.00 ± 2.38	88.00 ± 3.73
KPCA	8	94.60 ± 2.35	91.90 ± 2.86	94.30 ± 2.85	87.70 ± 4.01
KECA	7	93.10 ± 2.58	91.60 ± 3.28	93.30 ± 2.83	88.70 ± 6.88
LDA	9	95.40 ± 1.60	93.50 ± 2.74	96.00 ± 2.83	89.30 ± 4.01
SDA (ours)	6	97.20 ± 2.19	95.00 ± 2.71	97.70 ± 2.18	91.20 ± 4.18

Table 3
Precision of different feature dimensionality reduction methods combined with multi-pattern recognition algorithms (%).

Feature set	Dimension	SVM	ELM	RF	LVQ
Original	30	92.64 ± 2.66	91.05 ± 2.85	93.38 ± 2.28	88.27 ± 2.63
PCA	12	94.13 ± 2.22	92.63 ± 2.53	94.00 ± 2.28	89.90 ± 2.65
KPCA	8	95.35 ± 2.21	93.19 ± 2.52	94.74 ± 2.22	90.07 ± 2.86
KECA	7	94.07 ± 3.37	92.25 ± 2.73	93.74 ± 2.02	90.03 ± 5.36
LDA	9	96.25 ± 1.32	94.30 ± 2.20	96.63 ± 2.15	90.93 ± 3.05
SDA (ours)	6	97.73 ± 1.69	95.70 ± 2.14	98.47 ± 1.10	92.38 ± 3.40

Table 4
Recall of different feature dimensionality reduction methods combined with multi-pattern recognition algorithms (%).

Feature set	Dimension	SVM	ELM	RF	LVQ
Original	30	90.68 ± 3.26	89.10 ± 3.69	92.20 ± 2.88	84.66 ± 3.57
PCA	12	93.33 ± 2.21	90.78 ± 3.50	92.57 ± 2.37	86.84 ± 4.24
KPCA	8	94.37 ± 2.47	91.53 ± 3.16	93.45 ± 3.12	86.43 ± 4.32
KECA	7	93.00 ± 3.58	91.35 ± 3.38	93.45 ± 3.34	88.86 ± 5.95
LDA	9	95.44 ± 1.53	93.38 ± 2.90	96.10 ± 1.61	88.31 ± 4.31
SDA (ours)	7	97.24 ± 2.11	94.98 ± 2.88	98.23 ± 1.04	91.11 ± 4.27

highest classification accuracy and classification stability. At the same time, the classification effect of the feature set obtained by the supervised dimensionality reduction method is clear, and the effect of supervised dimensionality reduction is greater than that of unsupervised dimensionality reduction.

To comprehensively express the classification performance of the multi-pattern recognition algorithm, Tables 2–4 show the classification accuracy, precision, and recall of the multi-pattern recognition algorithm for different feature sets, respectively. It can be seen that for the multi-pattern recognition algorithm, the RF achieves the highest classification accuracy of 97.70%, the highest precision of 98.47%, and the highest recall of 98.23% for the SDA feature. Overall, the classification performance decreases in the order RF > SVM > ELM > LVQ. At the same time, as the recognition algorithms of neural networks, ELM and LVQ have poor classification stability. The SDA feature set obtained the best classification performance according to our comparison of the classification effects of different feature sets. Taking the RF classification algorithm as an example, the classification accuracy for the SDA feature set is 1.10% higher than that for the LDA feature set, 4.40% higher than that for the KECA feature set, 3.40% higher than that for the KPCA feature set, and 4.70% higher than that for the PCA feature set. The performance of other pattern recognition algorithms was also improved.

4. Conclusion

To achieve the effective dimensionality reduction of gas features, we proposed an SDA method combined with multi-pattern recognition technology to effectively identify beer gas information in different batches. The main conclusions of this paper are as follows:

(1) The data smoothing method using the RBF effectively avoids the disadvantage of dimensionality reduction caused by the nonlinear relationship between the features. Compared with other feature dimensionality reduction methods, it significantly increases the inter-class spacing and reduces the intra-class difference, improving the dimensionality reduction effect.

(2) In the classification results of the multi-pattern recognition algorithm, the RF achieved the highest classification accuracy of 97.70%, the highest precision of 98.47%, and the highest recall of 98.23% for the SDA feature set, enabling it to effectively identify beer gas information in different production batches.

The dimension reduction effect based on SDA can show the difference of unknown beer gas information. The classification results of pattern recognition are used to judge whether the gas information of different batches of beer is the same. Furthermore, the production process parameters of beer are adjusted. In this work, it provides a fast and effective detection method for beer quality control.

References

1 Y. Shi, F. Gong, M. Wang, J. Liu, Y. Wu, and H. Men: J. Food Eng. **263** (2019) 437. https://doi.org/10.1016/j.jfoodeng.2019.07.023

2 G. A. Helfer, J. L. V. Barbosa, E. Hermes, B. J. Fagundes, R. O. Santos, and A. B. d. Costa: Food Chem. **367** (2022) 130681. https://doi.org/10.1016/j.foodchem.2021.130681

3 L. F. Castro and C. F. Ross: J. I. Brewing. **121** (2015) 197. https://doi.org/10.1002/jib.219
4 G. Helfer, J. Barbosa, E. Hermes, B. Fagundes, R. Santos, and A. DaCosta: Food Chem. **365** (2022) 130681. https://doi.org/10.1016/j.foodchem.2021.130681
5 H. Men, Y. Shi, Y. Jiao, F. Gong, and J. Liu: Anal. Methods **10** (2018) 2016. https://doi.org/10.1039/C8AY00280K
6 H. Lin, H. Chen, C. Yin, Q. Zhang, Z. Li, Y. Shi, and H. Men: IEEE Sens. J. **22** (2022) 11463. https://doi.org/10.1109/JSEN.2022.3174251
7 S. Kang, Q. Zhang, Z. Li, C. Yin, N. Feng, and Y. Shi: Postharvest Biol. Technol. **197** (2023) 112214. https://doi.org/10.1016/j.postharvbio.2022.112214
8 Y. Shi, M. Liu, A. Sun, J. Liu, and H. Men: IEEE Sens. J. **21** (2021) 21175. https://doi.org/10.1109/jsen.2021.3079424
9 K. Qian, Y. Bao, J. Zhu, J. Wang, and Z. Wei: J. Food Eng. **290** (2021) 110250. https://doi.org/10.1016/j.jfoodeng.2020.110250
10 J. Zhang, Y. Xiong, and S. Min: Anal. Chim. Acta **1080** (2019) 43. https://doi.org/10.1016/j.aca.2019.06.054
11 S. Hazarika, R. Choudhury, B. Montazer, S. Medhi, M. P. Goswami, and U. Sarma: IEEE T. Instrum. Meas. **69** (2020) 9010. https://doi.org/10.1109/TIM.2020.2997064
12 M. Barriault, I. Alexander, N. Tasnim, A. O'Brien, H. Najjaran, and M. Hoorfar: Sens. Actuators, B **326** (2021) 129012. https://doi.org/10.1016/j.snb.2020.129012
13 J. H. Cho and P. U. Kurup: Sens. Actuators, B **160** (2011) 542. https://doi.org/10.1016/j.snb.2011.08.027
14 Y. Nozaki and T. Nakamoto: PLoS ONE **13** (2018) e0198475. https://doi.org/10.1371/journal.pone.0198475
15 S.-C. Wu, A. L. Swindlehurst, and Z. Nenadic: J. Neurosci. Meth. **253** (2015) 262. https://doi.org/10.1016/j.jneumeth.2015.07.003
16 W. Shin, T. Itoh, Y. Koyama, T. Akamatsu, A. Tsuruta, Y. Masuda, and K. Uchiyama: ECS Meeting Abstracts (2020) 1844. https://doi.org/10.1149/MA2020-01261844mtgabs
17 D. R. Wijaya and F. Afianti: IEEE Access **8** (2020) 33944. https://doi.org/10.1109/ACCESS.2020.2974982
18 G. Mishra, S. Srivastava, B. K. Panda, and H. N. Mishra: Anal. Methods **10** (2018) 5687. https://doi.org/10.1039/C8AY01921E
19 D. R. Wijaya and F. Afianti: IEEE Sens. J. **21** (2021) 476. https://doi.org/10.1109/JSEN.2020.3000756
20 S. A. Hosseini: J. Instrum. **14** (2019) P02008. https://doi.org/10.1088/1748-0221/14/02/P02008
21 Y. Shi, X. Jia, H. Yuan, S. Jia, J. Liu, and H. Men: Meas. Sci. Technol. **32** (2021) 025107. https://doi.org/10.1088/1361-6501/abb9e7
22 D. Alex, D. Chandy, A. Christinal, A. Singh, and M. Pushkaran: Int. J. Pattern Recognit. Artif. Intell. **36** (2022) 1. https://doi.org/10.1142/S0218001422560109
23 P. Melin, J. Amezcua, F. Valdez, and O. Castillo: Inf. Sci. **279** (2014) 483. https://doi.org/10.1016/j.ins.2014.04.003