

Article

A Framework for the Multi-Level Fusion of Electronic Nose and Electronic Tongue for Tea Quality Assessment

Ruicong Zhi ^{1,2,*}, Lei Zhao ³ and Dezheng Zhang ^{1,2}

¹ School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China; zdzchina@126.com

² Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing 100083, China

³ China National Institute of Standardization, Beijing 100191, China; zhaolei@cnis.gov.cn

* Correspondence: zhirc_research@126.com or zhirc@ustb.edu.cn; Tel.: +86-10-6233-4547

Academic Editor: Manel del Valle

Received: 2 March 2017; Accepted: 20 April 2017; Published: 3 May 2017

Abstract: Electronic nose (E-nose) and electronic tongue (E-tongue) can mimic the sensory perception of human smell and taste, and they are widely applied in tea quality evaluation by utilizing the fingerprints of response signals representing the overall information of tea samples. The intrinsic part of human perception is the fusion of sensors, as more information is provided comparing to the information from a single sensory organ. In this study, a framework for a multi-level fusion strategy of electronic nose and electronic tongue was proposed to enhance the tea quality prediction accuracies, by simultaneously modeling feature fusion and decision fusion. The procedure included feature-level fusion (fuse the time-domain based feature and frequency-domain based feature) and decision-level fusion (D-S evidence to combine the classification results from multiple classifiers). The experiments were conducted on tea samples collected from various tea providers with four grades. The large quantity made the quality assessment task very difficult, and the experimental results showed much better classification ability for the multi-level fusion system. The proposed algorithm could better represent the overall characteristics of tea samples for both odor and taste.

Keywords: multi-level fusion; feature fusion; decision fusion; electronic nose; electronic tongue; tea quality assessment

1. Introduction

Tea quality assessment is crucial for both producers and consumers, and it is a very challenging task due to the presence of innumerable compounds and their diverse contribution to tea quality. The evaluation of tea quality is usually carried out through human sensory analysis, which can provide direct and integrated measurements of various attributes. The gradation of tea is yielded on the basis of experienced tea tasters' scores, and tea tasters assign the perceived intensities of appearance, aroma, and taste captured by their sense organs [1]. However, human sensory evaluation is subjective, as even the same person may provide different evaluation information in different experiments, and sensory panel are inconsistent due to the individual variability. Moreover, there are various affective factors (e.g., physical or mental states) which make the human sensory evaluation inaccurate [2].

With the increased expectation of high quality and a large quantity of the product, there are many requirements for objective measurements in a fast, accurate, and cost-effective manner [3]. Conventional flavor analysis techniques are high cost and not suitable for online quality control, such as gas chromatography (GC), high-performance liquid chromatography (HPLC), plasma atomic emission spectrometry, and capillary electrophoresis [4].

Electronic intelligent systems (e.g., E-nose and E-tongue) have received considerable attention during the last two decades, since numerous applications have increased in the food industry [3,5]. The E-nose and E-tongue can mimic the sensory perception of human smell and taste, and they are widely applied in tea quality evaluation by utilizing fingerprints of response signals representing the overall information of tea samples. E-tongue can detect the taste of samples, whereas e-nose can identify the odor. Both of them are based on the principle of identifying odor/taste by extracting an overall signature from the comprehensive mixture of the compounds. Usually, an electronic intelligent system consists of a sensor array for chemical detection, along with advanced pattern recognition systems so that the sensor signal data can be automatically and reliably processed.

E-noses and E-tongues have been widely used in many food quality assessment applications, such as fruits (grapes, pears, apples, mangos, pineapples, and peaches), beverages (juice, coffee, tea, wine, and milk), meat, fish, and rice, etc. [3,5–7]. There are a great number of studies focusing on tea quality assessment by electronic intelligent instruments. Both E-nose and E-tongue are successfully applied to provide fast and reliable results for tea grade identification, storage time and fermentation processes, where the odor of tea or the taste of tea is considered as the sole attribute [8–19].

The intrinsic part of human perception is the fusion of sensors, as multiple sensory organs provide more information which helps to make a better decision than a single sensory pipeline. The electronic intelligent system is carried out in a similar way as human perception, with diverse sensors generating different signature phenomena to fully exploit the characteristics of test samples. Data fusion aims to combine information acquired from multiple sources through different strategies to achieve a better description and to enhance the probability of an accurate classification. Usually, fusion can be categorized into data fusion (low-level fusion), feature fusion (intermediate-level fusion), and decision fusion (high-level fusion). The data fusion strategy employs raw data level fusion by combining the raw sensor response signals into a single signal. Feature fusion concatenates the features extracted from the signal values of the sensors through a variety of feature extraction and selection methods. Decision fusion presents the output of multiple classifiers to achieve a final prediction, and each classifier is trained for each signal source. The decision level fusion combines decisions from multiple sensory pipelines, and this is similar to the fusion mechanism in the human brain [5].

A number of researches reported that a system combining an E-nose and E-tongue improved the performance of an individual system for quality assessment [20–29]. Most of the researches focused on data fusion and feature fusion, by combining E-nose and E-tongue raw sensor signals or features extracted from raw signal data, and a few researches reported on the decision fusion of multiple electronic intelligent instruments [30–33].

In this study, a framework for a multi-level fusion strategy of E-nose and E-tongue was proposed to promote the tea quality prediction accuracies by simultaneously modeling feature fusion and decision fusion, and the procedure is illustrated in Figure 1. Two different features of time-domain based and frequency-domain based representations were extracted from the E-nose and E-tongue sensor responses to better represent individual sensor signal characteristics. The merged features were analyzed by a nonlinear dimensionality reduction algorithm for feature selection. Finally, the classification results obtained by the K-nearest neighbor classifier from the E-nose and E-tongue were fused by D-S evidence, which was an effective decision fusion algorithm.

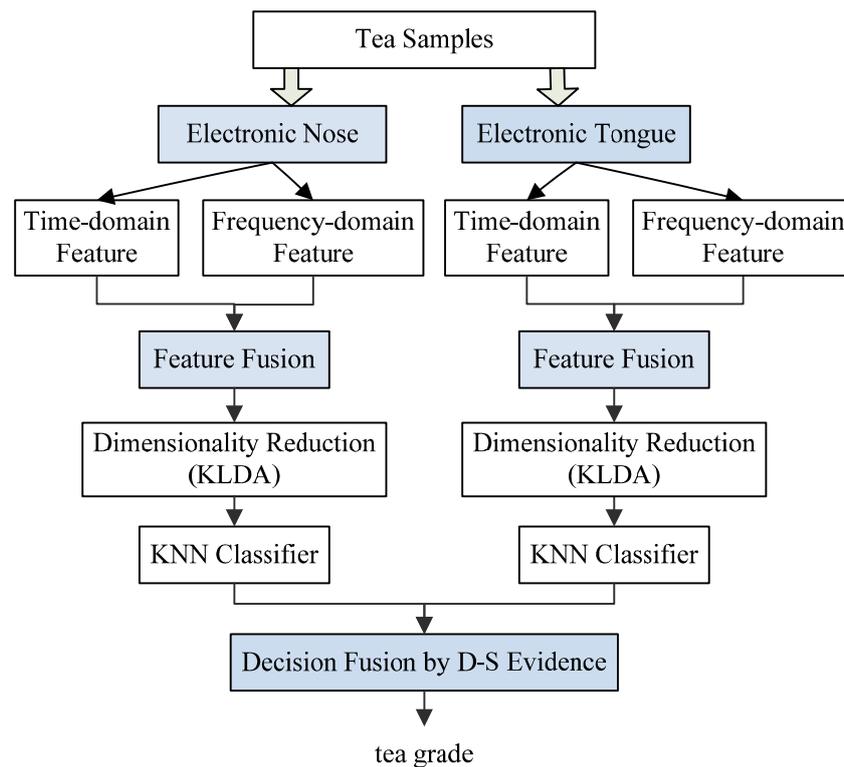


Figure 1. Illustration of the strategy flow of multi-level fusion tea quality assessment.

2. Materials and Methods

2.1. Sample Collection

All of the tea samples belong to a category of green tea, which is called “Longjing tea”, with four different grade levels (grade 1/T, grade 2/Y, grade 3/E, and grade 4/S) selected from seven tea producers (which are denoted by MC, MG, ML, MS, MX, MY, and QD) in the Xihu producing area, Hangzhou, China. The quality of the tea samples was evaluated by national certified tea experts according to the national standard “GB/T 23776-2009 Methodology of sensory evaluation of tea”. Thirty-five tea samples of each grade and each company were used, and there were 980 samples in total. The tea samples were individually vacuum packed with aluminum foil materials, and stored in the refrigerator at $-4\text{ }^{\circ}\text{C}$ [34].

2.2. Electronic Nose Measurement

An E-nose (Fox 4000, Alpha M.O.S. Co., Toulouse, France) was utilized to acquire the odor fingerprints of the tea samples. The sensor array consists of 18 metallic oxide sensors. The tea sample preparation procedure is as follows: In the preprocessing stage, each tea sample was weighted for 1 g, and mixed with 5 mL ultrapure water at room temperature in a 20 mL headspace bottle. The headspace bottle was arranged in an automatic sampling device after sealing. Cross testing was conducted for different grades of tea samples through a cyclic cross sequence, i.e., all four grades of tea samples were tested in the order of “T, Y, E, S; T, Y, E, S; T, Y, E, S; ...”. Therefore, the tea samples with various qualities could be tested alternately, thus avoiding the adaptability of sensors to certain type of tea samples. In the experiments, the headspace bottle was first sent to the preheating area for 900 s, with an oscillator rotation rate of 500 rpm/min and temperature of $60\text{ }^{\circ}\text{C}$. Then, 2 mL gas was injected into the sensory array room with a speed of 2 mL/s. Before the tea sample detection, pure nitrogen was pumped into the sensor array chamber with a speed of 2 mL/s for 300 s, to ensure that the former gas molecule was thoroughly cleaned out [35,36]. The tea sample reaction time was 120 s, and the

responses of the sensors were recorded every 0.5 s. Therefore, there were 241 values for each sensor of the electronic nose.

2.3. Electronic Tongue Measurement

An E-tongue (α -ASTREE, Alpha M.O.S. Co., Toulouse, France) with an array of seven electrodes was utilized in this study. The tea samples were prepared by boiling 150 mL of ultrapure water with every 1 g tea and covered. Then, the beaker was put in a boiling water bath (100 °C) and shaken every 10 min. After 45 min, the tea infusion was filtered and cooled down for measurement after 2 h [2]. The electronic tongue is not as stable as the electronic nose, and some preprocessing is needed to enhance the stability and reliability of the tea sample measurement. Such preprocessing includes self-testing, activation, training, calibration, and diagnostics. The procedure is as follows: The connection of the hardware is automatically tested by the equipment software. The sensors are cleaned by deionized water for 10–20 s, and activated in another cup of deionized water for 30 min to enrich H⁺ on the coating film of the sensor. Then, the sensors are trained using HCl (0.01 mol/L, training solution) and ultrapure water (cleaning solution). First, the sensors are cleaned in ultrapure water for 10 s, and then the sensors are moved to the HCl solution for 300 s. The same procedure is repeated for three times which is controlled by the software. Calibration is consequently conducted to check whether the coating film of the sensor is balanced after training. The sensors are cleaned in ultrapure water (10 s) and reacted in 0.01 mol/L HCl (120 s), and the operation is repeated three times. Moreover, the instrument is diagnosed every 20 days in our test to guarantee the sensitivity and effectiveness of the sensors. The procedure is controlled and implemented by the software. Each of the electrodes produced a signal curve, and each signal consisted of 241 points as the signal value was acquired every 0.5 s in the period of 120 s. Therefore, the data formed a matrix of 7 × 241.

2.4. Multi-Level Fusion System

2.4.1. Feature Extraction and Fusion

The response signals collected from the E-nose and E-tongue are sequence signals changing over time. It is not feasible to deal with the original signal directly as it is time consuming and there is usually a noisy signal mixed in with the sensor signal sequences. Therefore, feature extraction is of great importance as it can efficiently analyze the sensor sequences and extract the features representing the characteristics of the signals.

Most researches employed time-domain based features, which could represent the inner characteristics from the intelligent sensors' response signals. In the signal processing field, more effective representations are utilized to denote the frequency-domain based features by filter processing, as they can differentiate the signals which are not well discriminated in the time domain. Therefore, both time-domain based representations and frequency-domain based representations are extracted from the E-nose and E-tongue response signals, and feature fusion is conducted to comprehensively adopt the superiority of both representations.

In this study, the time-domain based representations include the maximum value (MV) and the average value (AV) of the original sensor responses of the E-nose and E-tongue, respectively. The maximum value of the i -th sensor is defined as:

$$MV_i = \max|x_{i,1}, x_{i,2}, \dots, x_{i,241}| \quad \left(\begin{array}{l} i = 1, 2, \dots, 18 \text{ for E-nose} \\ i = 1, 2, \dots, 7 \text{ for E-tongue} \end{array} \right) \quad (1)$$

The average value of the i -th sensor is defined as:

$$AV_i = \frac{x_{i,1} + x_{i,2} + \dots + x_{i,241}}{241} \begin{pmatrix} i = 1, 2, \dots, 18 \text{ for E-nose} \\ i = 1, 2, \dots, 7 \text{ for E-tongue} \end{pmatrix} \quad (2)$$

where i is the sensor label and $x_{i,1}, x_{i,2}, \dots, x_{i,241}$ are the absolute signal of the i -th sensor for the E-nose and E-tongue, respectively.

The frequency-domain based representation is extracted by wavelet packet analysis, and the maximum energy (ME) and average energy (AE) of the wavelet packet coefficients are explored to denote the mainstream traits and the overall level of the sensor signals. The maximum energy (ME) of the i -th sensor is defined as:

$$ME_i = \max \left| E_{30}^i, E_{31}^i, \dots, E_{37}^i \right| \begin{pmatrix} i = 1, 2, \dots, 18 \text{ for E-nose} \\ i = 1, 2, \dots, 7 \text{ for E-tongue} \end{pmatrix} \quad (3)$$

The average energy (AE) of the i -th sensor is defined as:

$$AE_i = \frac{E_{30}^i, E_{31}^i, \dots, E_{37}^i}{8} \begin{pmatrix} i = 1, 2, \dots, 18 \text{ for E-nose} \\ i = 1, 2, \dots, 7 \text{ for E-tongue} \end{pmatrix} \quad (4)$$

where $E_{30}^i, E_{31}^i, \dots, E_{37}^i$ are the energies of the i -th sensor at each frequency band, and $E_{3j} = \sum_{k=1}^m |C_{3jk}|^2$ ($j = 0, 1, \dots, 7$). m is the number of coefficients and C_{3jk} is the k -th coefficient of the j -th frequency band corresponding to three scale orthogonal wavelet decomposition.

The feature-level fusion is conducted by connecting the time-domain based features and frequency-based features in series, i.e., the four different features of each sensor are individually arrayed as (MV_i, AV_i, ME_i, AE_i) , and the fused features of each sensor are concatenated in order of the E-nose and E-tongue, respectively [36]. The fused feature is standardized for further analysis.

2.4.2. Nonlinear Subspace Embedding

Linear separation in original data space is rare in the practical application of quality classification due to the correlation and redundancy among multiple electronic sensors. Therefore, the fused features concatenating time-domain representation and frequency-domain representation extracted from the original sensor signals are embedded in a high dimensional subspace which is linearly separable for the samples. Kernel-based algorithms widely concern nonlinear embedding methods for dimensionality reduction. The effectiveness of kernel-based dimensionality reduction methods was evaluated in [19]. KLDA (Kernel Linear Discriminant Analysis) is utilized in this study to reduce the fused feature dimension [37].

The fused features for sample k ($k = 1, 2, \dots, N$) are denoted by $\mathbf{x}_k = [MV_i, AV_i, ME_i, AE_i]$ for the E-nose ($i = 1, 2, \dots, 18$) and E-tongue ($i = 1, 2, \dots, 7$). In the following, we use a single symbol $\mathbf{X} = [\mathbf{x}_1; \mathbf{x}_2; \dots; \mathbf{x}_N]$ to represent the E-nose and E-tongue data matrix for simplicity. The data matrix \mathbf{X} is implicitly mapped to a high dimensional feature space F with a nonlinear transformation Φ , i.e., $F = \{\Phi(\mathbf{X}) : \mathbf{X} \in R^N\}$, where N is the sample number. The objective of the KLDA algorithm is to maximize the ratio of the between-class scatter and the within-class scatter in F , where the within-class scatter is defined as $\mathbf{S}_w^\Phi = \sum_{i=1}^c \sum_{j=1}^{n_i} (\Phi(\mathbf{x}_{ij}) - \mathbf{m}_i)(\Phi(\mathbf{x}_{ij}) - \mathbf{m}_i)^T$, where c is the number of classes (four grades), n_i is the number of samples in class i , and \mathbf{m}_i is the mean of the i -th class sample. The between-class scatter is defined as $\mathbf{S}_b^\Phi = \sum_{i=1}^c n_i(\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T$, where \mathbf{m} is the mean of all the tea samples.

2.4.3. Classification

Classification is the grade identification process on the basis of fused and selected features. The K -nearest neighbors (KNN) classifier is a non-parametric method which is among the simplest of the machine learning algorithms employed for classification. The output of the classifier is the class label for the new unclassified sample by a majority vote of its K -nearest neighbors in the training set, and no explicit training step is required.

2.4.4. Decision Fusion Based on D-S Evidence

D-S evidence has been widely applied to artificial intelligence and multi-sensor fusion [38]. The D-S evidence theory is a more general and flexible method than traditional probability approaches, as it supports both the imprecision and uncertainty representation [39]. The basic principle of the D-S evidence is stated in detail, as follows:

Let $\Theta = \{\theta_1, \theta_2, \dots, \theta_c\}$ be the set of $\theta_c (\theta_c \in \Theta)$ corresponding to c identifiable classes, then Θ is the space of the hypotheses called a frame of discernment.

A key issue of the D-S evidence theory is the definition of basic probability assignment. The basic probability assignment function m is defined on 2^Θ as $m : 2^\Theta \rightarrow [0, 1]$ for every element A of 2^Θ . The probability assignment value $m(A)$ satisfies the following properties:

$$\begin{aligned} m(\phi) &= 0 \\ \sum_{A \in 2^\Theta} m(A) &= 1 \end{aligned} \quad (5)$$

For any $A \in 2^\Theta$, the quantity $m(A)$ represents a measure of belief that is exactly committed to A . ϕ is the empty set. The element A of 2^Θ with $m(A)$ being greater than zero is called the focal element of m .

The other key issue of the D-S evidence theory is the aggregate multiple evidence from different sources defined on the same frame of discernment by means of the basic probability assignment function.

Two bodies of evidence, m_1 and m_2 , with focal elements A_1, A_2, \dots, A_i and B_1, B_2, \dots, B_j , respectively, can be merged to obtain a new basic probability assignment function m by a combination rule. The D-S evidence combination rule is defined as:

$$m(A) = \frac{\sum_{A_i \cap B_j} m_1(A_i) m_2(B_j)}{1 - Q} \quad A \neq \phi \quad (6)$$

where $Q = \sum_{A_i \cap B_j = \phi} m_1(A_i) m_2(B_j)$ and $Q < 1$.

One of the main difficulties in the D-S evidence is how to initialize the basic probability assignment function as well as possible. There is no general answer to the key problem of the basic probability assignment function definition [40,41]. Generally, the probability values are determined artificially depending on the specific applications.

2.4.5. Decision Fusion for the E-Nose and E-Tongue

The grade identification is conducted by D-S evidence decision fusion for the E-nose and E-tongue, and the KNN classifier is utilized for the E-nose and E-tongue, respectively.

Let the four grades of the tea samples be denoted by T, Y, E, and S. Therefore, the identifiable set is defined as $\Theta = \{T, Y, E, S\}$. In our experiment, the basic probability assignment function is defined by the output of the classifier. The results of the KNN classifier are utilized to calculate the membership degree from the K -nearest neighbors of each class, that is:

$$m(i) = \frac{x_neigh_i}{K}, \quad i = T, Y, E, S \quad (7)$$

where x_neigh_i is the number of neighbors belonging to class i from the K -nearest neighbors. The parameter K is determined according to the recognition accuracies of the experiments.

Assume that $m_T^i, m_Y^i, m_E^i, m_S^i$ and $n_T^i, n_Y^i, n_E^i, n_S^i$ are the basic probability assignment values for the four grades of the i -th tea samples by the E-nose and E-tongue, respectively. The final probability values of the tea sample $M_T^i, M_Y^i, M_E^i, M_S^i$ are identified by the D-S evidence combination rule, expressed as Equation (6).

The grade of a test tea sample is classified according to the maximum membership rule by the final probability values, that is:

$$class(i) = suffix\left(\max\{M_j^i\}\right), \quad j = T, Y, E, S \quad (8)$$

The suffix of the maximum value of $M_T^i, M_Y^i, M_E^i, M_S^i$ is the grade that the test sample is assigned. The algorithms are implemented by Matlab R2015b (Mathworks, Natick, MA, USA).

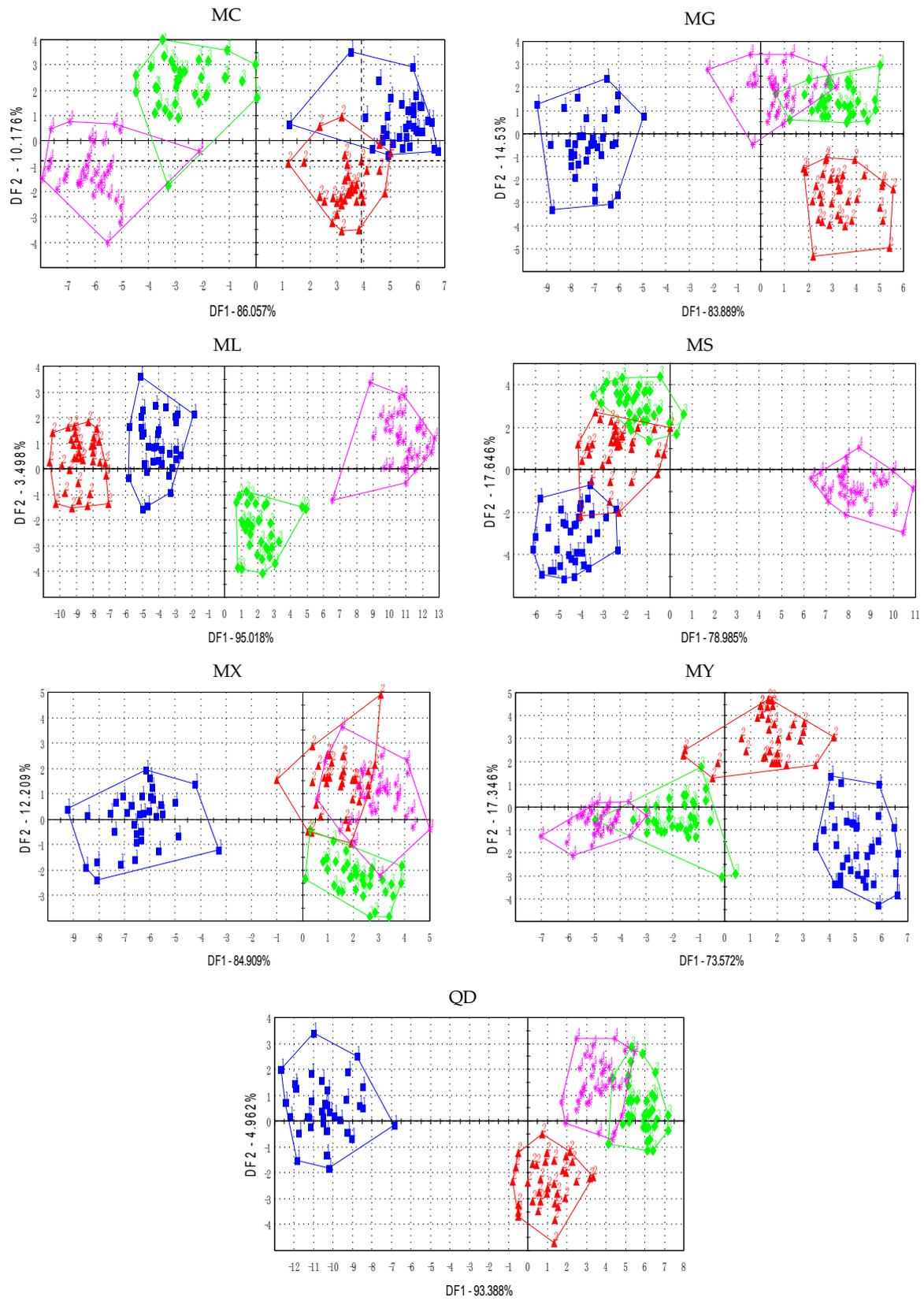
3. Results and Discussion

3.1. Feature Representation from Sensors

Both the time-domain based features (MV, AV) and the frequency-domain based features (ME, AE) were calculated from the sequence signal of the E-nose and E-tongue to enhance the discriminative ability of the features. The feature fusion was individually conducted for the E-nose and E-tongue. The fused features of all the 18 sensors of the E-nose were concatenated in order, and obtained a feature vector of 72 dimensions. Similarly, a vector of 28 dimensions was obtained by fusing the features of all the seven sensors of the electronic tongue.

The discriminant ability of the features was evaluated by cluster scatters performed by linear discriminant analysis. The cluster trends of the tea samples were visualized by a 2-D scatter plot. Figure 2a shows the cluster scatter of the seven tea sample providers by single features (maximum values which were most commonly used) for the electronic nose, and Figure 2b shows the cluster scatter by the fused features. Scatters for the average value, maximum energy, and average energy of the E-nose are provided in the supplementary material.

The results of the electronic nose demonstrated that the tea samples with four grades were not correctly discriminated by a single feature. The boundaries of the classes were not explicit and the samples with different tea grades overlapped. Moreover, the samples in the same class were diverse (i.e., located over a large area). This may decrease the probability of correctly assigning a certain test sample to one of the four grades. In Figure 2b, the scatter plots depict that the fused features significantly improved the performances. It could be seen that for most of the tea samples (i.e., producers of MC, MG, ML, MS, MX, and MY), the tea samples with the same class were much more compact, and the tea samples with different classes were clearly distinguished. Although a number of tea samples from tea producer QD with grade E and grade S overlapped, the samples within the same class became much closer than the scatters of a single feature. The results meet the objective of the dimensionality reduction methods which is to maximize the between-class scatter and minimize the within-class scatter.



(a)

Figure 2. Cont.

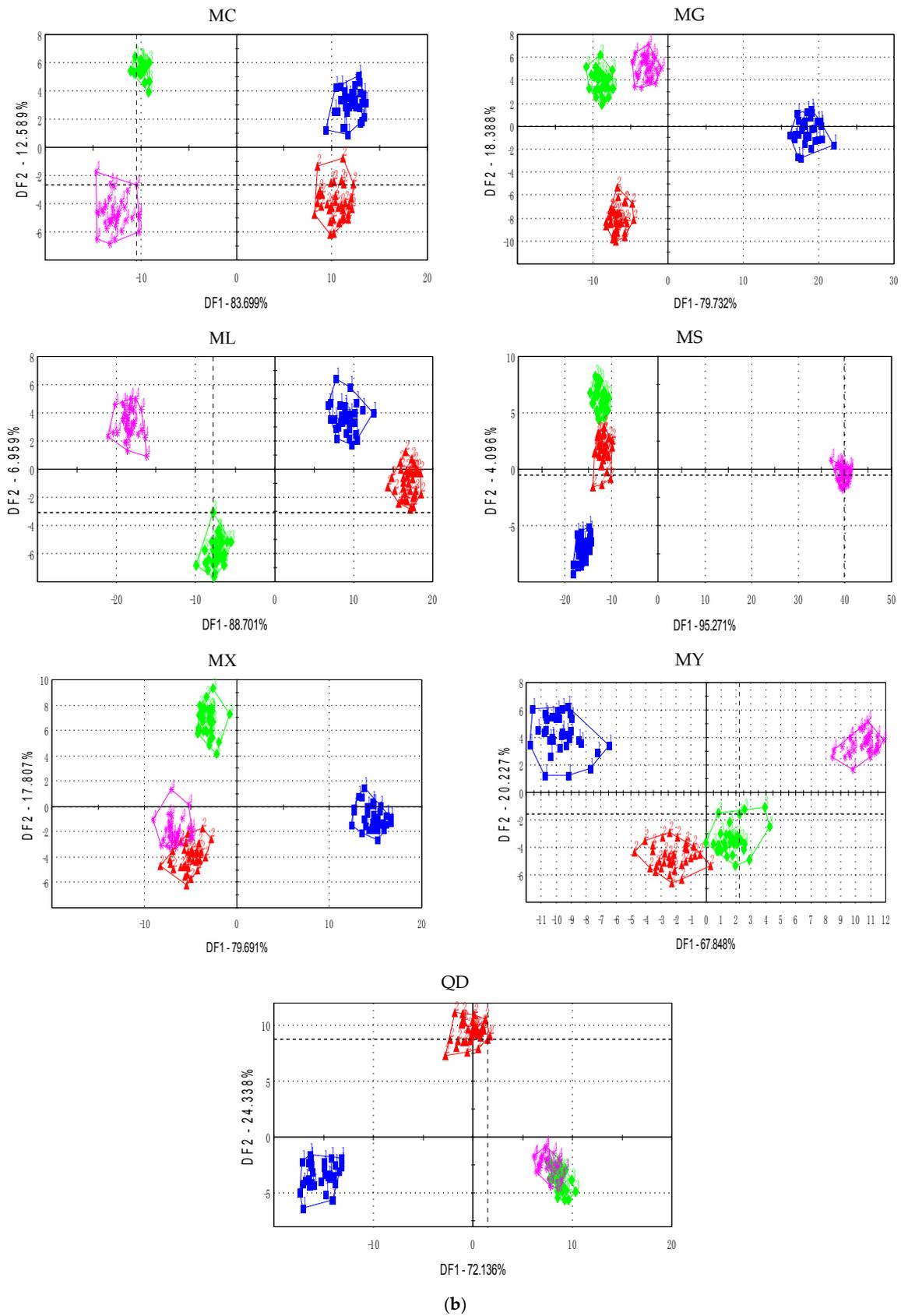


Figure 2. Score plot of the E-nose (a) maximum value (MV) (b) fused feature.

Moreover, Figure 3a illustrated the cluster scatter of the seven tea sample providers by the average values which were commonly used for the electronic tongue, and Figure 3b illustrated the cluster scatter by the fused features.

According to the results of the electronic tongue, the tea samples of the four grades significantly overlapped for single features. The fused feature improved the performance, as shown in Figure 3b. Most of the four grades of tea samples were separated correctly, and few of them were close to the class boundaries. The scatter plots of Figure 3a show that the tea samples from producers MG and MS overlapped for more than two classes, while the fused feature separated the samples clearly and the class boundaries were explicit for each class. The tea samples from producers ML and MX were a mass for a single AV feature, and Figure 3b illustrates that the classes were separated more clearly. Moreover, the samples were closely located within the same class for all the cases.

The results of both the E-nose and E-tongue demonstrate the superiority of the time-domain based and frequency-domain based feature fusion in representing comprehensive signal characteristics. The fused feature could represent the overall characteristics of the sensor signals, and it is more suitable for tea grade identification. However, the results are not clear when putting all of the tea samples from diverse companies together, even in the case of feature fusion. The tea samples collected from this study are much more than state-of-art, and the tea providers are various too. This is the real problem faced in tea gradation in China. To better identify the green tea grades with a large number of tea samples with diverse sources, advanced technologies should be introduced, and that is the purpose of proposing this study.

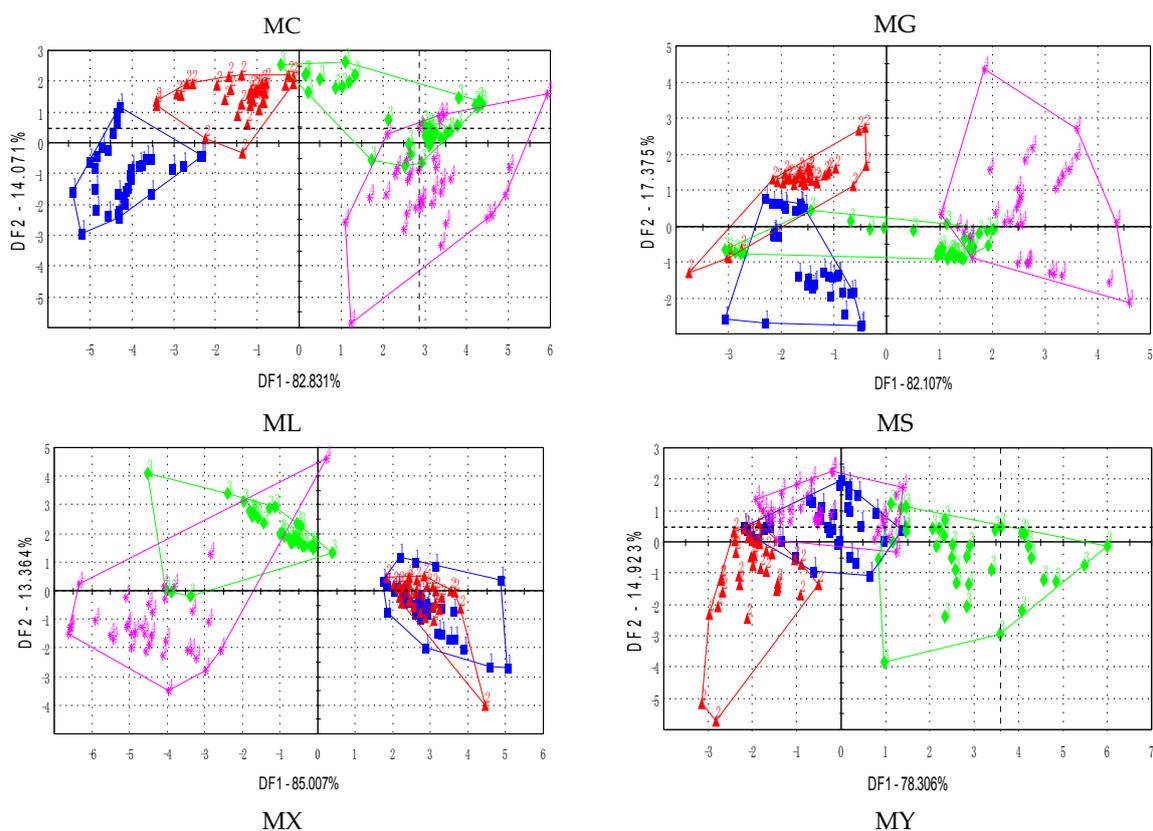


Figure 3. Cont.

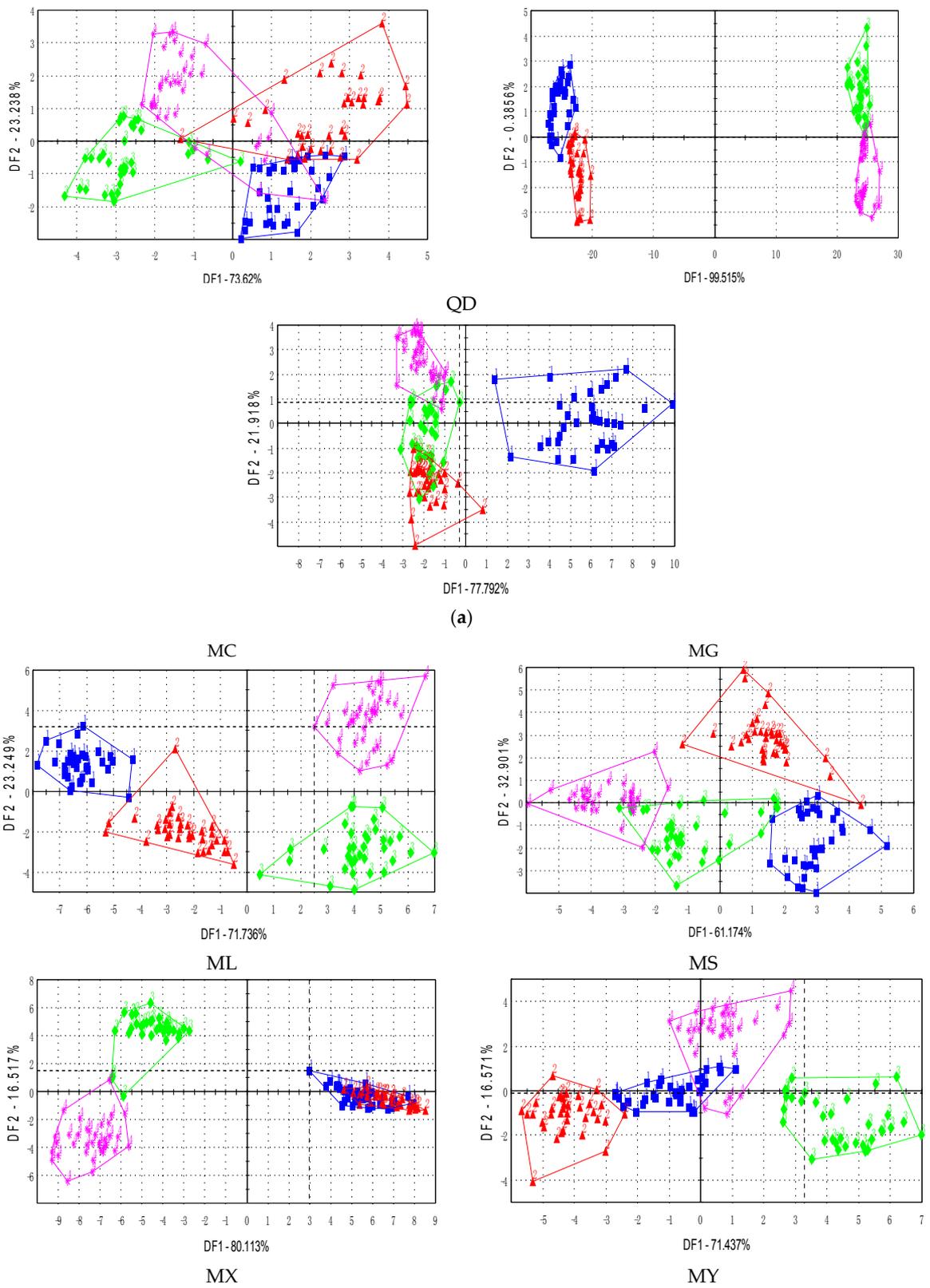


Figure 3. Cont.

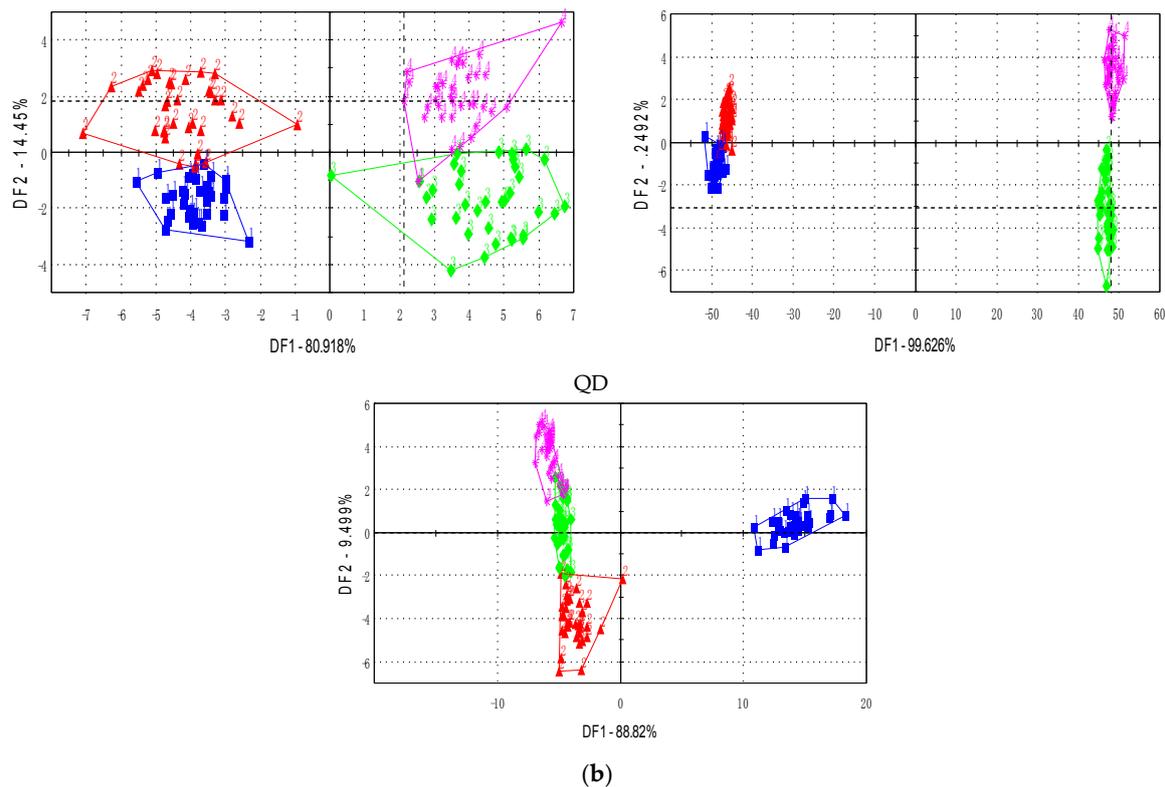


Figure 3. Score plot of the E-tongue (a) average value (AV) (b) fused feature.

3.2. Dimensionality Reduction of Fused Feature

In this section, the experimental results of the classification performed by feature fusion, the KLDA-based dimensionality reduction, and the KNN classifier were illuminated.

The tea samples were divided into two subsets, the training set was utilized to develop the model and the testing set was utilized to verify the performance of the model. The training set consisted of twenty samples of each grade from each tea provider ($20 \times 4 \times 7 = 560$ in total) by random selection, and the testing set was constructed from the remaining tea samples. The recognition accuracies versus the dimensionality according to different K values (parameter of the KNN classifier) are shown in Figure 4 for the E-nose and E-tongue, respectively. The recognition accuracies variable versus the reduced dimensions and for different K values. The top recognition accuracy of the E-nose was 71.3%, and it was 82.7% for the E-tongue. The top recognition rate of the E-nose was obtained with $Dim = 5$ and $K = 3$, while that of the E-tongue was obtained with $Dim = 4$ and $K = 3$.

Furthermore, the corresponding confusion matrices of the E-nose and E-tongue are shown in Table 1. Each column of the matrix represents the instances in a predicted class, while each row represents the instance in an actual class. It showed that the tea samples were confused among classes differently. A large number of grade E tea samples were misclassified as grade Y, and many tea samples of grade S were misclassified as grade Y and grade E for the E-nose. While many tea samples of grade T were misclassified as grade Y, grade E samples were misclassified as grade Y, and a great number of tea samples of grade S were misclassified as grade E for the E-tongue.

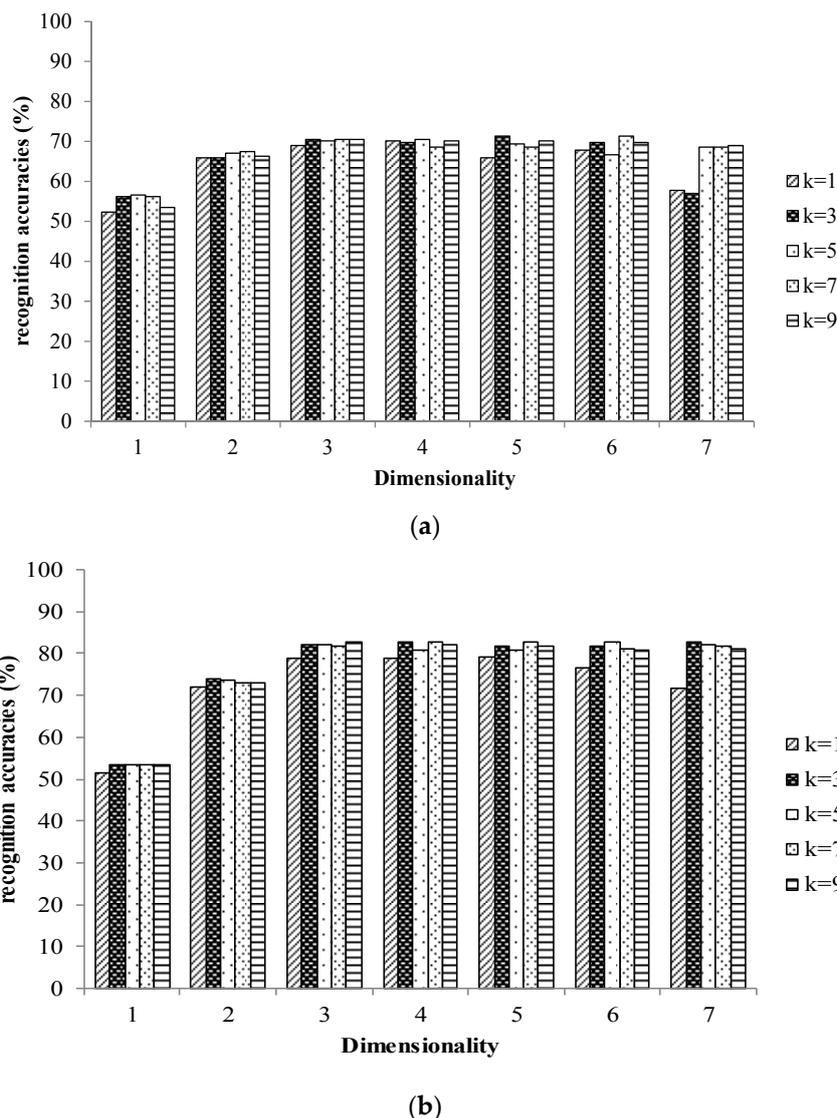


Figure 4. Recognition accuracies of the KLDA-KNN model for (a) E-nose (b) E-tongue.

Table 1. Confusion matrix (%) for KLDA-KNN classification of the E-nose, E-tongue, and decision fusion.

	E-Nose Feature				E-Tongue Feature				Decision Fusion			
	T	Y	E	S	T	Y	E	S	T	Y	E	S
T	92.0	8.0	0	0	78.7	16.0	5.3	0	93.3	2.7	4.0	0
Y	19.0	78.6	1.2	1.2	8.3	85.7	4.8	1.2	1.2	94.0	4.8	0
E	2.5	26.6	58.2	12.7	1.3	8.9	82.3	7.6	1.3	3.8	86.1	8.9
S	4.9	18.3	20.7	56.1	2.4	4.9	8.5	84.1	0	0	8.5	91.5

3.3. Decision Level Fusion for Tea Quality Identification

The feature-level fusion, together with the kernel dimensional reduction method, could correctly identify most of the tea samples. However, it was not ideal for all the tea samples collected from the various companies. The D-S evidence-based decision fusion was conducted to improve the classification. As we discussed in Section 2.4.5, the basic probability assignment function was defined by the membership degree from the K -nearest neighbors. In this study, the value K related to the highest recognition accuracy was selected, where the parameter K was set to three for the KNN classifier and the subsequent analysis.

The D-S evidence was conducted on the classification results obtained by feature fusion + KLDA + KNN for both the E-nose and E-tongue. The maximum membership rule defined in Equation (8) was utilized to identify the grade of the tea samples. Based on the multi-level fusion method, the average recognition accuracy achieved for the testing set was 91.3%, compared to 71.3% for the E-nose, and 82.7% for the E-tongue. The corresponding confusion matrix is shown in Table 1. The diagonal of the confusion matrix denoted the correct recognition rates of the four grades, and the other elements denoted the misclassified rates among the four grades. Some tea samples with grade T were misclassified as grade Y (8.0%) for the E-nose, and it was even worse for the E-tongue (16.0%), where the misclassification accuracy reduced to 2.7% for the decision fusion scheme. For the E-nose, the samples of grade Y were misclassified to grade T (19.0%), and for the E-tongue, the misclassification rate was 8.3%. This was improved by the fusion method, producing a value of 1.2%. Moreover, the recognition accuracy of the grade Y samples increased to 94% for the fusion output, and there was no sample misclassified to grade S. The tea samples of grade E were easily misclassified to the other three categories. The recognition accuracy of decision fusion for grade E was enhanced compared to the E-nose and E-tongue. For the E-nose, the number of grade S samples that were misclassified as grade Y and grade E were 18.3% and 20.7%, respectively. The number of grade S samples that were misclassified as grade Y and E were 4.9% and 8.5% for the E-tongue, respectively. For the fusion method, the grade S tea samples were well identified with a recognition accuracy of 91.5%, while no misclassification occurred for grade T and grade Y samples. It seemed that the multi-level fusion strategy combining the E-nose and E-tongue data significantly improved the recognition accuracy. The grade T, Y, and S samples could be distinguished effectively, and the recognition accuracy for these three grades was above 90% for the large quantity of tea samples.

4. Conclusions

A multi-level fusion strategy which combines an electronic nose and electronic tongue was proposed and applied to tea quality identification. The procedure included feature-level fusion (fuse the time-domain based feature and frequency-domain based feature) and decision-level fusion (D-S evidence to combine the classification results from KNN). The experiments were carried out on tea samples collected from various tea providers with four grades which made it very difficult, and the results showed a much better classification ability for the multi-level fusion system. The proposed algorithm could better represent the overall characteristics of the tea samples for both odor and taste, and that was the reason why the system outperformed the single feature method or single classifier.

Furthermore, the general idea of this paper can be improved by selecting better feature representing methods, and more attributes like appearance may be considered for this multi-level fusion system. This paper provides a framework for the combination of features and decisions, and it is important consider the difference sources to better identify the tea quality.

Supplementary Materials: The supplementary materials are available online at <http://www.mdpi.com/1424-8220/17/5/1007/s1>.

Acknowledgments: This work is supported by the National Natural Science Foundation of China (No. 61673052, No. 31201358), and the Fundamental Research Fund for the Central Universities of China (06116070), the National Research and Development Major Project (SQ2017YFNC010027), National High Technology Research and Development Program 863 (No. 2011AA1008047), and the grant from Chinese Scholarship Council (CSC).

Author Contributions: R. Zhi and L. Zhao conceived and designed the experiments, and performed the experiments; R. Zhi and D. Zhang analyzed the data, and R. Zhi wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bhattacharyya, N.; Bandyopadhyay, R.; Bhuyan, M.; Tudu, B.; Ghosh, D.; Jana, A. Electronic Nose for Black Tea Classification and Correlation of Measurements with “Tea Taster” Marks. *IEEE Trans. Instrum. Meas.* **2008**, *57*, 1313–1321. [[CrossRef](#)]

2. Dutta, R.; Hines, E.L.; Gardner, J.W.; Kashwan, K.R.; Bhuyan, M. Tea Quality Prediction Using a Tin Oxide-Based Electronic Nose: An Artificial Intelligence Approach. *Sens. Actuators B Chem.* **2003**, *94*, 228–237. [[CrossRef](#)]
3. Kiani, S.; Minaei, S.; Ghasemi-Varnamkhasti, M. Fusion of Artificial Senses as a Robust Approach to Food Quality Assessment. *J. Food Eng.* **2016**, *171*, 230–239. [[CrossRef](#)]
4. Ghasemi-Varnamkhasti, M.; Mohtasebi, S.S.; Siadat, M.; Razavi, S.H.; Ahmadi, H.; Dicko, A. Discriminatory Power Assessment of The Sensor Array of an Electronic Nose System for the Detection of Non-Alcoholic Beer Aging. *Czech J. Food Sci.* **2012**, *30*, 236–240.
5. Banerjee, R.; Tudu, B.; Bandyopadhyay, R.; Bhattacharyya, N. A Review on Combined Odor and Taste Sensor Systems. *J. Food Eng.* **2016**, *190*, 10–21. [[CrossRef](#)]
6. Baldwin, E.A.; Bai, J.; Plotto, A.; Dea, S. Electronic Noses and Tongues: Applications for the Food and Pharmaceutical industries. *Sensors* **2011**, *11*, 4744–4766. [[CrossRef](#)] [[PubMed](#)]
7. Persaun, K. Electronic Noses and Tongues in the Food Industry. In *Electronic Noses and Tongues in Food Science*, 1st ed.; Mendez, M.L.R., Preedy, V.R., Eds.; Elsevier Academic Press: London, UK, 2016; Chapter 1; pp. 1–12.
8. Chen, Q.; Zhao, J.; Chen, Z.; Lin, H.; Zhao, D.A. Discrimination of Green Tea Quality Using the Electronic Nose Technique and the Human Panel Test, Comparison of Linear and Nonlinear Classification Tools. *Sens. Actuators B Chem.* **2011**, *159*, 294–300. [[CrossRef](#)]
9. Kaur, R.; Kumar, R.; Gulati, A.; Ghanshyam, C.; Kapur, P.; Bhondekar, A.P. Enhancing Electronic Nose Performance: A Novel Feature Selection Approach Using Dynamic Social Impact Theory and Moving Window Time Slicing for Classification of Kangra Orthodox Black Tea (*Camellia sinensis* (L.) O. Kuntze). *Sens. Actuators B Chem.* **2012**, *166*, 309–319. [[CrossRef](#)]
10. Yu, H.; Wang, Y.; Wang, J. Identification of Tea Storage Times by Linear Discrimination Analysis and Back-Propagation Neural Network Techniques Based on the Eigenvalues of Principal Components Analysis of E-nose Sensor Signals. *Sensors* **2009**, *9*, 8073–8082. [[CrossRef](#)] [[PubMed](#)]
11. Yu, H.; Wang, J.; Zhang, H.; Yu, Y.; Yao, C. Identification of Green Tea Grade Using Different Feature of Response Signal from E-Nose Sensors. *Sens. Actuators B Chem.* **2008**, *128*, 455–461. [[CrossRef](#)]
12. Bhattacharyya, N.; Seth, S.; Tudu, B.; Tamuly, P.; Jana, A.; Ghosh, D.; Bandyopadhyay, R.; Bhuyan, M. Monitoring of Black Tea Fermentation Process Using Electronic Nose. *J. Food Eng.* **2007**, *80*, 1146–1156. [[CrossRef](#)]
13. Ivarsson, P.; Holmin, S.; Höjer, N.; Krantz-Rülcker, C.; Winqvist, F. Discrimination of tea by means of a voltammetric electronic tongue and different applied waveforms. *Sens. Actuators B Chem.* **2001**, *76*, 449–454. [[CrossRef](#)]
14. Palit, M.; Tudu, B.; Dutta, P.K.; Dutta, A.; Jana, A.; Roy, J.K.; Bhattacharyya, N.; Bandyopadhyay, R.; Chatterjee, A. Classification of Black Tea Taste and CORRELATION with Tea Taster's Mark Using Voltammetric Electronic Tongue. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 2230–2239. [[CrossRef](#)]
15. Xiao, H.; Wang, J. Discrimination of Xihulongjing Tea Grade Using an Electronic Tongue. *Afr. J. Biotechnol.* **2009**, *8*, 6985–6992.
16. Chen, Q.; Zhao, J.; Vittayapadung, S. Identification of the Green Tea Grade Level Using Electronic Tongue and Pattern Recognition. *Food Res. Int.* **2008**, *41*, 500–504. [[CrossRef](#)]
17. Wu, J.; Liu, J.; Fu, M.; Li, G. Classification of Chinese Green Tea by a Voltammetric Electronic Tongue. *Chin. J. Sens. Actuators* **2006**, *19*, 963–965.
18. Lvova, L.; Legin, A.; Vlasov, Y.; Cha, G.S.; Nam, H. Multicomponent Analysis of Korean Green Tea by Means of Disposable All-Solid-State Potentiometric Electronic Tongue Microsystem. *Sens. Actuators B Chem.* **2003**, *95*, 391–399. [[CrossRef](#)]
19. Zhi, R.; Zhao, L.; Shi, B.; Jin, Y. New Dimensionality Reduction Model (Manifold Learning) Coupled with Electronic Tongue for Green Tea Grade Identification. *Eur. Food Res. Technol.* **2014**, *239*, 157–167. [[CrossRef](#)]
20. Haddi, Z.; Mabrouk, S.; Bougrini, M.; Tahri, K.; Sghaier, K.; Barhoumi, H.; Bari, N.E.; Maaref, A.; Jaffrezic-Renault, N.; Bouchikhi, B. E-Nose and E-tongue Combination for Improved Recognition of Fruit Juice Samples. *Food Chem.* **2014**, *150*, 246–253. [[CrossRef](#)] [[PubMed](#)]
21. Rudnitskaya, A.; Kirsanov, D.; Legin, A.; Beullens, K.; Lammertyn, J.; Nicolai, B.M.; Irudayaraj, J. Analysis of Apples Varieties—Comparison of Electronic Tongue with Different Analytical Techniques. *Sens. Actuators B Chem.* **2006**, *116*, 23–28. [[CrossRef](#)]

22. Natale, C.D.; Paolesse, R.; Macagnano, A.; Mantini, A.; D'Amico, A.; Legin, A.; Lvova, L.; Rudnitskaya, A.; Vlasov, Y. Electronic Nose and Electronic Tongue Integration for Improved Classification of Clinical and Food Samples. *Sens. Actuators B Chem.* **2000**, *64*, 15–21. [[CrossRef](#)]
23. Sole, M.; Covington, J.A.; Gardner, J.W. Combined Electronic Nose and Tongue for a Flavour Sensing System. *Sens. Actuators B Chem.* **2011**, *156*, 832–839.
24. Hong, X.; Wang, J. Detection of Adulteration in Cherry Tomato Juices Based on Electronic Nose and Tongue: Comparison of Different Data Fusion Approaches. *J. Food Eng.* **2014**, *126*, 89–97. [[CrossRef](#)]
25. Ouyang, Q.; Zhao, J.; Chen, Q. Instrumental Intelligent Test of Food Sensory Quality as Mimic of Human Panel Test Combining Multiple Cross-Perception Sensors and Data Fusion. *Anal. Chim. Acta* **2014**, *841*, 68–76. [[CrossRef](#)] [[PubMed](#)]
26. Qiu, S.; Wang, J.; Tang, C.; Du, D. Comparison of ELM, RF, and SVM on E-nose and E-tongue to Trace the Quality Status of Mandarin (Citrus unshiu Marc.). *J. Food Eng.* **2015**, *166*, 193–203. [[CrossRef](#)]
27. Buratti, S.; Ballabio, D.; Giovanelli, G.; Dominguez, C.M.Z.; Moles, A.; Benedetti, S.; Sinelli, N. Monitoring of Alcoholic Fermentation Using Near Infrared and Mid Infrared Spectroscopies Combined with Electronic Nose and Electronic Tongue. *Anal. Chim. Acta* **2011**, *697*, 67–74. [[CrossRef](#)] [[PubMed](#)]
28. Zakaria, A.; Shakaff, A.Y.; Masnan, M.J.; Ahmad, M.N.; Adom, A.H.; Jaafar, M.N.; Ghani, S.A.; Abdullah, A.H.; Aziz, A.H.; Kamarudin, L.M.; et al. A Biomimetic Sensor for the Classification of Honeys of Different Floral Original and the Detection of Adulteration. *Sensors* **2011**, *11*, 7799–7822. [[CrossRef](#)] [[PubMed](#)]
29. Rodriguez-Mendez, M.L.; Apetrei, C.; Gay, M.; Medina-Plaza, C.; De Saja, J.A.; Vidal, S.; Aagaard, O.; Ugliano, M.; Wirth, J.; Cheynier, V. Evaluation of Oxygen Exposure Levels and Polyphenolic Content of Red Wines Using an Electronic Panel Formed by an Electronic Nose and an Electronic Tongue. *Food Chem.* **2014**, *155*, 91–97. [[CrossRef](#)] [[PubMed](#)]
30. Zakaria, A.; Shakaff, A.Y.M.; Adom, A.H.; Ahmad, M.N.; Masnan, M.J.; Aziz, A.H.A.; Fikri, N.A.; Abdullah, A.H.; Kamarudin, L.M. Improved Classification of *Orthosiphon stamineus* by Data Fusion of Electronic Nose and Tongue Sensors. *Sensors* **2010**, *10*, 8782–8796. [[CrossRef](#)] [[PubMed](#)]
31. Fikri, N.A.; Adom, A.H.; Md. Shakaff, A.Y.; Ahmad, M.N.; Abdullah, A.H.; Zakaria, A.; Markom, M.A. Development of Human Sensory Mimicking System. *Sens. Lett.* **2011**, *9*, 423–427. [[CrossRef](#)]
32. Banerjee, R.; Chattopadhyay, P.; Tudu, B.; Bhattacharyya, N.; Bandyopadhyay, R. Artificial flavor perception of black tea using fusion of electronic nose and tongue response: A Bayesian statistical approach. *J. Food Eng.* **2014**, *142*, 87–93. [[CrossRef](#)]
33. Banerjee, R.; Tudu, B.; Shaw, L.; Jana, A.; Bhattacharyya, N.; Bandyopadhyay, R. Instrumental Testing of Tea by Combining the Responses of Electronic Nose and Tongue. *J. Food Eng.* **2012**, *110*, 356–363. [[CrossRef](#)]
34. Xu, Y.; Chen, W.; Yin, B.; Huang, L. Green Tea Storage and Preservation. *Agric. Mach. Technol. Ext.* **2004**, *9*, 29–32.
35. Shi, B.; Zhao, L.; Zhi, R.; Xi, X. Optimization of Electronic Nose Sensor Array by Genetic Algorithms in Xihu-Longjing Tea Quality Analysis. *Math. Comput. Model.* **2013**, *58*, 752–758. [[CrossRef](#)]
36. Dai, Y.; Zhi, R.; Zhao, L.; Gao, H.; Shi, B.; Wang, H. Longjing Tea Quality Classification by Fusion of Features Collected from E-nose. *Chemom. Intell. Lab. Syst.* **2015**, *144*, 63–70. [[CrossRef](#)]
37. Ma, B.; Qu, H.; Wong, H. Kernel clustering-based discriminant analysis. *Pattern Recognit.* **2007**, *40*, 324–327. [[CrossRef](#)]
38. Lin, T. Improving D-S evidence Theory for Data Fusion System. Available online: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.652&rep=rep1&type=pdf> (accessed on 10 April 2017).
39. Guan, J.W.; Bell, D.A. *Evidence Theory and its Applications*; Elsevier Science Inc.: New York, NY, USA, 1992.
40. Hegarat-Masclé, S.L.; Bloch, I.; Vidal-Madjar, D. Application of Dempster-Shafer Evidence Theory to Unsupervised Classification in Multisource Remote Sensing. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 1018–1031. [[CrossRef](#)]
41. Walley, P.; Moral, S. Upper Probabilities Based Only on the Likelihood Function. *J. R. Stat. Soc.* **1999**, *61*, 831–847. [[CrossRef](#)]

