International Journal of Engineering Science Invention
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org ||Volume 4 Issue 1 || January 2015 || PP.30-38

Primary Odor on Consideration of Reducing the Number of Compositions

Jing Li[1], Dehan Luo[1], Yunlong Sun[1]

[1]*School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, P.R. China*

ABSTRACT: *As is known to all, network coding has a byte limit. So when the chemical formula of odors is applied to the digital network coding, keep the number of parameters to a minimum for meeting the requirements of network coding byte limitation, less is better. The main purpose of this paper is to reduce the number of chemical compositions, namely the process of identifying "primary odor". The primary odors are emitted by the same 6-10 key chemical components of two different selected odors, which can affect the human olfactory senses to judge the class of odor identification. Testing the primary odors of the rational proportion under the chemical formula of two different odors diffused by two selected fruits by the electronic nose, the relationship between the primary odors in different composition ratio and the corresponding fruit smells has been investigated by Principal Component Analysis (PCA) according to the comparison of different sensors response. The results showed that after the feature extraction, two groups of samples have been classified as the homologous original fruit. Thus, it is feasible that simplified chemical formula is obtained based on primary odors, which makes reducing the number of parameters in network transmission possible.*

***KEY WORDS*：** *Odor coding; Olfactory components; Machine olfaction; PCA*

I. INTRODUCTION

Historically, unlike visual and audio information, olfactory information cannot be directly processed with a computer due to its difficulty digitization [1]. A proper data presentation method to present the sense of smell into multimedia information still does not exist, which may be an exploratory field in human computer interaction [2]. Although, Jeong-Do Kim et al. of Hoseo University argued that about forty emotional adjectives or smell characteristics such as odor threshold, intensity, persistence, and hedonic tone could be used in describing odor characteristics [2-3], there are still many problems when using such highly subjective methods to convert the odor information into digital data for representation so far. However, there is a relatively objective and accurate method of digitizing odor---the proven techniques that some fragrance products (such as flavors and fragrances, perfumes) made in the factory actually are the process of indirectly quantified smells into chemical composition and proportion [4] as a reference---that the chemical formula and the percentage of the ingredients are used as the characteristic parameters and characteristic values in the network transmission of odors. Here, for the first time, the chemical compositions of the smell and its ratio are applied to the network transmission of odor coding.Taking the chemical compositions of all smells into account, too many parameters can lead to a large amount of data; furthermore, some smells does not contain the characteristic parameters that will cause the waste of space and result in the large data streams.

As we know, the standard of Ethernet frame that the Ethernet II and IEEE802.3 are the most common LAN frames is unchanged and the data of the Ethernet transmission is limited [5-6]. Ethernet II can be mounted in the length of the data is 46---1500; IEEE802.3 SAP can be installed with the length of the data is 43---1497; IEEE 802.3 SNAP length of data that can be loaded is 38---1492 [7-9]. Hence, the length of the data is being transmitted as short as possible. How to reduce the number of chemical formulations and proportions of the selected odors, which are the only way that can change the number of data [10], for being well indicated for the smell of network transmission data frame format?

In this paper, reducing the chemical compositions of two different selected odors for controlling the length of tele-olfactory based on the primary odors, which are emitted by the same 6-10 key chemical components of the two experimental odors. To explore this hypothesis, two selected odors using in this study are being respectively circulated by two common fruits---the apricot and pear, so the primary odors are determined on the basis of 6 kinds of chemicals ingredients--- isopentyl acetate ($C_7H_{14}O_2$), ethyl acetate ($C_4H_8O_2$), butyric acid ethyl ester ($C_6H_{12}O_2$), eugenol ($C_{10}H_{12}O_2$), citrus sinensis oil, vanillin [11]. Detected and analyzed by the electronic nose (PEN3) on four kinds of smells, which are came out from four different samples---the apricot and pear and different proportions of the primary odors based on the chemical formula of the two chosen fruits, the original data will generated that is reflected the corresponding value by the sensors response.

Comparing the difference between the four groups response value, the obtained original data has been processing by PCA algorithm to reduce the dimension of data and generate respectively plentiful feature subsets to form feature parameter vectors [12]. According to the distribution map of four samples after dimension reduction based on PCA, the primary odors and the corresponding fruit smell are closer. Furthermore, the classifications about the different ratio of primary odors into the fruit smells are more precisely demonstrated that there are some properties in the primary odors to be classified as the original odor. These results verify that reducing the number of encoding parameters is possible. Hence, these features can be used to the smell encoding.

II. MATERIALS AND METHODS

Experimental samples : This study was carried out using apricots and pears and different proportions of the primary odors based on the chemical formula of the two chosen fruits as described in Table 1. Viewing the chemical ratio of apricots and pears, the primary odors respectively account for 57.3% and 96.5% in the proportion. In order to make the external variables are consistent, namely to make two samples achieve the same overall capacity, also to encode parameters taking into account of 100% per cent, water which is chemically defined colorless and odorless add to sample 1 and sample 2 to make the chemical ingredients of dissolution that should join the part of ingredients. Moreover, in order to make the smell of apricots and pears better come out, we use the mixer to stir apricots and pears, which obtained as samples 3 and 4.

Table 1 Information of samples used in this experiment

Sample / Composition(g)	1	2	3	4
sopentyl acetate	19	40		
ethyl acetate	16	40		
Butyric acid ethyl ester	19	7		

Eugenol	0.3	0.5		
Citrus sinensis Oil	0.5	7		
Vanillin	2.5	2		
Water	42.7	3.5		
Others			The mixture after stirring apricots	The mixture after stirring pears
Total capacity	100	100	100	100

PEN3 : Samples were tested on the electronic nose (PEN3) made by AIRSENSE Analytics GmbH in Schwerin, Germany. PEN3 is equipped with an array of 10 different MOS sensors positioned into a small chamber (V=1.8ml), a sampling apparatus and a pattern recognition software named WinMuster [13]. Table 2 summarizes the sensitivity list of all sensors in PEN3.

Table 2 The sensitivity list of 10 sensors in PEN3

Number in array	Sensor name	Sensitive to	Detection range/ ppm
S1	W1C	Aromatic components	10
S2	W5S	Nitrogen oxides, very sensitive	1
S3	W3C	Ammonia and aromatic components	10
S4	W6S	Mainly hydrogen, selectively, (breath gases)	100
S5	W5C	Alkanes and aromatic components	1
S6	W1S	Propane	100
S7	W1W	Sulfur organic compounds	1
S8	W2S	Ethanol	100
S9	W2W	Aromatic components and organic-sulfides	1
S10	W3S	Propane (selective sometimes)	100

The data of sensors response collected by PEN3 is defined as the ratio of conductance: G/G0. G represents the resistance of each sensor in the chamber after the exposition to the headspace gas in the vial and G0 represents the resistance while the sensors expose to the zero gas filtered by active carbon. The initial value of the response data is 1, and the normal floating is range from 0.5 to 1.5. The experiment for data collection was carried out in an air-conditioned laboratory that the environment parameters set up as follows: the temperature was kept at 26℃-28℃ and the humidity at 52%-56%. Static headspace sampling (SHS) that is the most common technique for its accessibility [14] was used in this experiment. In the process of measuring, the headspace of each vial was pumped over the sensors in PEN3 at a constant flow speed of 400 ml/min. Furthermore, the collected data set would be voluntarily preserved in a computer connected to PEN3 in advance. The headspace gas of each vial of samples was measured 12 times continuously. Thus 48 data sets were collected for all 4 groups of samples.

Packets constituting odor : When we began to encode the two selected odors using the way of "primary odors", the number of the primary odors was the first thing being determined to n(6<=n<=10), and then, the chemical ingredients and the ratio were being corresponding defined to C1,C2,C3....Cn and R1,R2,R3 ... ... Rn. The relationship between the primary odors with the chemical formula of two selected odors showed in figure 1 that there were not the components of the primary odors in the observed odors, in the other word that the sum of the defined ratios R1, R2, R3 ... ... Rn was less than 1 which is absolutely broke the law that each item must be added up to 100%.

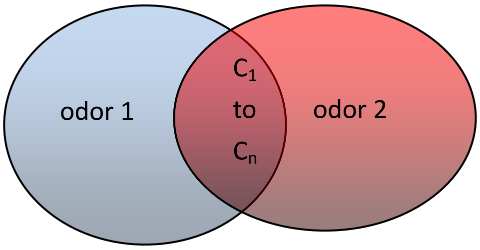


Figure 1 the primary odors with the chemical formula of two selected odors

In order to obey the rule, water which is chemically defined as colorless and odorless was added in the coding parameters, namely, the primary odors and water were the different factors in encoding. The ratio of water set to R, and the relationship between R and R_1, R2, R3 ... ... Rn was described in formula (1).

$$R+R1+R2+\ldots\ldots+Rn=1 \tag{1}$$

PCA : Principal component analysis (PCA) algorithm is one of the most commonly used methods for the electronic nose data processing because its dimension reduction [15-17]. The process of reducing the dimensionality is to take a minority of linear combinations, namely comprehensive indexes, from the original multiple variables. Although these indicators cannot be directly observed, they are linearly independent between any two. In addition, they can maximally keep the original variable information, thus determining the internal clustering features between the samples. In order to gain a good visualization, principal component analysis (PCA) is used to analyze the extracted data sets by Matlab. PCA score plots are conducive to view where measurements spread and whether the electronic nose has enough resolutions to distinguish different samples [18]. PCA analysis also helps to find out the sample that is different from others and variables contribute most to this difference [19]. The role of PCA in this experiment is two aspects: on the one hand, plot the comparison chart about the projection space of the samples after reducing dimensions, on the other hand is used for the identification of similar samples. The process of dimensionality reduction is following:

Q is defined as the covariance matrix which is calculated according to the sample matrix T achieved from the original data. And the formula of definition as follows:

$$\mathrm{Q} = \mathrm{T} \times \mathrm{T}^{T} \tag{2}$$

The definition of Q is a real number, namely $\mathrm{Q} \in R^{r \times r}$.

According to the value of Q, eigenvalues and eigenvectors also have been got. Defined the PCA coefficient matrix P_c is comprised by the eigenvectors determined by the coefficient larger eigenvalues, which the number of values is defined a. P_c is matrix composed by many real numbers, and $P_c \in R^{r \times a}$. Training the sample matrix T projects onto the PCA feature subspace, and then the PCA recognition feature is defined as follows:

$$P_f = P_c^T \times T \tag{3}$$

The definition of Q is a real number, namely $P_f \in R^{a \times N}$. N is determined by the dimension of the matrix of the sample matrix T.

III. RESULTS AND DISCUSSION

The disturbances of water : Water is colorless transparent liquid under atmospheric pressure [20]. There is no doubt that the human nose cannot be able to detect the flavor of water. In order to eliminate the interference of water on this experiment, figure 2 shows the response data of ten sensors in PEN3 on 100ml water and all air in 200ml bottle. Diagram showed that the ratio of conductance---G/G0 range from 0.8 to 1.2, which the response data was in the normal floating, that there was not without response on air and water. Viewing the curve of the two pictures, excluding the disturbances of air and the errors made by PEN3 itself, water hardly caused the reaction of sensors.

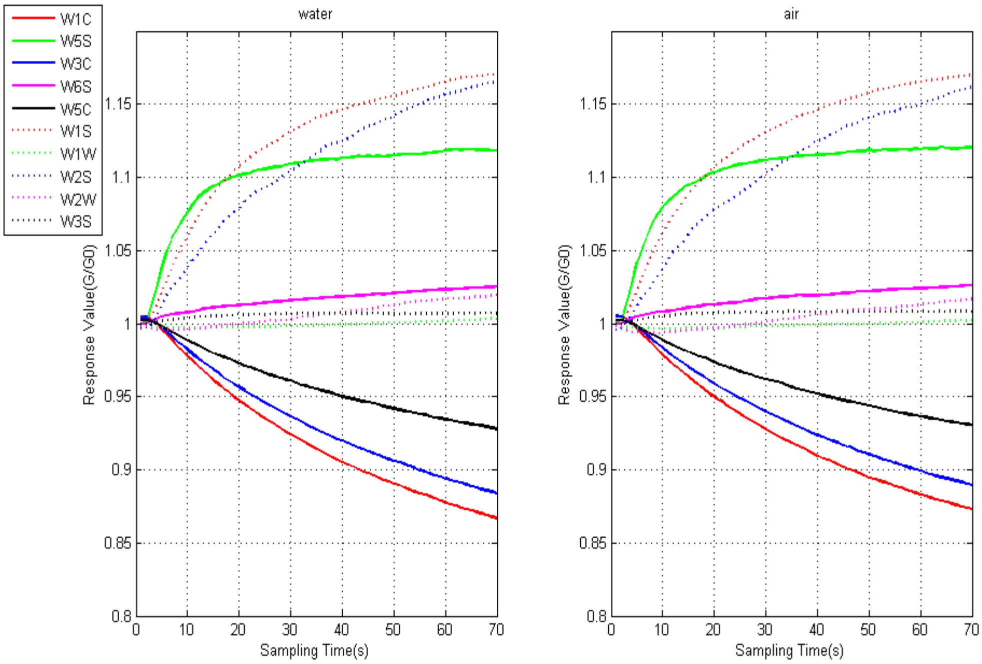


Figure 2 the response curve of water and air

The fingerprints : In order to comprehensively observe the response values of four samples measured by PEN3, the fingerprint that was four four-order curve fitting made by function polyfit in matlaB based on the arithmetic mean of various sensor response values was showed in figure 3. The fitted curves showed that the relationship between sample 1 and sample 2 was similar to two parallel curves without any intersections, which was depending on the ingredients used in experiments that sample 1 and sample 2 using the same chemical materials. Compared the distance of four curves, the gap between sample 1 and sample 3 was smaller, apparently, while the response data of sample 2 and sample 4 was more similar. Obviously one point about four curves is that the sensors response data of sample 1 and sample 3 is clearly higher than sample 2 and sample 4, which is because the reaction of sensors is more sensitive to pear.

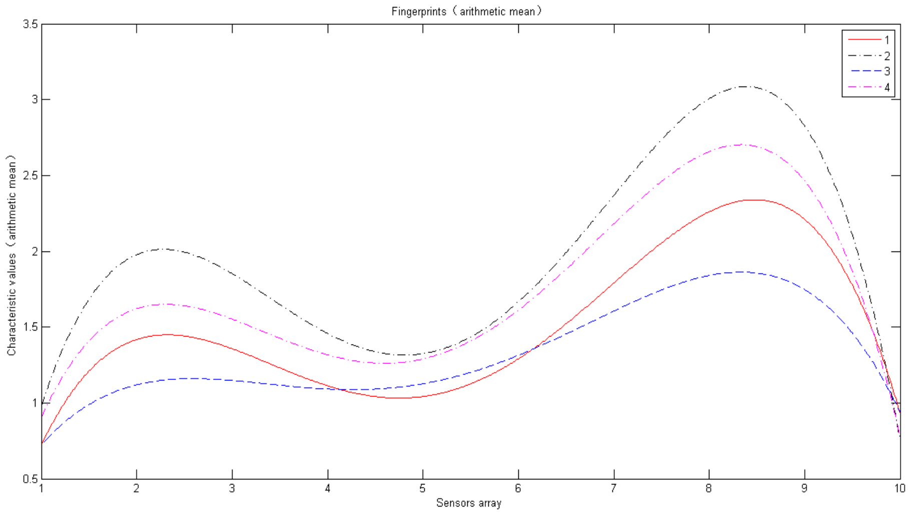


Figure 3 the fingerprints of four samples

Principal Component Analysis : Initially, the original data measured by PEN3 was large numbers and higher dimensions that there were a lot of useless information that has nothing to do with the data analysis. Hence, in the premise of without losing the useful information about classification as far as possible, the data which contains the most effective features for classification and a lower dimensions compared with the customary data would be got through reasonable principles that were set up to transform combine, and select from the original data, which the following characteristics that the response data of 30s, 40s, 50s, 60s that were the stability time in every 70s measurement time, the arithmetic mean and variance of each sensor response value, the differential value and integral value of each curve of fingerprints and the coefficient value of four curve fittings satisfied those conditions were selected. After analyzing by PCA, multiple variables through linear transformation to choose less number important variables of a multivariate statistical analysis were showed in figure 4, the envelope interval of sample 2 and sample 4 had a limited overlap for the reason that the similarity of formula is 96.5%, while a huge gap between the interval of sample 1 and sample 3, because the similarity is just only 57.3%.

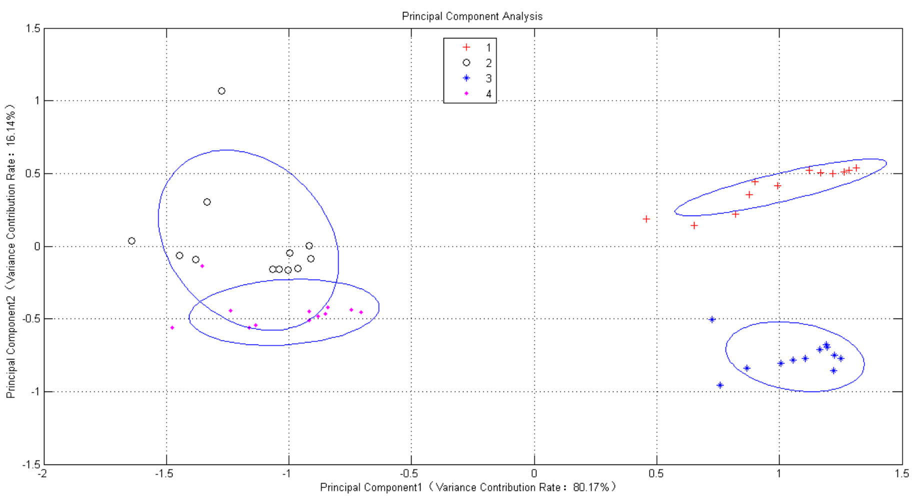
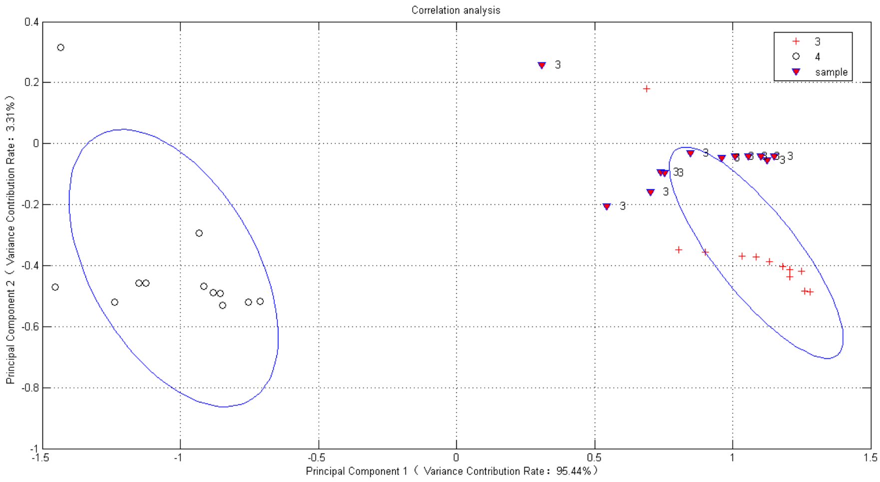


Figure 4 Principal Component Analysis of four samples

Correlation Analysis : In order to judge whether sample 1 and sample 2 that were the synthetic odors according to water and the primary odors based on the chemical formula of sample 3 and sample 4 dropped into the corresponding original smell by PCA, sample 1 and 2 were chosen as the tested specimens, and the sample 3 and 4 were chosen as the standard specimens. Figure 4 was respectively shown that sample 1 and 2 as the tested specimens were classified by PCA, obviously in the pictures that the tested sample was identified as the corresponding standard sample. In addition, the projected points of sample 2 were closer to the center of the envelop curve of the sample 4 than sample 2 to sample 4, which was on account for the different ratios of the primary odors. The results show that in the PCA identification, the samples comprised by the primary odors can be projected into the corresponding fruit samples.



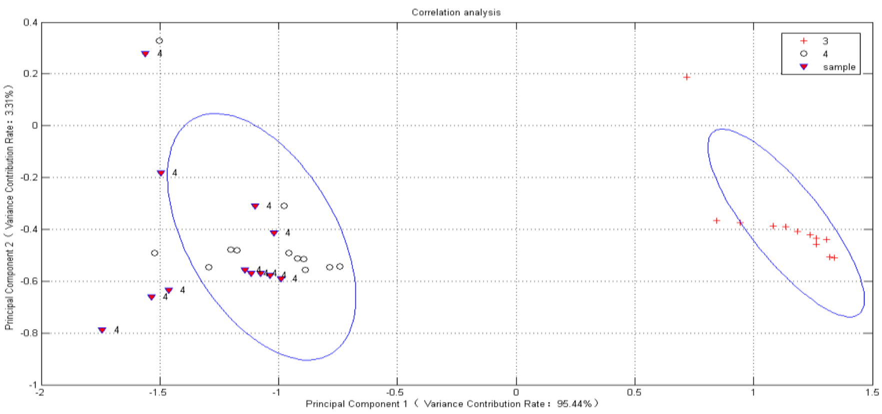


Figure 4 Principal Component Analysis of four samples

IV. CONCLUSION

For the purpose of reducing the number of chemical compositions of two selected odors according to its chemical formula, the four samples (two samples comprised by the primary odors and two by the original odors) have been tested by PEN3. The results show that the sensors response data set is a visual representation of the gap on different samples and the relationship, too. Besides, the outcomes of the collected data analyzed by PCA are more precisely demonstrated that the samples comprised by the primary odors are similar to the original odors; indeed, the identification between the synthetics and the originals is the more evidence. Therefore, the primary odor that is put forward on the consideration of reducing the number of chemical ingredients is successfully proved useful.In this paper, on the premise of a conclusion verified on theory that the chemical formula can be applied to the network transmission of data format, it is feasible that the original odor can be reducing the number of chemical components by the primary odor. And the characteristic parameters and characteristic values in the data format of network transmission can be the chemical formulations after reducing by the primary odors, which is theoretically validated. There is no doubt that the primary odor is benefit for the process of digitizing smell. In the future, the e-nose systems as a general measurement tool for identifying and classifying odors might be more widely used. Further work needs to be carried out in order to prove that the primary odor can be the characteristic parameters and characteristic values in the data format of network transmission.

V. ACKNOWLEDGEMENTS

The authors acknowledge the financial support of the Guangdong Province Foundation of Nature and Science through Project S2011020002906.

REFERENCES

[1] Imahashi M, Hayashi K. Odor clustering and discrimination using an odor separating system[J]. Sensors and Actuators B: Chemical, 2012, 166: 685-694.

[2] Jeong-Do Kim and Hyung-Gi Byun , A Proposal of the Olfactory Information Presentation Method and Its Application for Scent Generator Using Web Service[J]. Journal of Sensor Science and Technology, Vol. 21, No. 4, pp249-255, 2012

[3] Jeong-Do Kim, Dong-Jin Kim, Dong-Won Han et, A Proposal Representation, Digital Coding and Clustering of Odor Information[C].IEEE SENSORS 2006 CONFERENCE,pp872-877, 2006

[4] Scott A. The French Stench, The English Pong, The Cheesy Norwegians [J]. 2013.

[5] Birrell A D, Nelson B J. Implementing remote procedure calls [J]. ACM Transactions on Computer Systems (TOCS), 1984, 2(1): 39-59.

[6] Fukuda K, Takayasu H, Takayasu M. Origin of critical behavior in Ethernet traffic[J]. Physica A: Statistical Mechanics and its Applications, 2000, 287(1): 289-301.

[7] Leland W E, Taqqu M S, Willinger W, et al. On the self-similar nature of Ethernet traffic (extended version)[J]. Networking, IEEE/ACM Transactions on, 1994, 2(1): 1-15.

[8] Kramer G, Pesavento G. Ethernet passive optical network (EPON): building a next-generation optical access network[J]. Communications magazine, IEEE, 2002, 40(2): 66-73.

[9] Lian F L, Moyne J R, Tilbury D M. Performance evaluation of control networks: Ethernet, ControlNet, and DeviceNet[J]. Control Systems, IEEE, 2001, 21(1): 66-83.

[10] Kramer G, Mukherjee B, Dixit S, et al. Supporting differentiated classes of service in Ethernet passive optical networks[J]. Journal of Optical Networking, 2002, 1(9): 280-298.

[11] Lin Xiangyun. Perfumery [M].Chemical Industry Press (CIP) ,2008

[12] Yang J, Zhang D, Frangi A F, et al. Two-dimensional PCA: a new approach to appearance-based face representation and recognition[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2004, 26(1): 131-137.

[13] Dehan Luo, Yu Sun, Jiajun Zhuang H. Classification of Hundred-grass-oil Samples Using E-nose[J]. Journal of Computational Information Systems, 2013, 9(7): 2659-2666.

[14] Antihus Hernández Gómez, Jun Wang, Guixian Hu, et al. Electronic nose technique potential monitoring mandarin maturity. Sensors and Actuators B, 113(2006):347–353.

[15] Jolliffe I. Principal component analysis[M]. John Wiley & Sons, Ltd, 2005.

[16] Abdi H, Williams L J. Principal component analysis[J]. Wiley Interdisciplinary Reviews: Computational Statistics, 2010, 2(4): 433-459.

[17] Moore B. Principal component analysis in linear systems: Controllability, observability, and model reduction[J]. Automatic

Control, IEEE Transactions on, 1981, 26(1): 17-32.

[18] Simona Benedetti, Susanna Buratti, Anna Spinardi, et al. Electronic nose as a non-destructive tool to characterise peach cultivars and to monitor their ripening stage during shelf-life. Postharvest Biology and Technology, 47(2008):181–188.

[19] Huichun Yu, Jun Wang. Discrimination of LongJing green-tea grade by electronic nose. Sensors and Actuators B, 122(2007):134–140

[20] Nockemann P, Binnemans K, Driesen K. Purification of imidazolium ionic liquids for spectroscopic applications [J]. Chemical physics letters, 2005, 415(1): 131-136.