Study on Key Technology for the Discrimination of Xihu Longjing Tea Grade by Electronic Tongue

Bolin Shi, Houyin Wang, Lei Zhao, Ruicong Zhi, Zhi Li,
Lulu Zhang, and Nan Xie

Food and Agriculture Standardization Institute, China National Institute of Standardization,
Beijing 100191, P.R. China

Abstract. Electronic tongue has the characteristics of sensitivity and instability. However the technical specification for its research has still not been formed. In this paper, it was introduced the key technology of electronic tongue qualitative discriminant analysis to micro-difference samples. The research objects were four different grades of Xihu Longjing Tea with less difference within small producing area in Hangzhou of Zhejiang Province in China. According to the research, the stability of equipment had not been shown until the fifth repetition for the same sample by the electronic tongue. Finally, the signal from seventh repetitive test was selected to represent the intelligent taste fingerprint for this sample. The fingerprints of electronic tongue collected at different days showed linearity drifting for the same sample. The special tea samples were set into each series of experiments to be considered as the reference sample. All samples' signals were calibrated with the difference value between the reference sample fingerprint of corresponding test to the designated reference sample. Principal component analysis (PCA) results showed that the same grade samples were clustered, while the different grades samples were more dispersion and non-overlapping. Through Mahalanobis Distance and Residual Method, the four abnormal tea samples were rejected from highest and 1st grade respectively. The electronic tongue's Longjing Tea Grade models were built by soft independent modeling of class analogy (SIMCA). The discrimination accuracy for tea sample grade were both 100% for correction set and prediction set. Through this study, it was established the technical specification and flow for the quick detection of tea by electronic tongue, which including the determination on intelligent taste spectrum repeatability performance, system error calibration for spectrum drifting, rejection for abnormal tea sample based on taste spectrum and establishment for the quality judgment model . The technical specification provides the theoretical base for reasonable use of electronic tongue.

Keywords: Electronic Tongue, Longjing Tea, Grade, Shifting Calibration, Repeatability, Abnormal Sample.

1 Introduction

The 30~48% contents of tea are water-soluble substances, which represents the taste of tea soup and directly reflect the tea quality and grade (Rahim etc., 2014). Currently, the commonly used analysis technologies for flavoring matter of tea are

D. Li and Y. Chen (Eds.): CCTA 2014, IFIP AICT 452, pp. 379–392, 2015.
DOI: 10.1007/978-3-319-19620-6_43

liquid chromatography, spectroscopy, mass spectra and nuclear magnetic resonance method as well as the combination and cooperation between them. However, the flavoring ingredients are extremely complex. The whole quality of tea is not composed by certain or some kinds of taste matters but the comprehensive performance by dozens, even by over hundreds kinds of flavoring matters. The researchers were difficult to detect these ingredients one by one, and they can only figure out some ingredients with main flavor, which makes it difficult to test the taste of tea comprehensively (Shi Bolin etc., 2009).

The electronic tongue is a kind of electronic intelligent identification system developed by simulating human body's taste mechanism and is a kind of novel food analysis, identification and testing technology developed within recent years. Being different from the common analyzer, the result obtained by the electronic tongue is not the qualitative or quantitative for the certain kind or some kinds of ingredients in the tested sample but is the overall taste information of sample which is also called as "Taste Fingerprint" data (Jiang Sha etc., 2009). IVARSSON etc. utilized the electronic tongue to appraise nine different kinds of tea and obtained the satisfied results through combining the pattern recognition method with multivariate analysis and principal component analysis. Lvova etc.(2003) used the electronic tongue microsystem with full solid status to quantitative determine the multiple ingredients of green tea from Korea. Chen Quansheng etc. (2008) utilized the electronic tongue technology to perform the classification and identification study for pan-fired green tea with different grades through combining the identification method for K nearest neighbor domain and neutral network mode. He Wei etc. (2009) applied the electronic tongue technology into the grading and classification study for Pu'er tea to study the correlation with sensory evaluation result.

In the former tea quality study, the electronic tongue was employed to test the difference among different types of tea (such as, red tea, black tea and green tea) (Wu Jian etc., 2006; Liu Shuang etc., 2014) or the difference among the same kind of tea with various geographic origins (such as, green tea classification in Zhejiang, Fujian, Anhui, Jiangxi Province) (Wu Xinyu etc., 2007; Runu Banerjee etc., 2012). Since the quality difference was greater among samples in the these studies, it was easy to obtain the satisfied prediction result. Although these studies were specially taken into account the excavation of electronic tongue signal and the application of electronic tongue, the corresponding technical specification and systematic key technology solution on electronic tongue application were not formed in accordance with the electronic tongue technology advantages and the existing technology level. In order to conduct the rapid detection research of electronic tongue for different grades of Xihu Longjing Tea, three key elements should be taken into account. They were repeatability performance analysis of intelligent taste spectrum for electronic tongue, error calibration for spectrum drifting system and rejection for outlier tea samples based on taste spectrum. It was helpful to explore the solution on repeatability and reproducibility of electronic tongue instrument as well as the importance of selecting representative samples in classified model. This paper revealed the key technology and technical specification in the rapid detection of electronic tongue and reflected the value of electronic tongue test for micro-difference sample (such as Xihu Longjing Tea with different grades) in the practical application.

2 Experiments and Methods

2.1 Sample Preparation

Xihu Longjing Tea Sample from 2013 year with four grades (highest grade, special grade, 1st Grade and 2nd Grade) were collected by Zhejiang Hangzhou Standardization Research Institute from Xihu Longjing Tea production area. The tea samples were sub-packaged in 3 g/ bag individually by the aluminum foil bag with good sealing performance (Beijing Huadun Plastics Co., Ltd., 10cm*10cm, food grade, avirulent and insipidity), and were stored in the refrigeration house below −4℃ (Xu Yanjun etc., 2004). According to the experimental dosage, several small bags were conveniently taken each time.

The preparation for tea soup used for electronic tongue detection was as follows. 1.00 g tea sample was brewed into 150 mL boiling ultrapure water. Covered the watch glass, the tea sample was put into 100℃ boiling water batch for continuous digestion (DK-98-11A water bath kettle was made by Tianjin Taisite Instruments Co., Ltd.), and was stirred every 10 min at a time. After 45 min, the sample was suction filtration. And the sample was tested by electronic tongue two hours later (Xue Dan etc., 2010).

2.2 Intelligent Taste Acquisition Method

ASTREE electronic tongue made by Alpha MOS in France was adopted, mainly including automatic liquid sampler LS16, electrochemical transducer array (including 7 chemical sensors and 1 reference electrode) and data acquisition unit. Prior to tea sample testing, the electronic tongue was subjected to such procedures as self-inspection, activation, training and calibration so as to ensure the equipment reliability and stability.

Every fingerprint of sample was collected according to the flow of “Tea Soup Sample (120s) Cleaning Fluid No. 1 (10s) Cleaning Fluid No. 2 (10s)”. Each sample was repeatedly tested for seven times and provided with two same cleaning fluids (Figure 1). After all samples were detected, the probe of electronic tongue sensor was stopped in the ultrapure water. Finally, the sensors were completely cleaned based on the setting cleaning procedure. The detection parameters of electronic tongue were presented in Table 1.

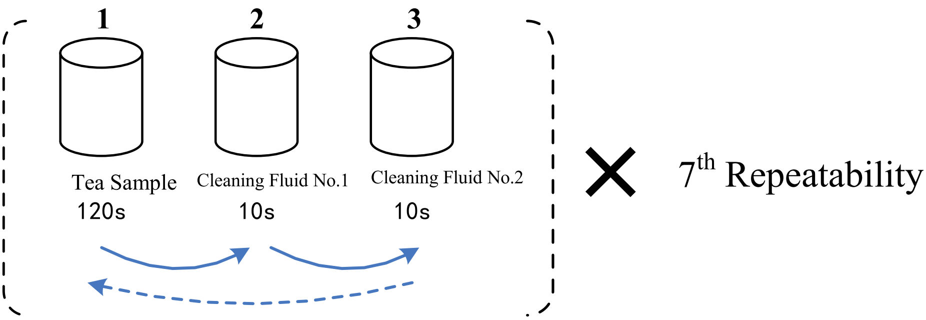

Fig. 1. The steps of sample test

Table 1. The detection parameters of electronic tongue

Parameter Items	Parameter Value
Activation, Training and Calibration Solution	0.0lmol/L HCl
Cleaning Solution	Ultrapure Water
Cleaning Time	10 s
Stirring Speed	1 r/s
Volume of Sample	25 mL
Volume of Sample Cup	40 mL
Sampling Time	120 s
Repeated times of Determination	7 times
Sampling Temperature	25℃ for Room Temperature

2.3 Multivariate Statistic Method

Mahalanobis was widely applied in the outlier judgment, because the method was used to measure whether the sample affects the whole samples set (Shi Bolin etc., 2010). If the Mahalanobis value of certain sample was too high, the regression model would have greater dependency on this sample, which would affect the stability of model. In other words, this sample was an outlier sample.

Principal component analysis（PCA）was used for data dimension reduction (Shi Bolin etc., 2012). PCA is a statistical technique that is used to analyze the interrelationships among a large number of variables and to explain these variables in terms of a smaller number of variables, called principal components, with a minimum loss of information. The principal component scoring not only reflected the similarity and peculiarity among tea samples, but also revealed the internal characteristics and clustering information of samples. It represented that whether the sample had greater difference in various categories of sample set, and whether the automatic clustering phenomenon was formed in accordance with the quality characteristics among samples.

Soft independent model classification analysis (SIMCA) was adopted to establish the classification identification model with different grades for Xihu Longjing Tea (Shi Bolin etc., 2011). The foundation of SIMCA was principal component analysis (PCA). Each class of samples had an independent model which was trained by PCA. Meanwhile the residuals of each class were generated. The unknown sample was brought into every class to compare the residuals in different models.

All algorithms were programmed by Matlab 7.0.

3 Results and Discussion

3.1 Selection of Stable Tea Spectrum for Electronic Tongue

The response value that the electronic tongue obtained was the difference value between electric potential of tested sample and electric potential of reference electrode (R-R0, where R0 was electric potential value for reference electrode, and R was electric potential value of sensor of tested sample). Figure 2 was the response curve of certain sample on seven sensors within 120s, from intense response originally to stability finally. The response on the final stage included the whole information for tea sample. Moreover, the steady value at 120 s reflected a minimum relative standard deviation (RSD) for the same sample, and a maximum distinction for the different samples. Therefore, the steady value at 120 s was the characteristic response signal for subsequent study.

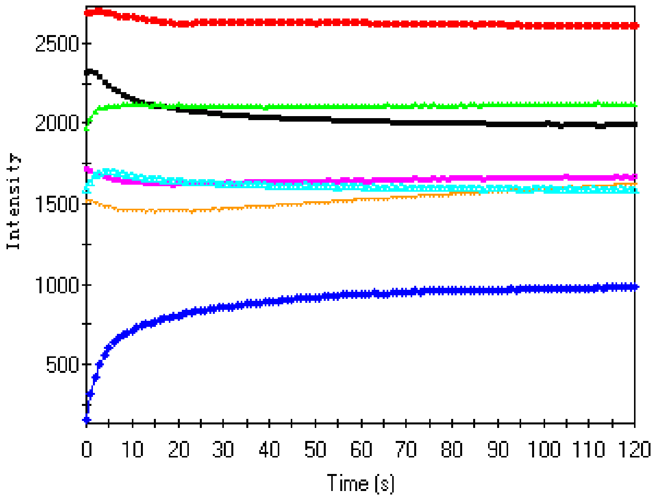

Fig. 2. The response curve for electronic tongue of a certain Sample

For electronic tongue equipment, it was required to guarantee the repeatability of detection signal, otherwise, the reliability of testing data may be queried. The animal's gustation can perfectly catch the taste characteristics of sample after taste training continuously. Similarly, the taste fingerprint from electronic tongue would be more reliable after repeated testing. In the previous researches, the fingerprint information of electronic tongue was used by only one time test result or the mean response signal after three-time repetition. In this study, 7-time tests were repeated simultaneously for each sample. Figure 3 was the PCA scoring for four different grades samples after 7-time repeated tests. The figure showed that the signal repeatability of electronic tongue was poor in the former three-time testing. However, with the increase of repetition time, the testing results were tended to be stable continuously. Especially, the testing results for final three times showed good repeatability to the same sample and great difference among the various grades samples. Therefore, the electric potential value at 120s acquired during 7th repetition was selected to be the original information of intelligent taste for Xihu Longjing Tea. In the repeatability test of electronic tongue, the signal choosing was solved.

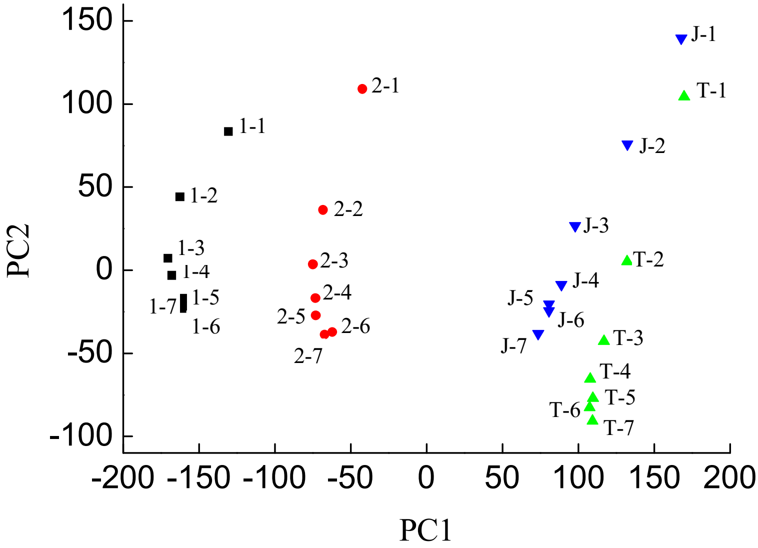

Fig. 3. PCA scores of four grading samples under seven repeating determination (The Highest Grade, Special Grade, 1st Grade and 2nd Grade Products were expressed as “J”, “T”, “1” and “2” in figure.)

3.2 System Drifting Correction for Electronic Tongue Spectrum of Tea

In order to give play to the advantage of electronic tongue’s rapid testing, a long-term prediction model should be established based on lots of samples that had the characteristics of representativeness, typicality and comprehensiveness, so as to make sure of the model robustness and stability. These tea samples were usually detected for several days by electronic tongue. In the same day, the 7th repeated result was used as the sample fingerprint to solve the repeatability. Due to influence by various complicated factors such as environment, the spectrums of the same sample were shifted seriously in different days test, leading to bad reproduction. Therefore, the results would be meaningless if the original data was used directly and the accurate conclusion would be difficult to be made in this case.

The same experimental operation was adopted on each tea soup sample during the whole test period, lasting for a month and acquiring one time every week. The principle analysis was applied to deal with the original data of 4 grades obtained in 4 different days. Figure 4 showed that the samples belonging to the same grade did not gather together properly and but had great dispersion; while for the samples of different grades, there was obvious trend of intersection and serious overlapping, which would not be helpful to the discrimination. Besides, contribution ratios of first principle component and second principle component were only 50% and 25%, showing that the cohesiveness of original data was worse and not suitable for the tea grade discrimination.

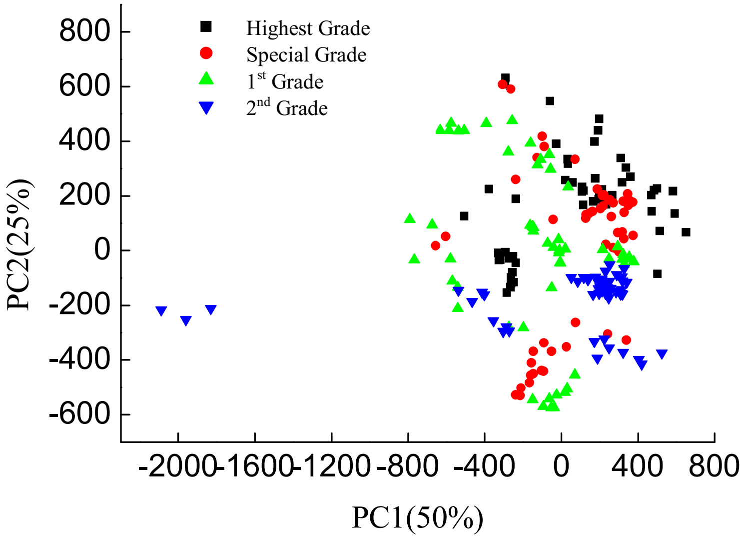

Fig. 4. PCA Scores of Original Date from Four Grading

In order to solve the problem of error at different times, the reference sample was introduced during the data collecting. 1st Grade tea soup was chosen as the reference sample whose preparation method and acquisition method kept the same as other experimental samples. The reference sample was arranged in each sequence experiment, which means for each experiments, there was always a corresponding reference existing. It was necessary to conduct the signal drifting treatment to remove the drifting by utilize the value of reference sample gained in each measurement before modeling. The method was shown below: firstly, fixed the response value of reference sample on certain sequence as the reference value(REF), then figured out different value (∆i=refi-REF) of response value of other sequence reference samples with REF, and then subtracted the value of the corresponding reference samples from the original data of the samples. In this way, the drift could be minimized as largely as possible and data by corrected in this way could be used to build the reliable prediction mode.

Figure 5 was the PCA scoring figure of the samples calibrated by the correction. When compared with the results before the correction, data after correction showed that the cohesiveness of the same grade of tea samples was enhanced and the discreteness between different grade samples was improved, which brought more significant difference between different grads of samples. By analyzing the value range of the horizontal and vertical coordinates of the figure before and after calibration, we could find that after the correction, the values of PC 1 and PC 2 were reduced. But the contribution rates of the first principal component and the second principal component were improved largely, whose contribution rates were 78% and 12%respectively, and the cumulative contribution rate reached 90%, representing the main information of original data. Meanwhile, the cumulative contribution rate of the first four principal components was 97%, representing that these PCs containing 97% of sensor information. According to the main principle of PCA, it could be known that the first four principal components represented the structure characteristics of electronic tongue data of the samples, which could be helpful to reduce the data dimensionality and simplify the data. The above proved that the problem of system drifting for electronic tongue at different times was solved, showing good test reproducibility of the electronic tongue.

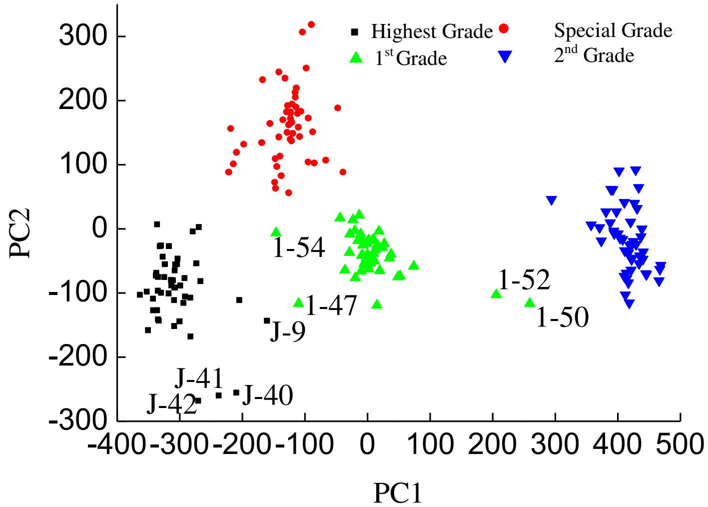

Fig. 5. PCA Scores Under Sensors Modification

3.3 Rejection for Abnormal Sample Point of Tea

The sample set used for pattern recognition was required to be provided with representativeness and correctness of original data, and without anomalism. The existence of the outlier would affect, even change the distribution trend of original data leading to the effect of the prediction model accuracy. From figure 5 it could be found there outliers in highest grade and 1st grade marked by the separation from the cluster. By analyzing (Figure 6) the samples of four grades with Mahalanobis distance in combination with residual errors, No. 54 sample was doubted as the outlier considering its largest Mahalanobis value, meaning that it was apart from the set. Meanwhile, the residual value of No. 50 was largest, which meant the prediction for this sample was inaccurate. In other words, the model cannot be able to explain the sample properly, so it should be treated as abnormal sample, too. However, there was not similar phenomenon for 2nd Grade and Special Grade samples.

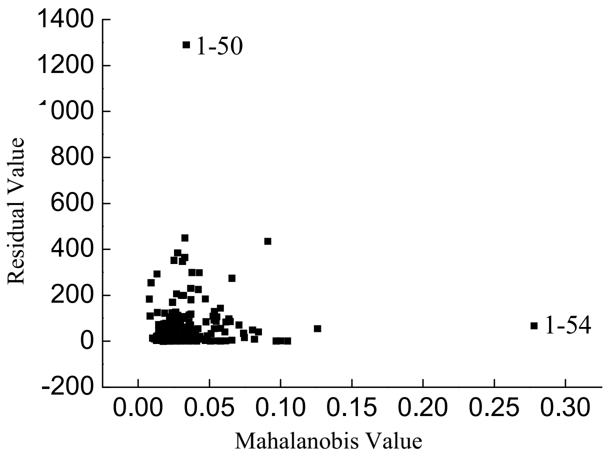

Fig. 6. PCA Scores (a) and Mahalanobis Residual (b) before Outlier Elimination

In order to comprehensively analyze the abnormal sample point, the principal component analysis and calculation of Mahalanobis distance value for the 1st Grade samples and highest Grade samples were conducted separately. From Figure 7, it could be found that the samples of No.47, No.50 and No. 54 were the three maximum Mahalanobis distance samples in 1st Grade, being considered to be the abnormal sample points. However, No. 52 sample not only detached from the samples set in the score figure but also had larger Mahalanobis distance value and residual value, also being of abnormal sample point. From Figure 8, it could be found that the Mahalanobis distance of No.9 sample in the competitive product sample set was maximum, being of abnormal sample point, but the samples of No.40, No.41 and No.42 were far from the samples set in the scoring chart, also being judged to be of abnormal sample points. So the quantity of abnormal sample points to be removed was 8 in total, including four primary samples respectively: 1-47, 1-50, 1-52 and 1-54; and four competitive product samples respectively: J-9, J-40, J-41 and J-42.

Table 2 was the number of the samples before and after outlier elimination for each grade, and the final division of the sample set. The quantity of sample used for discriminating the grade of tea were 209 in total, among which, two-thirds were randomly chosen to be served as calibration set samples to establish the qualitative classification model so as to enable the modified samples to not only have good representativeness but also broaden the prediction range of the model to strengthen the adaptive capacity of the model. The rest one-thirds were used for prediction set samples to inspect the veracity and reliability of the established model.

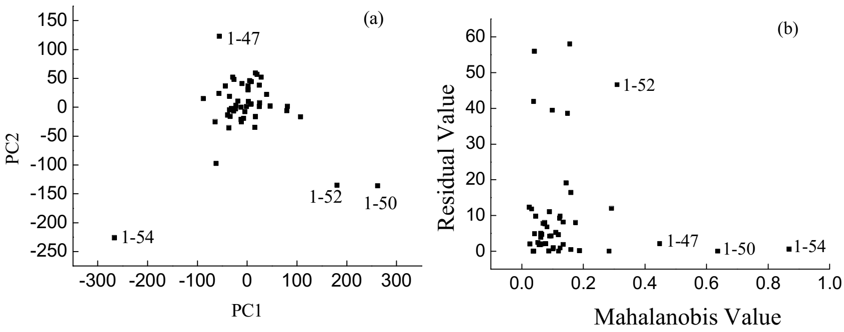

Fig. 7. PCA Score (a) and Mahalanobis Residual (b) Before Outlier Elimination of Grade-I Sample

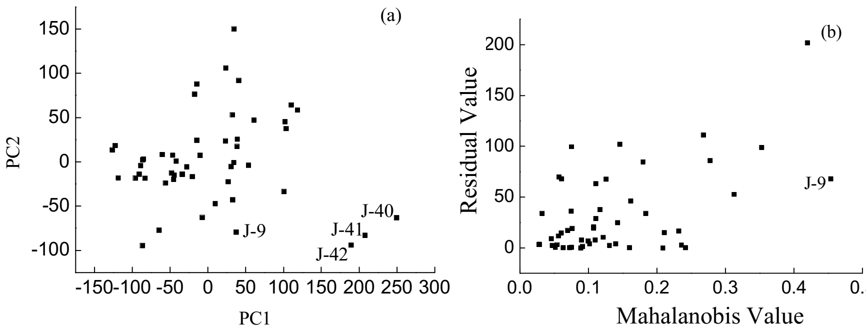

Fig. 8. PCA Score (a) and Mahalanobis Residual (b) before Outlier Elimination of Competitive Product Grade Samples

Table 2. Quantity for Different Sample Sets of Graded Model

Name of specimen clustering	After Rejection			Before Rejection
	Sample quantity of training set	Sample quantity of prediction set	Total sample amount	
Highest Grade	33	17	50	54
Special Grade	36	18	54	54
1st Grade	34	16	50	54
2nd Grade	36	19	55	55
Quantity of all graded sample	139	70	209	217

3.4 Grading Model Establishment of Xihu Longjing Tea Based on Electronic Tongue

Although the individual sensor was provided with high sensitivity and selectivity for identifying the different grades of tea samples, but the component of tea soup was complex, whose information were inevitably mixed together. Through choosing repetitive signal of electronic tongue, after calibrating signal drift at different time and eliminating outlier tea samples, the performance of PCA for extracting the independent information from large numbers of data was improved effectively. According to Figure 9, the score chart of PC 1 and PC 2 could make the distinction between different grade of samples more obvious (the cumulative contribution rate of the first two principal components had been up to 90%), among which the performance of PC 1 was the greatest, being able to distinguish different grade samples obviously. With the decrease of grade level, the score value in PC 1 increased gradually; while the variance of score value between principal component 3 and principal component 4 was very small (the cumulative contribution rate of principal component 3 and principal component 4 was 6%).

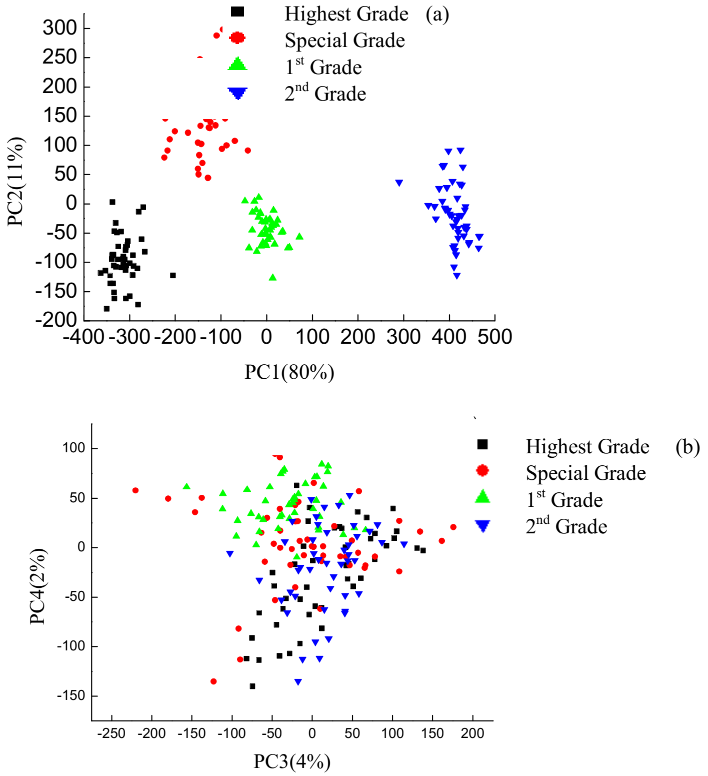

Fig. 9. PCA Scores between Different PCs after Outlier Elimination: PC1-PC2, (b) PC3 -PC4

SIMCA discrimination model was built based on the basis of PCA analysis, where the choice of principal component numbers was very important to SIMCA modeling. The best principal component number of models for the different grades was determined by cross-validation. Figure 10 was the relationship between the sum of square of prediction residual of different grade samples model and the number of principal components. In the case of a little change of SIMCA value, selecting the relatively less principal components, and finally the number of principal components of SIMCA modeling for Highest Grade, Special Grade, 1st Grade and 2nd Grade were selected as 2, 1, 2, 3, respectively. The SIMCA tea grade discrimination model was established according to the selected number of principal components of each grade of samples. Table 3 was the final model result. As what was found in the table, the electronic tongue sensor could distinguish the four grades of Xihu Longjing Tea effectively, not only for cross-validation for 139 modeling samples(33 Pcs. of Highest Grade, 36 Pcs. of Special Grade, 34 Pcs. of 1st Grade, 36 Pcs. of 2nd Grade) reaching 100%, but also for 70 prediction samples with the unknown grade level (17 Pcs. of Highest Grade, 18 Pcs. of Special Grade, 16 Pcs. of 1st Grade, 19 Pcs. of 2nd Grad) reaching 100%, and the models were provided with good adaptability and robustness. So the samples of same grade gather into a group severally, and the samples of different groups didn't overlap each other to reflect the good susceptibility of electronic tongue.

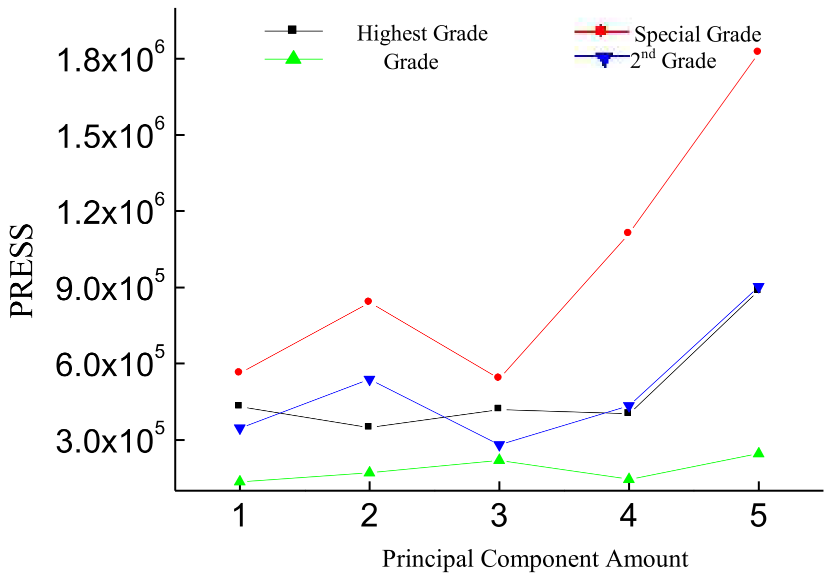

Fig. 10. Relationship between PRESS and Number of PCs under Grading Model

Table 3. Results of SIMCA Modeling for Four Grades of Samples

Model Type	**Number of PC**	Integral Identification Ratio（%）	
		Correction Set	Prediction Set
Highest Grade, Special Grade, 1st Grade, 2nd Grade	2,1, 1, 3	100	100

4 Conclusions

The study chose the graph of samples in the seventh repetition to stand for the intelligent taste fingerprint of the sample since in this case the stability and authenticity of the tea soup could be guaranteed effectively. The corresponding reference on each sequence was adopted to be served as the basis of calibration to solve the problem of system error generated by linear drifting which guaranteed the stability of the graph signal, making Mahalanobis distance and residual method could be run quickly, factually and accurately, eliminating the abnormal samples in tea sample set. Finally, we used SIMCA modeling with of the selected number of principal components, and the accuracy of discrimination for Xihu Longjing Tea reached 100%. In conclusion, the study focusing on three key points of electronic tongue: repeatability, reproducibility and representative. By handling these three points we could obtain signal graph with high signal to noise ratio, and the authenticity of the characteristics of the samples. The study developed the process of confirming the repeatability, correcting graph drift system error, eliminating the outlier samples and building the discrimination model with technical specifications and procedure. All these will provide the theory basis for application of electronic tongue in other fields.

Acknowledgment. This study had been sponsored by Beijing-sponsored programs for excellent talents training (No. 2012D009999000001), National High Technology Research and Development Plan (863 Plan) for the Project of No. 2011AA1008047.

References

1. Rahim, A.A., Nofrizal, S., Saad, B.: Rapid tea catechins and caffeine determination by HPLC using microwave-assisted extraction and silica monolithic column. Food Chemistry 147, 262–268 (2014)
2. Shi, B., Zhao, L., Wang, H., Yin, J.: Application of intelligent sensory technique in evaluation of tea quality. Food Science 30(19), 351–355 (2009)
3. Jiang, S., Chen, Q., Hu, X., Yang, Y., Ni, Y.: Application of electronic tongue on black tea beverage discrimination. Transactions of the CSAE 25(11), 345–348 (2009)
4. Patrik, I., Susanne, H., Nils-Eriker, H., et al.: Discrimination of tea by means of a voltammetric electronic tongue and different applied waveforms. Sensors and Actuators B 76, 449–454 (2000)

5. Lvova, L., Legin, A., Vlasov, Y., et al.: Multicomponent analysis of Korean green tea by means of disposable all-solid-state potentiometric electronic tongue microsystem. Sensors and Actuators B 95, 391–399 (2003)
6. Chen, Q., Jiang, S., Wang, X.: Discrimination of tea's quality level based on electronic tongue and pattern recognition. Food and Machinery 24(1), 124–126 (2008)
7. He, W., Hu, X., Zhao, L., Liao, X., Zhang, Y., Wu, J.: Application of electronic tongue in the Pu-er tea quality grade analysis. Science and Technology of Food Industry 30(11), 125–127 (2009)
8. Wu, J., Liu, J., Fu, M., Li, G.: Classification of Chinese green tea by a voltammetric electronic tongue. Chinese Journal of Sensors and Actuators 19(4), 963–965 (2006)
9. Liu, S., Tan, J., Lin, Z., Lu, C., Wang, X.: Application of electronic tongue in green tea sensory estimation and grade evaluation. China Tea (5), 19–20 (2014)
10. Wang, X., Chen, Q.: Discrimination on grades of roasted green tea using electronic tongue. Journal of Anhui Agricultural Science 35(28), 8872–8873 (2007)
11. Roy, R.B., Tudu, B., Shaw, L., Jana, A., Bhattacharyya, N., Bandyopadhyay, R.: Instrumental testing of tea by combining the responses of electronic nose and tongue. Journal of Food Engineering (110), 356–363 (2012)
12. Xu, Y., Chen, W., Yin, B., Huang, L.: Green tea store and preservation technology. Agriculture Machinery Technology Extension 9, 29 (2004)
13. Xue, D., Shi, B., Zhao, L., Yin, J.: Classification of tea's grade level based on electronic tongue. Food Science and Technology 35(12), 278–281 (2010)
14. Bolin, S., Lei, Z., Wen, L., Houyin, W., Dazhou, Z., Ingyuan, Y.J.: Outlier sample analysis on near infrared spectroscopy determination for apple interior quality. Transactions of the CSAM 41(2), 132–137 (2010)
15. Bolin, S., Lei, Z., Ruicong, Z., Xingjun, X., Dazhou, Z.: Quality recognition of Xihu-Longjing tea based on intelligent olfactory. Transactions of CSAM 43(12), 130–135 (2012)
16. Shi, B., Zhao, L., Zhi, R., Xi, X., Zhu, D.: Aroma quality discrimination of Xihu-Longjing tea by electronic nose. Transactions of the CSAE 27(Supp. 2), 302–306 (2011)