Sensors and Materials, Vol. 19, No. 5 (2007) 287–298
MYU Tokyo

S & M 0683

Discrimination of Red Wine Age Using Voltammetric Electronic Tongue Based on Multifrequency Large-Amplitude Voltammetry And Pattern Recognition Method

Shi-Yi Tian, Shao-Ping Deng*, Chun-Hui Ding[1], Chun-Li Yin[1] and Hua Li[1]

Food Sensory Science Laboratory of Zhejiang Gongshang University, Food Safety Key Lab of Zhejiang Province, Zhejiang Gongshang University, Hangzhou 310035, P. R. China
[1]College of Enology, Northwest Science and Technology University of Agriculture and Forestry, Shaanxi, Yangling 712100, P. R. China

(Received April 2, 2007; accepted May 1, 2007)

Key words: electronic tongue, multifrequency large-amplitude voltammetry, multivariate data analysis, red wine

Three methods of multivariate data analysis (MVAD), principal component analysis (PCA), soft independent modeling of class analogy (SIMCA) and partial least squares discriminating analysis (PLS-DA), were used for processing data from a multifrequency large-amplitude pulse electronic tongue (MLAP-ET) in this paper. The dry red wine samples from the same company, produced by the same type of grape from the same vineyard, but with different vintages were studied using MLAP-ET. The results showed that these three methods were all effective for the data treatment of MLAP-ET to assess the vintage of red wine samples but differ in their discriminating ability. PLS-DA had the best classification property and was most suitable for processing the data from MLAP-ET.

1. Introduction

Although the traditional measurement technique in the food industry can be used to determine a specific parameter such as the conductivity, viscosity, pH or concentration of one component accurately, it has two obviously drawbacks: (i) one or several specific parameters cannot represent the integral quality of a complex food sample; for example, conductivity can only display the character of solutions' resistance and pH can only reveal the content of solutions' dissociated protons; (ii) generally, one instrument can

*Corresponding author: e-mail: spdeng@zjgsu.edu.cn

only give a specific parameter of the solution; thus, numerous instruments are needed for the evaluation of one system.

With the development of computer technology, multivariate data analysis (MVDA) has progressed considerably in the last two decades. A kind of analysis technique based on MVDA has emerged, in which lots of information variables with low selectivity, and partial overlapping are collected and processed. These analysis techniques were effectively applied for evaluating the integral character of products in the food industry, instead of one or several specific parameters of the samples. One of the available analysis techniques based on MVDA is called electronic tongue analysis, which has attracted great interest among researchers in recent years. The electronic tongue is a multisensor system, which consists of a number of low-selectivity sensors and uses advanced mathematical procedures for signal processing based on pattern recognition (PARC) and/or multivariate analysis [artificial neural networks (ANNs), principal component analysis (PCA), etc.].(1) It was a very good analysis tool for evaluating samples qualitatively and could replace the expert panelist for evaluating the quality of products due to the high cost of hiring expert panelists. It could also be used for quantitative analysis, i.e., for predicting the concentrations of compounds in the samples in some specific cases.

Recently, several kinds of electronic tongue have been developed. These devices could be classified into three types on the basis of the different electrochemical measurement techniques, namely, potentiometry, impedance spectroscopy and voltammetry. A device based on potential, called taste sensor, was first presented by Toko.(2) It was composed of several kinds of lipid/PVC membranes for transforming the taste quality information, such as sweetness, bitterness, sourness, saltiness, or umami into an electric signal. Different types of food such as soy sauce, beer, coffee and mineral water were investigated using the taste sensor. Not only the differences in food type were identified but also the taste interactions, such as the suppression effect, were detected.(3–7) Another type of potential electronic tongue was presented by Legin,(8) which was configured by using several nonspecific sensors based on chalcogenide glasses as transducers. It has been used for the discrimination of various foodstuffs and the analysis of some specific ions or species in solutions.(9–12) The second kind of electronic tongue based on impedance spectroscopy was first described by Riul, in which the sensors were constructed using supramolecular thin films of conducting polymers with a lipid-like material that were analyzed by impedance spectroscopy. This kind of electronic tongue could distinguish different brands of beverages and detect the low levels of taste, inorganic contamination in water and the suppression of the sense of taste.(13–15) The third kind of electronic tongue based on voltammetry, was first designed by Winquist.(16) It comprised several metallic electrodes (platinum, gold, palladium, iridium, rhenium and rhodium) as working electrodes, an Ag/AgCl reference electrode and a stainless-steel electrode as a counter electrode for standard three-electrode systems. It has been successfully used to analyze milk, tea, juice, drinking water and mold growth in liquid media.(17–20)

Among the three types of devices mentioned above, the voltammetric electronic tongue has some special advantages, such as very high sensitivity, versatility, simplicity and robustness and it has been extensively used in analytical chemistry. Recently, we have reported a novel electrochemical method: multifrequency large-amplitude pulse voltammetry (MLAPV) with three different frequency segments, 1 Hz, 10 Hz and 100 Hz, which was very useful in the voltammetric electronic tongue for discriminating samples. A novel voltammetric electronic tongue based on this MLAPV, called the multifrequency large-amplitude pulse electronic tongue (MLAP-ET), was also developed and it displayed excellent ability in the discrimination of six different Chinese liquors and seven Chinese Longjing teas. The results showed that MLAPV had a better classification property than commonly used large-amplitude voltammetry (LAPV).[21]

To increase the selectivity and sensitivity of the discriminating process, the data should be processed reasonably. As extended work on our electronic tongue based on MLAPV, we examined the effect of the data processing method on discriminating ability. In the present study, 147 red wine samples from the same factory, with the same typical grapes and the same vineyards, but from different vintages were studied using the MLAP-ET. Three kinds of MVAD, i.e., principal component analysis (PCA), soft independent modeling of class analogy (SIMCA) and partial least squares discriminating analysis (PLS-DA), were used for data treatment. PCA was used for reduction of the data and displaying the classification property of red wine samples using the MLAP-ET in the first step. SIMCA and PLS-DA were used for studying which method was more suitable for MLAP-ET in pattern recognition. The result showed that MLAP-ET could discriminate red wine samples with different vintages well by PCA, and that PLS-DA was more suitable for MLAP-ET to evaluate the age of red wine than SIMCA.

2. Materials and Methods

2.1 *Experimental setup*

Figure 1 shows the voltammetric electronic tongue (MLAP-ET) used in the experiments, which was described previously.[21] It consisted of three different metallic

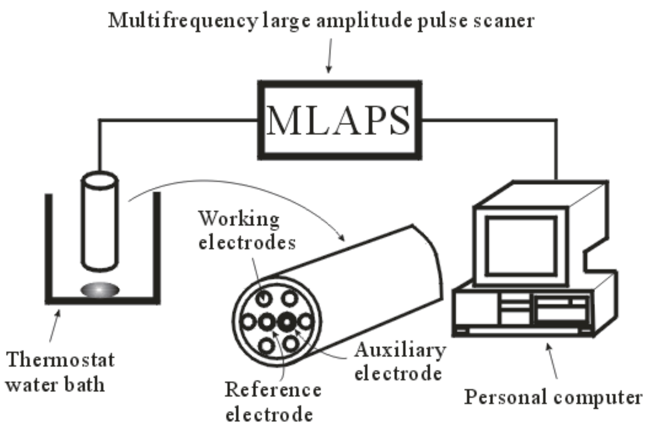

Fig. 1. Setup of the voltammetric electronic tongue.

disc electrodes (gold, silver, and titanium) as working electrodes, an Ag/AgCl electrode, a reference electrode (saturated KCl, diameter 2 mm), and a platinum counter electrode with a length of 5 mm and a diameter of 1 mm for standard three-electrode systems. The metal wire that served as working electrodes had a diameter of 2 mm and a purity of 99.9%. All the electrodes were made by Tianjing Aida Co. Ltd., China. A device called the "multifrequency large-amplitude pulse scanner (MLAPS)" (built at the lab of Sensory Science Zhejiang Gongshang University) was a potentiostat with six channels, which was controlled by a personal computer (PC). It could make the potential pulses/steps on the working electrodes and enable the working electrodes to work consecutively one at a time in three-electrode configurations, which were controlled by a relay in MLAPS. The PC was used to set and control the potential pulses, and measure and store current responses. A thermostat water bath was also applied to make the cell at a constant temperature.

2.2 *Treatment of samples*

147 bottles of red wine samples (see Table 1) was provided by Huaxia Great Wall Co. Ltd., which were stored in the cellar of the company in bottles. All the red wine samples were of the same type (dry). They were all produced from the same type of grape (cabernet sauvignon) and from the same vineyard. The only difference lies in their vintage.

Eighty milliliters of each sample was used for analysis using the electronic tongue at room temperature (approximately 22°C) in random. Each sample was analyzed three times, and then we averaged the data to produce one point in the PCA score plot.

2.3 *Voltammetric procedure and data extraction method*

The applied potential waveform was an MLAPV one, which was described in the previously paper.[21] It consisted of three segments with frequencies of 1 Hz, 10 Hz and 100 Hz. The waveform of each segment of the LAPV waveform had the maximal value at 1.0 V and the minimal value at –1.0 V. The amplitude of each pulse was 0.1 V. A step of 0 V was inserted immediately before and after the pulse.

Four points of each cycle were collected, which were related to the concentration and diffusion coefficients of the charged and electroactive compounds in the solution (see Fig. 2).

Table 1
Red wine samples used in the treatments.

Samples	Brand	Type	Grape typical	Vintage	Amount (bottle)
S1	Great wall	Dry	Cabernet sauvignon	1993	25
S2	Great wall	Dry	Cabernet sauvignon	1996	25
S3	Great wall	Dry	Cabernet sauvignon	2001	26
S4	Great wall	Dry	Cabernet sauvignon	2003	26
S5	Great wall	Dry	Cabernet sauvignon	2004	24
S6	Great wall	Dry	Cabernet sauvignon	2005	21

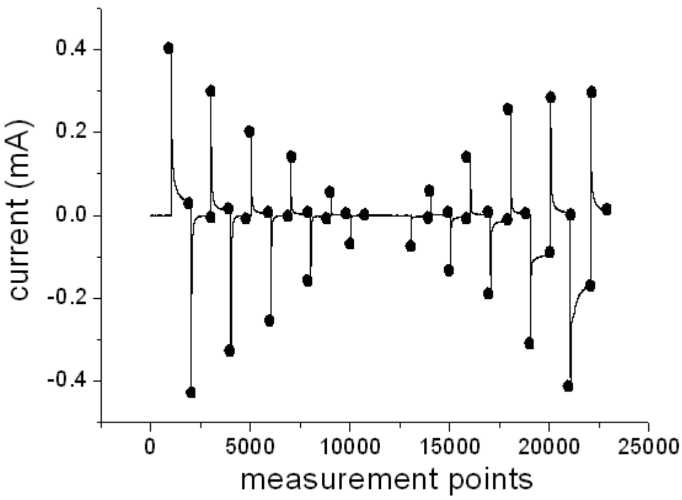

Fig. 2. Method of extracting data from the original data of one segment of one working electrode.

2.4 *Data processing method*

2.4.1 *Principal component analysis*

Principal component analysis (PCA) is a popular and useful MVDA method for the analysis of experimental data.[1,4,18] It can decompose the experimental data matrix into latent variables and explain the variables in data using the loading plot from the loading vectors and the score plot from the score vectors. The loading plot could elucidate the relationships between the original variables and their influence on the system. The score plot displays the relationships between the samples of the experiments, and could be used for grouping and classifying the observations.

2.4.2 *Soft independent modeling of class analogy*

Soft independent modeling of class analogy (SIMCA) is another kind of MAVD, which is popularly applied in pattern recognition.[1,19,22–23] It builds a class model on the basis of the principal component analysis of each separated category. Some parameters would be obtained from the model of separated class samples, i.e., the number of significant components “A”, the mean standard deviation of each class “Sc” and the degree of freedom of each class, among others. These parameters are then used to classify external samples (the prediction set) on the basis of the same variables. If appropriate, the model can then be used to classify unknown samples.

2.4.3 *Partial least squares discriminating analysis*

Partial least squares discriminating analysis (PLS-DA) is a supervised pattern recognition method based on partial least squares (PLS). Y-variables could be set for different groups. How many groups respond to how many Y-variables. The Y-variables that respond to the samples with the same group are valued 1, otherwise they are valued 0. PLS-DA builds the model by relating the variations in one or several group variables (Y-variables) to sample variables (X-variables) based on PLS. Then it uses the X-variables of the unknown samples to predict the group variables of Y-variables. The samples with the group of Y-variables larger than 0.5 and a deviation that does not cross the 0.5 line are predicted to be in this group. This method works particularly well when the various X-variables express common information, i.e., when there is a large amount of correlation, or even collinearity due to its being based on PLS.[1,24–25]

The above three methods were performed using SPSS11.0 and the unscrambler 9.1.

3. Results

3.1 *Performance of PCA*

The data from the MLAP-ET was treated as described in the previous paper.[21] The MLAPV that was applied in MLAP-ET was a useful electrochemical method for the electronic tongue and it made one individual working electrode display different classification properties for samples at different frequency segments. The data of each of the three MLAPV frequency segments from one working electrode was processed by PCA, and the score plot of each frequency segment of MLAPV from one working electrode would show different sample discrimination abilities. Then the data of one frequency segment of MLAPV from one working electrode, which showed the best classification property, were picked out and merged as the data of the sensor array of MLAP-ET. In the present work, the score plot of the silver working electrode exhibited the best discriminating ability from the 100 Hz frequency segment of all three frequency segments of MLAPV, while the score plots of the titanium working electrode and gold working electrode showed the best classification property from the 1 Hz frequency segment and 10 Hz frequency segment of their own three frequency segments of MLAPV, respectively. As the data treatment of MLAPV, the data of the 100 Hz frequency segment from the silver working electrode, the 1 Hz frequency segment from the titanium working electrode and the 10 Hz frequency segment from the gold working electrode were merged as the data from the sensor array of MLAP-ET for PCA processing (see Figs. 3 and 4). Also, the other choices of the combination of the data from each frequency segment were tried. The score plot from the merged data of the above frequency segments of individual working electrodes had the best discriminating ability.

Figures 3 and 4 showed the final PCA results from the data of the sensor array of MLAP-ET. The first three components contained 78.9% of all the information. It can be seen that the score plot of PC1 and PC2 could not discriminate samples with six different kinds of vintage well. The sample clusters with vintages of 1993, 2001, 2003 and 2005 overlapped each other. However, the red wine samples were classified well in the score plot of PC1 and PC3. They were well separated according to the different vintages. Only the sample with a vintage of 2005 and the sample with a vintage of 2001 overlapped with the cluster with a vintage of 2003.

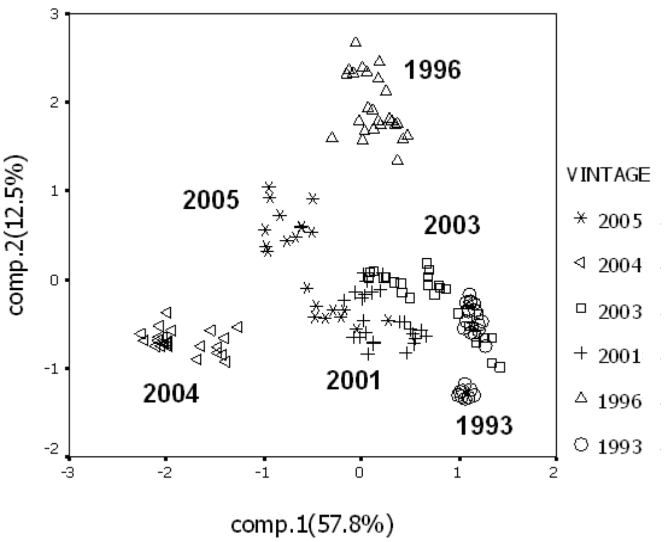

Fig. 3. Scope plot from MLAP-ET by PC1 vs PC2.

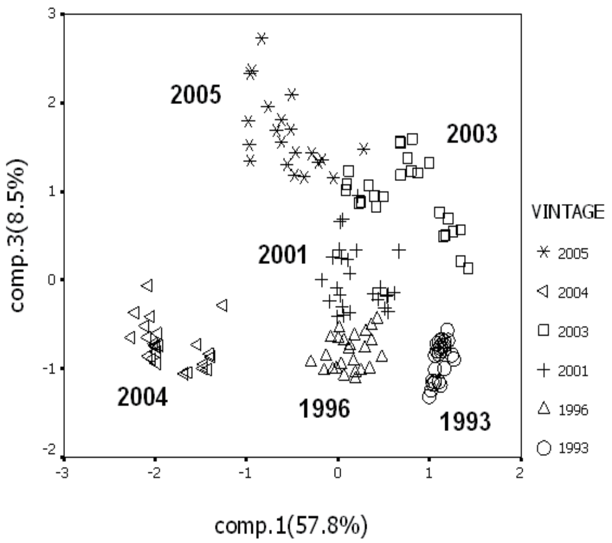

Fig. 4. Score plot from MLAP-ET by PC1 vs PC3.

3.2 *Performance of SIMCA*

The PCA results showed that MLAP-ET can classify red wine samples according to their vintage. PCA was a simple unsupervised pattern recognition method and it is popularly used because it can be used when samples are few. However, supervised pattern recognition methods were needed when the electronic tongue was applied in industry. SIMCA as a supervised recognition method was used. Also, the data of each sample obtained using MLAP-ET was the same as that obtained by PCA.

In the SIMCA procedure, the data of five samples of every vintage were selected at random as the test samples for classification. Thus, 20 samples with vintages of 1993 and 1996, 21 samples with vintages of 2001 and 2003, 19 samples with a vintage of 2004 and 16 samples with a vintage of 2005 remained for building the SIMCA models. Seven components for the model of the samples with the vintages of 1993 and 1996, 5 components for the model of the samples with the vintages of 2001 and 2005, 4 components for the model of the samples with the vintages of 2003 and 6 components for the model of the samples with the vintage of 2004 were chosen, to make the total explained variance more than 85%. The results of the classification of test samples are shown in Table 2. Two samples with the vintage of 2003 and one sample with the vintage of 2001 were incorrectly classified to be the model of the samples of 2001 and the model of the samples of 2003, respectively. Compared with those in Fig. 4, the samples with the vintage of 2001 overlapped with the samples of the vintage of 2003. However, the SIMCA method discriminated the samples with the vintage of 2005 well, while one sample with the vintage of 2005 was overlapped with the cluster with the vintage of 2003.

3.3 *Performance of PLS-DA*

PLS-DA as another supervised pattern recognition method was also used for the classification of samples with different vintages. The x-variables of the samples for building models and predicting groups were the same as those for SIMCA. Six y-variables were created for the presented six groups with different vintages, which were labeled 1993, 1996, 2001, 2003, 2004 and 2005, respectively. The y-variables of the samples with the same vintage were assigned 1, otherwise they were assigned 0.

Table 2
Results of SIMCA in the significance limited of 5%.

Name of models	Vintage of the predicted samples: 1993	1996	2001	2003	2004	2005
1993	*****					
1996		*****				
2001			*** *	**		
2003			*	***		
2004					*****	
2005						*****

"*" represents the number of samples that the model discriminated.

Ten components of the PLS model were chosen to make the total explained variance more than 85%. The results showed that all the samples with different vintages were effectively classified by PLS-DA (Table 3 and Fig. 5). It can be seen that the predicted samples with the vintages of 1993, 1996, 2004 and 2005 were well classified by PLS-DA, whose deviations were less than 0.1 (see Table 3 and Figs. 5(a), 5(b), 5(e), and 5(f)). The samples with the vintages of 2001 and 2003 are less effectively classified by PLS-DA due to its relatively large deviation (see Figs. 5(c) and 5(d) and Table 3). However, they were still discriminated by PLS-DA, while two samples with the vintage of 2003 and one sample with the vintage of 2001 were not successfully classified by SIMCA.

4. Discussion

In this work, 147 bottles of red wine with vintage differences, were studied using MLAP-ET based on MLAPV. The PCA showed that the sensor array data were merged with the data of the 100 Hz frequency segment from the silver working electrode, the 1 Hz frequency segment from the titanium working electrode, and the 10 Hz frequency segment from gold working electrode was the best for classifying the samples according to vintage (see Fig. 4). This result verified that MLAPV was a useful electrochemical method for the voltammetric electronic tongue, and different working electrodes needed specific frequency segments for the analysis of samples.

Three MAVD techniques, i.e., PCA, SIMCA and PLS-DA, were used for processing the data from MLAP-ET. PCA is an unsupervised pattern recognition method. It is very popularly used for processing the data of the electronic tongue,(1,4,18) since it can give good results when there is a spot of samples with a large number of variables and can easily discriminate the samples with different characters in 2 D or 3 D plots with the help of human eyes. It can be seen from Fig. 4 that all the samples can be separated by the score plot of PC1 vs PC3 from PCA. However, the clusters of the samples with the vintages of 1996, 2001, 2003 and 2005 were close to each other, and one sample with the vintage of 2005 and one sample with the vintage of 2001 were both overlapped

Table 3
Result of PLS-DA.

Vintage of the predicted samples	PLS-DA		
	Discrimination ratio	Mean of predicted	Mean of deviation
1993	100%	0.93 (0)	0.08 (0.09)
1996	100%	0.98 (0)	0.05 (0.04)
2001	100%	0.78 (0.03)	0.14 (0.15)
2003	100%	0.84 (0.05)	0.13(0.14)
2004	100%	0.96 (0.02)	0.07 (0.07)
2005	100%	0.97 (0)	0.09 (0.08)

The number in the brackets represents the predicted Y-variable responding to the samples that did not belong to this group.

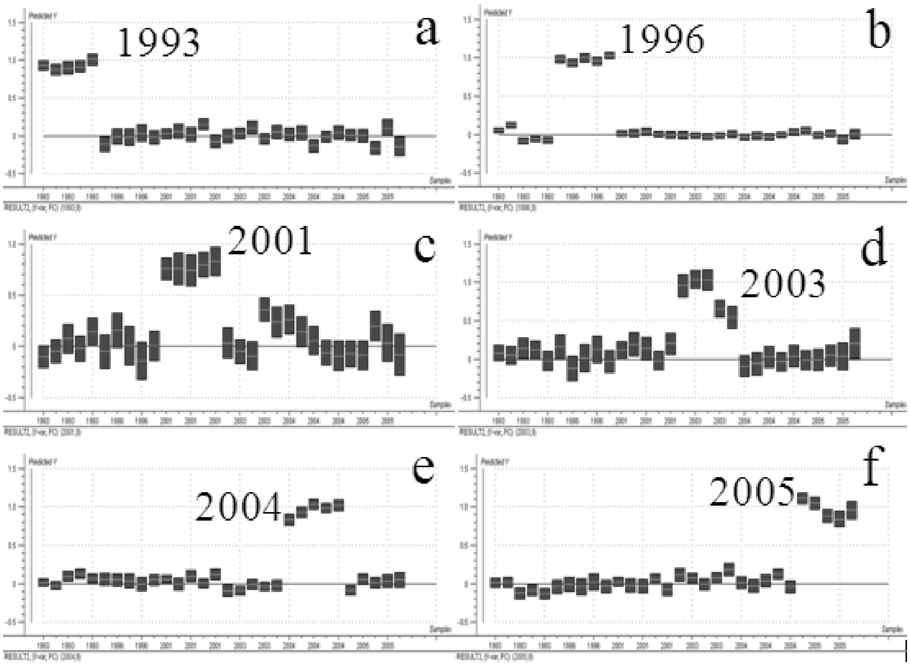

Fig. 5. Plots from PLS-DA of predicted Y of various samples. (a) Predicted Y of the samples with the vintage of 1993; (b) predicted Y of the samples with the vintage of 1996; (c) predicted Y of the samples with the vintage of 2001; (d) predicted Y of the samples with the vintage of 2003; (e) predicted Y of the samples with the vintage of 2004; (f) Predicted Y of the samples with the vintage of 2005.

with the sample with the vintage of 2003. The reason lies in the fact that the first three components only contain 78% of all the data information. PCA had some drawback that it could not classify samples well when the first three components had less than 80% of the information and the mean of the specific component usually was not clear. Maybe the score plot would be better when the first three components contain more information after some advanced data treatment skills are incorporated in PCA.

Two supervised pattern recognition methods, SIMCA and PLS-DA, were used for further data processing. SIMCA was the first class modeling technique introduced in chemistry and is commonly used for classifying samples.[1,19,22–23] PLS-DA was another supervised pattern recognition method that extended from PLS, which has been rapidly developed in recent years.[24,25] The results of two methods showed that the samples with the vintages of 2001 and 2003 could not be easily classified. These results were similar to those in Fig. 4; the cluster of the samples with the vintages of 2001 and 2003 had a partial overlap and a large dispersion. The discrimination ratio of PLS-DA (100%) is higher than that of SIMCA (90%). In the SIMCA method, two samples with the vintage of 2003 and one sample with the vintage of 2001 were not classified correctly. All the results showed that PLS-DA was more effective here. The reason may be that the applied waveform MLAPV was designed to promote the interaction of potential pulses with each other to produce new information, and the data derived from MLAPV which changed the length of the large-amplitude pulse voltammetry (LAPV) signal contain a lot of nonlinear information. It is well known that PLS is more effective for nonlinear data processing than PCA;[24,25] thus, PLS-DA pattern recognition based on PLS shows a better discriminating ability than SIMCA, which is based on PCA. The analysis of the vintage of red wine herein showed that PLS-DA was more suitable than SIMCA for MLAP-ET data processing. Further work is needed to extend this method to the evaluation of more samples, and investigate other nonlinear data processing methods for MLAP-ET.

Acknowledgement

The project was supported by the Science and Technology Department of Zhejiang Province (Grant No. 2006c23082). The authors gratefully acknowledge Hua Xia Great Wall Co., Ltd. for providing the red wine samples for the experiment, Xiao-Ping Wang (Zhejiang University) for the technical support for hardware fabrication and adjustments and Zhong-Xiu Chen (Zhejiang Gongshang University) for lots of useful advice.

References

1 Yu. Vlasov, A. Legin, A. Rudnitskaya, C. Dinatale and A. D’amico: Pure Appl. Chem. **77** (2005) 1965.
2 S. Iiyama, S. Ezaki, K. Toko, T. Matsuno and K. Yamafuji: Sens. Actuators B **24** (1995) 75.
3 K. Toko: Sens. Actuators B **64** (2000) 205.
4 H. Yamada, Y. Mizota, K. Toko and T. Doi: Mater. Sci. Eng. C **5** (1997) 41.
5 S. Iiyama, M. Yahiro and K. Toko: Sens. Actuators B **66** (2000) 205.
6 H. Sakai, S. Iiyama and K. Toko: Sens. Actuators B **66** (2000) 251.
7 M. Habara, H. Ikezaki and K. Toko: Biosens. Bioelectron. **19** (2004) 1559.
8 Y. Vlasov, A. Legin and A. Rudnitskaya: Sens. Actuators B **44** (1997) 532.
9 A. Legin, A. Rudnitskaya, Y. Vlasov, C. D. Natale, F. Davide and A. D’Amico: Sens. Actuators B **44** (1997) 291.
10 C. D. Natale, A. Macagnano, F. Davide, A. D’Amico, A. Legin, Y. Vlasov, A. Rudnitskaya and B. Selezenev: Sens. Actuators B **44** (1997) 423.
11 A. Legin, A. Rudnitskaya, Y. Vlasov, C. D. Natale, E. Mazzone and A. D’Amico: Sens. Actuators B **65** (2000) 232.
12 A. Legin, A. Rudnitskaya, L. Lvova, Y. Vlasov, C. D. Natale and A. D’Amico: Anal. Chim. Acta. **484** (2003) 33.
13 A. Riul Jr., R. R. Malmegrim, F. J. Fonseca and L. H. C. Mattoso: Biosens. Bioelectron. **18** (2003) 1365.
14 A. Riul Jr., A. M. Gallardo, S. V. Mello, S. Bone, D. M. Taylor and L. H. C. Mattoso: Synthetic Met. **132** (2003) 109.
15 A. Riul Jr., H. C. D. Sousa, R. R. Malmegrim, D. S. D. Santos Jr., A. C. P. L. F. Carvalho, F. J. Fonseca, O. N. Oliveira Jr. and L. H. C. Mattoso: Sens. Actuators B **98** (2004) 77.
16 F. Winquist, P. Wide and I. Lundström: Anal. Chim. Acta **357** (1997) 21.
17 F. Winquist, S. Holmin, C. Krantz-Rülcker, P. Wide and I. Lundström: Anal. Chim. Acta **406** (2000) 147.

18 P. Ivarsson, S. Holmin, N. E. Höjer, C. Krantz-Rülcker and F. Winquist: Sens. Actuators B **76** (2001) 449.
19 C. Söderström, F. Winquist and C. Krantz-Rülcker: Sens. Actuators B **89** (2003) 248.
20 F. Winquist, R. Bjorklund, C. Krantz-Rülcker, I. Lundström, K. Östergren and T. Skoglund: Sens. Actuators B **111** (2005) 299.
21 S. Y. Tian, S. P. Deng and Z. X. Chen: Sens. Actuators B: Chem. **123** (2007) 1049.
22 V. Parra, Á. A. Arrieta, J. A. Fernández-Escudero, M. Íñiguez, J. A. Saja and M. L. Rodríguez-Méndez: Anal. Chim. Acta **563** (2006) 229.
23 M. Casale, C. Armanino, C. Casolino and M. Forina: Anal. Chim. Acta (2006) in press.
24 S. Wold, M. Sjöström and L. Eriksson: Chemometr. Intell. Lab. **58** (2001) 109.
25 P. Ciosek, Z. Brzózka, W. Wróblewski, E. Martinelli, C. Di Natale and A. D'Amico: Talanta **67** (2005) 590.