

Article

A Non-Destructive Distinctive Method for Discrimination of Automobile Lubricant Variety by Visible and Short-Wave Infrared Spectroscopy

Lulu Jiang ¹, Fei Liu ^{2,*} and Yong He ^{2,*}

¹ Zhejiang Technology Institute of Economy, Hangzhou 310018, China; E-Mail: abbyu111@sina.com

² College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China

* Authors to whom correspondence should be addressed; E-Mails: fliu@zju.edu.cn (F.L.); yhe@zju.edu.cn (Y.H.); Tel.: +86-571-8898-2143; Fax: +86-571-8898-2143.

Received: 6 January 2012; in revised form: 27 February 2012 / Accepted: 5 March 2012 /

Published: 12 March 2012

Abstract: A novel method which is a combination of wavelet packet transform (WPT), uninformative variable elimination by partial least squares (UVE-PLS) and simulated annealing (SA) to extract best variance information among different varieties of lubricants is presented. A total of 180 samples (60 for each variety) were characterized on the basis of visible and short-wave infrared spectroscopy (VIS-SWNIR), and 90 samples (30 for each variety) were randomly selected for the calibration set, whereas, the remaining 90 samples (30 for each variety) were used for the validation set. The spectral data was split into different frequency bands by WPT, and different frequency bands were obtained. SA was employed to look for the best variance band (BVB) among different varieties of lubricants. In order to improve prediction precision further, BVB was processed by UVE-PLS and the optimal cutoff threshold of UVE was found by SA. Finally, five variables were mined, and were set as inputs for a least square-support vector machine (LS-SVM) to build the recognition model. An optimal model with a correlation coefficient (R) of 0.9850 and root mean square error of prediction (RMSEP) of 0.0827 was obtained. The overall results indicated that the method of combining WPT, UVE-PLS and SA was a powerful way to select diagnostic information for discrimination among different varieties of lubricating oil, furthermore, a more parsimonious and efficient LS-SVM model could be obtained.

Keywords: lubricant; visual and short-wave spectroscopy; wavelet packet transform; uninformative variable elimination; simulated annealing algorithm

1. Introduction

Automobile lubricating oil is a kind of efficient anti-friction agent, mainly used to reduce the friction between the surfaces of moving parts. At present, there are many different varieties of lubricating oil, and these varieties greatly affect the market price and quality. Recently, consumption of lubricating oil has increased, as a result, instances of fraud have increased as well. To make enormous profits, some illegal factories mix varieties of lubricating oil of different quality, and such behaviors infringe on the rights and interests of consumers and legal factories. Therefore, there is a need to develop an accurate and rapid method to discriminate between varieties and qualities of lubricants, which may be also utilized for the detection of adulteration.

Recently, some researchers have devoted much attention to the study of lubricants. Zhao *et al.* studied the content of wear metal in lubrication oil by ICP-AES [1]. Wei *et al.* analyzed the additive element contents in lube oil by atomic emission spectra [2]. Borin *et al.* applied mid-infrared spectroscopy for quantification of contaminants in lubricating oil [3]. However, there are few reports on the discrimination of different varieties of lubricants using VIS-SWNIR. With the development of modern chemometrics and instrumentation, VIS-SWNIR has proved to be useful in providing a cheap and rapid detection technology [4–7], which is widely applied to discriminate agricultural and food production, such as yogurts [4], fruit vinegars [5], Chinese bayberries [6] and tea plants [7]. However, the spectra typically consist of broad, weak, non-specific and overlapping bands [8], and the data matrix of the spectra is often very large, so how to obtain the most useful information accurately from the data matrix is of great importance.

Wavelet packet transform (WTP), an extension of wavelet transform (WT) [9], is a powerful signal processing technique. It transforms the raw spectral data into different frequency bands, and the frequency component in different bands has different contribution to the multivariate model [10], so how to find the most useful band that represents the most variant spectral information is an important issue.

Uninformative variable elimination by PLS (UVE-PLS) is a method for variable selection [11]. The method evaluates the reliability of each variable in the model based on analysis of regression coefficients of PLS and selection criterion. It has been widely applied in analytical chemistry for removing the low-frequency varying background and the high-frequency noise [10], retention prediction of peptides [12], and analysis of steroids [13]. In these researches, they proposed to use an artificial random variable matrix with very small amplitude, added to the original data set to estimate the cutoff. This method is experiential.

The simulated annealing (SA) algorithm, a simulation of the annealing process used for metals, was put forward by Kirkpatrick in 1983 [14]. It offers arguably the simplest and the most elegant solution with the “best” record for solving combinatorial optimization problems. Unlike other algorithms, the SA algorithm allows various types of transitions in which some of them may be opposite the goal [15]. Hence, the SA algorithm has been widely applied to many optimization problems, such as multi-objective optimization of a constrained problem [16], the maximum clique problem [17] and multiparameter analysis of water optical properties from above-water remote-sensing reflectance [18]. However, the application of WTP combined with SA in extraction of variance spectra has received little attention with no contributions reported in the literature.

In this work, WPT with SA was used to search for BVB, then irrelevant variables in BVB were eliminated by UVE-PLS, and SA was employed to search for the cutoff threshold of UVE-PLS. Finally, the variables were mined as input sets for a LS-SVM to build a lubricant recognition model.

2. Experimental Section

2.1. Sample Preparation

A total of 180 lubricant samples were used as the whole data set. The calibration set of 90 samples was selected randomly for the optimal parameters. The remaining 90 samples were selected as validation set to evaluate the performance of discrimination model. The 180 samples were purchased in local market including the following three varieties: Changcheng (Cc), Huaxiayyanguang (Hxyg), Caltex (Ca). All samples were stored in the lab at a constant temperature of 25 ± 1 °C to equalize the temperature.

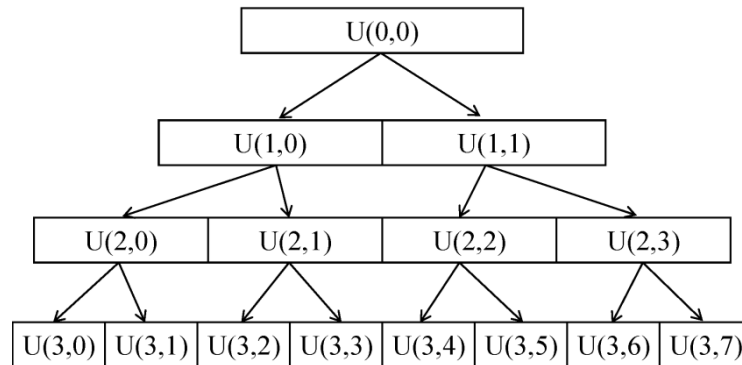
2.2. Spectral Measurement

For each sample, three reflectance spectra were scanned by a handheld FieldSpec Pro FR (325–1,075 nm)/A110070 (Analytical Spectral Devices Inc., Boulder, CO, USA). The light source consists of a Lowell pro-lam interior light source assemble/128930 with a Lowell pro-lam 14.5 V Bulb/128690 tungsten halogen bulb that could be used both in the visible and near infrared regions. The field-of-view of the spectral radiometer is 10°. The spectroradiometer was placed at a height of approximately 250 mm and at a 45° angle away from the center of the sample. The light source was placed at a height of approximately 150 mm and 45° angle away from the sample. The spectrum of each sample was the average of 30 successive scans with 1.5 nm intervals. Three spectra were collected for each sample and the average spectrum of these two measurements was used in the later analysis. All spectral data were stored in a computer and processed using the RS3 software for Windows (Analytical Spectral Devices) designed with a Graphical User Interface.

2.3. Algorithm

2.3.1. Wavelet Packet Transform

WPT is a derivative of WT. In the fast wavelet transform (WT) [19], a partial multi-resolution analysis is performed. Only approximation coefficients (low-pass node) are employed to deduce both scale and wavelet coefficients at the next resolution level. However, WPT allows a full multi-resolution analysis; both the approximation and detailed coefficients (high-pass node) are involved to decompose at the next resolution level at the same time [20]. As a result, a library of sub-bands including low frequency and high frequency is obtained. A schematic diagram for the WPT decomposition was shown in Figure 1. In the diagram, each node is identified by the pair of indices $U(j,k)$, where j is the level of decomposition and k is the position of the node at that level of decomposition.

Figure 1. Full WPT binary tree.

2.3.2. Uninformative Variable Elimination by Partial Least Squares (UVE-PLS)

In linear least squares models, the predictions \hat{y} are computed by the equation:

$$\hat{y} = Xb + e \quad (1)$$

where X is a $n \times p$ matrix containing p spectral responses of n samples, $b(1, p)$ is the vector of PLS regression coefficients and $e(n, 1)$ is vector of errors that cannot be explained by the model.

In UVE-PLS method, a PLS regression coefficient matrix $b = [b_1, \dots, b_n]$ is calculated through a leave-one-out validation [8,11], then the reliability of each variable (wavelength) can be quantitatively measured by its stability; the stability of variable j can be calculated as:

$$S_j = \text{mean}(\beta_j) / \text{std}(\beta_j) \quad (2)$$

where $\text{mean}(\beta_i)$ and $\text{std}(\beta_i)$ are the mean and standard deviation of the regression coefficients of variable j , so, the larger the stability, the more important the corresponding variables is, and the variables whose stability is lower than a cutoff threshold are regarded as uninformative and should be eliminated.

2.3.3. Principle of SA Algorithm

In simulated annealing, a problem starts with an initial solution, and this solution can be easily changed. But as the temperature t is decreased, changing the configuration is increasingly difficult. Finally, if t is lowered sufficiently, no further changes in the solution space are possible. To avoid being frozen at a local optimum, the SA algorithm moves slowly through the solution space. This controlled improvement of the objective value is accomplished by accepting non-improving moves with a certain probability that decreases as the algorithm progresses. This criterion is a Boltzman's probability distribution (Metropolis criterion) as a function of temperature t :

$$p(\Delta d) = \exp\left(\frac{-\Delta d}{t}\right) \quad (3)$$

where

$$\Delta d = d(x') - d(x_i) \quad (4)$$

d is a distance operator, and x' and x_i are current values.

The main processes of the SA algorithm for optimal parameters are explained as follows:

Step 1: Initialize parameters: set initial temperature T_I , termination temperature T_s , cooling coefficient τ of temperature, and select an initial solution x_I in solution space X . Finally, input the best value F^{best} and the corresponding solution:

$$\begin{aligned} x^{best} \\ x^{best} &= x_1 \\ F^{best} &= F(x_1) \\ T &= T_I \end{aligned}$$

Step 2: Another solution x_2 is generated at random in the neighbor to x^{best} .

If $F(x_2) < F^{best}$ then $F^{best} \leftarrow F(x_2)$ and $x^{best} \leftarrow x_2$

If $F(x_2) > F^{best}$ then, consider R (a random number in $[0, 1]$):

$$\begin{aligned} \Delta &= F(x_2) - F^{best} \\ p &= \frac{1}{1 + \exp\left(\frac{-\Delta d}{T}\right)} \sigma^2 \end{aligned}$$

If $R < P$ then $x^{best} \leftarrow x_2$

Else change = false

Step 3: Check whether the stopping criteria are satisfied. If satisfied, end the SA algorithm; otherwise, change annealing temperature:

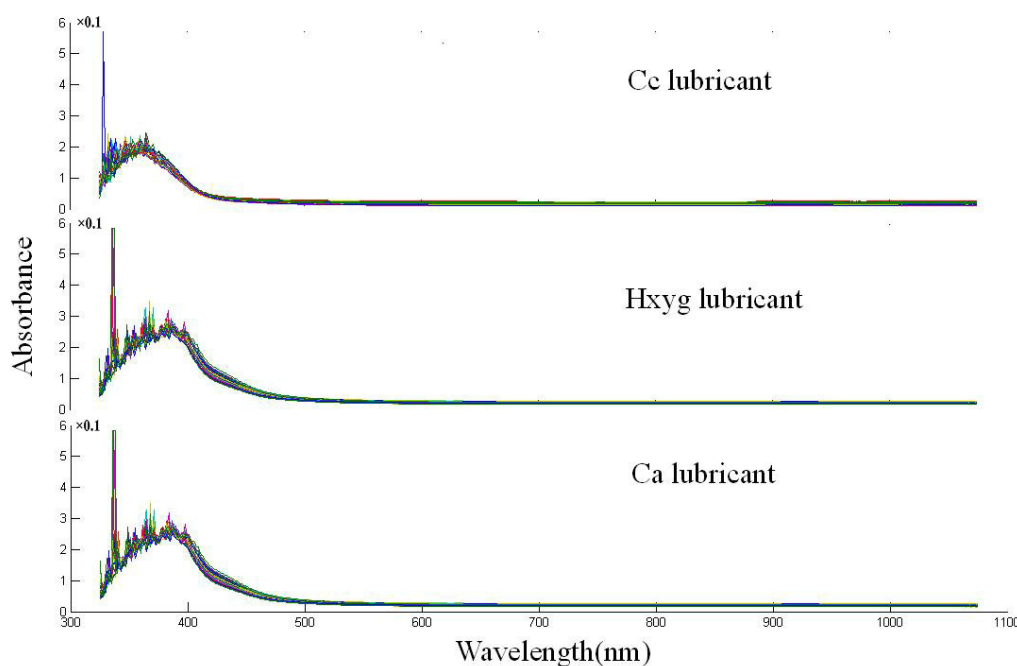
$$T = T * \tau$$

and back to step 2.

3. Results and Discussion

3.1. Absorbance Spectra of Three Varieties of Lubricants

Typical spectra of three varieties of lubricants are shown in Figure 2. Due to potential system imperfection, the scattering ray affects the accuracy of measurement; there is lots of noise in the 325–400 nm region. The spectra of samples from all three varieties have similar gross patterns of absorbance, there is a sharp absorption peak round 380 nm. After 400 nm, the absorbance values begin to decrease, and the spectral curve is flat. Though the trend of the spectra is similar, some latent differences and features exist according to the chemical components and color variance which could not be distinguished by the naked eye. With a closed observation, a small difference between Cc lubricant and the two other lubricants existed in the 400–450 nm region, which might have resulted from the color variance.

Figure 2. Vis-SWNIR spectra of three varieties of lubricant.

3.2. WPT Decomposition

Before SA was employed to search for BVB, the searching range of SA should be defined; the range includes lower and upper bound constraints of decomposition level; the max node number in each level. Lower and upper bound constraints of WPT decomposition should be determined first, upper bound, namely, for a signal with a length of N , the theoretically maximum decomposition level J , which could be calculated by the Matlab command of “wmaxlev”. wmaxlev is a one- or two-dimensional wavelet or wavelet packets-oriented function. wmaxlev can help avoid unreasonable maximum level values. wmaxlev gives the maximum allowed level decomposition, but in general, a smaller value is taken. After some performance comparison, the decomposition scale could be determined. The lower bound of decomposition was defined as zero, namely, the signal has not been decomposed. Then, constraint range of node number was determined according to decomposition level, because the WTP decomposed signal is the full binary tree (see Figure 1), the upper bound of node number was set as 2^l , where l is the decomposition level. The low bound number of the node was zero. In this work, the wavelet functions within the range of *Db1-Db9* were computed for a better performance, and the *Db4* wavelets function was employed as the optimal one after performance comparison and all the calculations were performed using MATLAB 7.6 (Mathworks, Natick, MA, USA).

3.3. SA Configuration Variables

3.3.1. Determination of SA Algorithm Parameters

After different bands were obtained by WPT decomposition, SA was employed to seek this optimal band in this study. Before using of SA, some parameters of SA algorithm should be preset:

- (1) Selection of initial temperature T_i , and termination temperature T_s : Initial temperature T_i is 100 °C, and termination temperature T_s is 0 °C.
- (2) State generate function and state accepting function: Student's t distribution was employed to generate new solutions in the SA algorithm. The random disturbance can be regarded as a jumping optimal model. State accepting function uses the change in function values between the current point and the new point to determine whether the new point is accepted or not.
- (3) Annealing schedule: Annealing schedule is an exponential annealing schedule which updates the current temperature based on the initial temperature and the current annealing parameter k (the number of evaluations of the objective function):

$$T_{n+1} = 0.95^k T_n \quad (5)$$

- (4) Algorithm termination criteria: In general, there are two stopping rules: one is that the number of temperature transitions satisfies the temperature termination rules, and the other one is that the neighbor solution was not improved after a certain period [21]. In our strategy, the algorithm stops when the average change on the value of the fitness function at current point is less than 1e-6 after 500 iterations.

3.3.2. Evaluation of Fitness Function

The performance of SA was evaluated through a fitness function, also known as objective function. The function value was the criterion for guiding SA to the global optimum. The prediction ability of the calibration model was evaluated with parameters of correlation coefficient (R) and root mean and root mean square error of calibration (RMSEC) or prediction (RMSEP), and the ideal model should have high R value, and low RMSEC and REMSEP values [22]. So the fitness function could be defined as follows:

$$\max f(X) = \frac{R}{1+RMSEC} = \frac{\sum_{i=1}^n (c_i^p - \bar{c}^p)(c_i^0 - \bar{c}^0)}{\left(1 + \left(\frac{1}{n} \sum_{i=1}^n (c_i^0 - \bar{c}^0)^2\right)^{\frac{1}{2}}\right) \cdot \left(\sum_{i=1}^n (c_i^p - \bar{c}^p)^2 \cdot \sum_{i=1}^n (c_i^0 - \bar{c}^0)^2\right)^{\frac{1}{2}}} \quad (6)$$

where c_i^0 is the actual concentration of spectrum i in the calibration set, \bar{c}^0 is the mean of c_i^0 , c_i^p is the concentration predicted by the model, \bar{c}^p is the mean of c_i^p , and n is the number of calibration set. The fitness function was measured in the form of leave-one-out cross-validation using partial least-square regression (PLSR), principal components (PCs) of PLSR according to accumulative reliabilities of PCs [23]. Thus the model was built using $n-1$ training samples and the one left out is used for prediction.

3.4. Optimization Results

After the parameters of WTP and SA were determined, SA was employed to search for the BVB. The optimal BVB and the best function value are shown in Figure 3, and the optimal node was (6,3). The best fitness function value of 0.8814 was obtained by SA.

Figure 4 shows the coefficients of node (6,3) of three varieties of lubricants. Compared with Figure 2, there was obviously variance among three varieties of lubricants, which could be discriminated by the naked eye. Moreover, it also could be found from Figure 4, that the trend of the curve of the same variety was very similar, except there was only an abnormal curve in the Cc lubricant.

Figure 3. Result of optimal node by SA. (a) Optimal node. (b) Best fitness function value.

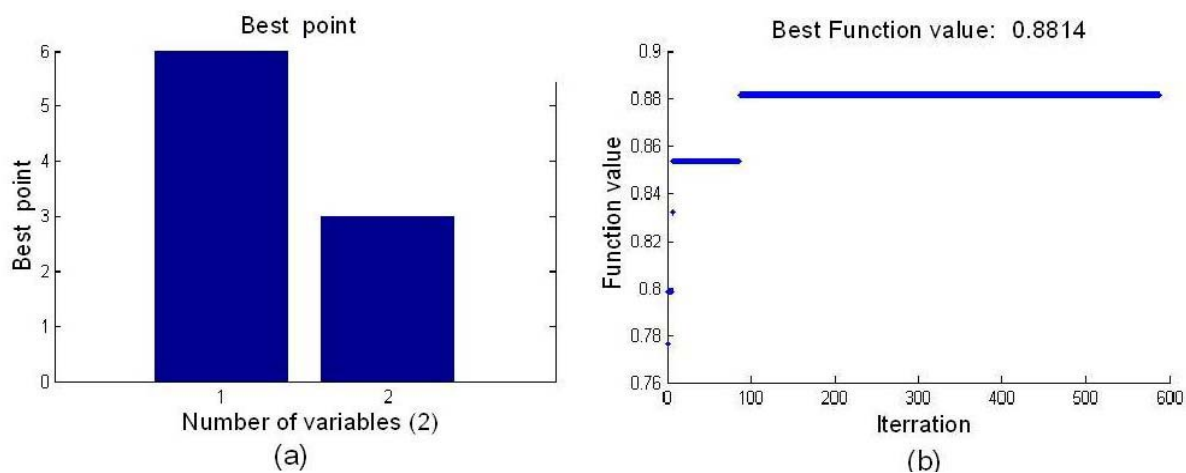
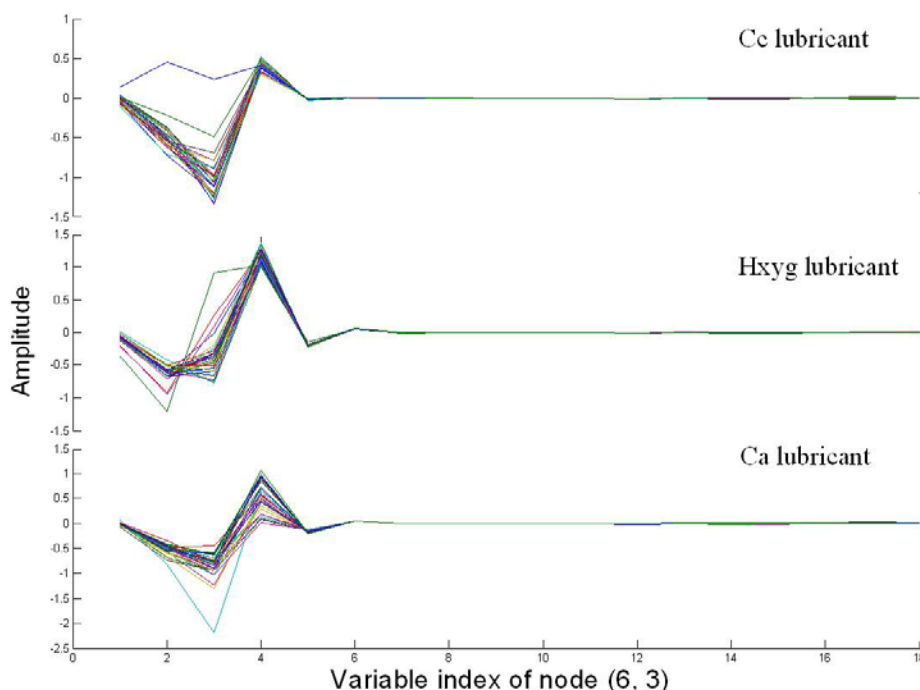


Figure 4. Characteristics of coefficients of node (6,3).

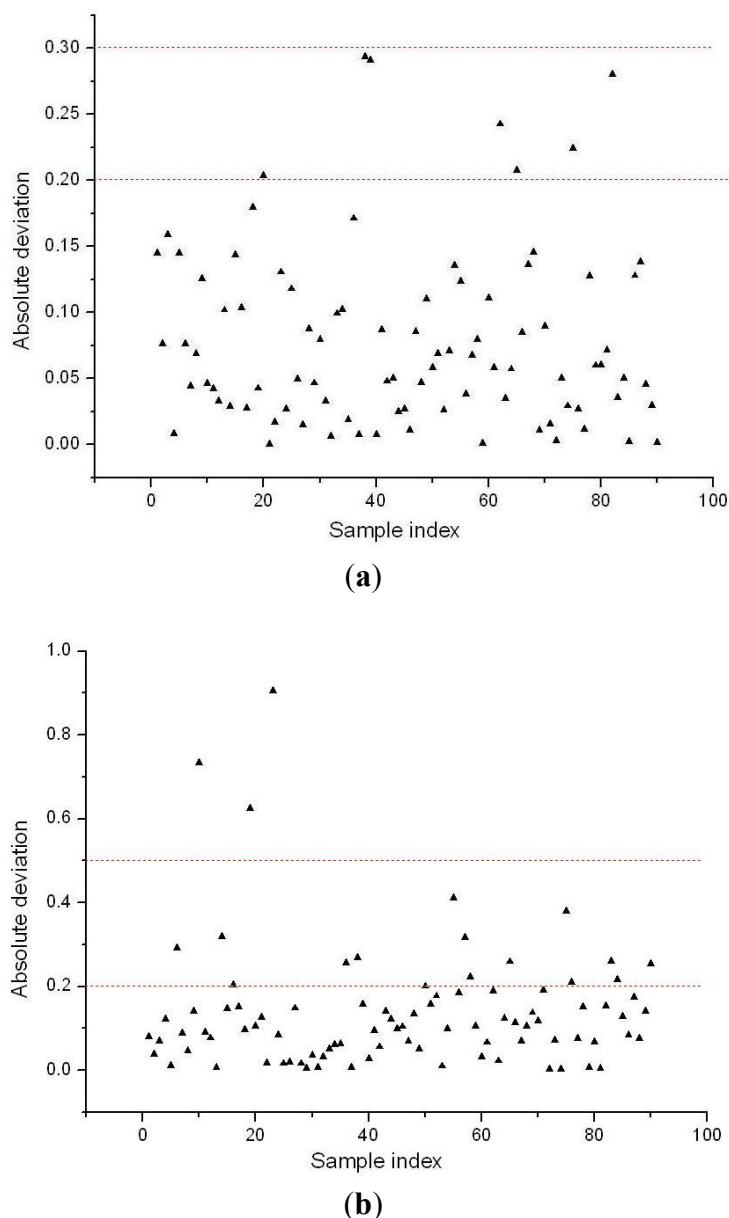


3.5. Development and Performance of LV-SVM

Support vector machine (SVM) is a state-of-the-art statistical learning method proposed by Vapnik [24], which has a good theoretical foundation in statistical learning theory based on the principal of structural risk minimum (SRM). LS-SVM can be considered as an extension of SVM [25], which is a new and attractive statistical learning method, and has the capability of dealing with linear and nonlinear multivariate calibration. When using LS-SVM, three crucial problems are required to be

solved, including the determination of the optimal input feature subset, proper kernel function, and the optimal kernel parameters [26]. Currently, there is no systematic methodology for selection of kernel functions. In our work, based on experience, a radial basis function (RBF) was adopted as the kernel function of LS-SVM, the regularization parameter γ within the region of $(2^{-1}, 2)$ and the parameter $sig2$ (σ^2) within the region of $(10^{-2}, 10^4)$ were set.

Figure 5. The absolute deviation value of prediction results of validation set (sample index 1–30 for Cc lubricant, 31–60 for Hxyg lubricant, and 61–90 for Ca lubricant). (a) BVB as the input set for LS-SVM. (b) Full spectra as the input set for LS-SVM.



Finally, BVB of 18 variables were extracted, and they constituted a new set of variables which could be used as inputs for LS-SVM (BVB-LS-SVM) to build the recognition model. In the application of BVB-LS-SVM, each variety of lubricant in the calibration set was assigned a dummy variable as a reference value (set Cc lubricant as 1, set Hxyg lubricant as 2, set Ca lubricant as 3).

A total of 90 samples in the validation set were predicted by the LS-SVM model. Good performance was achieved, and the prediction results for R and RMSEP were 0.9920 and 0.1042. The absolute deviation of prediction results of each sample is shown in Figure 5(a). The threshold error of recognition was set as ± 0.1 (the red dotted lines in the Figure 5). Taking Cc lubricant for example, the standard value for Cc is 1, and if the predicted value is lower than 0.9 or higher than 1.1, it will be marked as a wrong identification, so when the threshold error was ± 0.3 , the total discrimination rate was 100%, if the threshold error was set as ± 0.2 , the prediction results indicated only seven samples (No. 20 Cc, No. 38, 39, Hxyg, and No. 62, 65, 75, 82, Ca) were wrongly misclassified. Figure 5(b) shown the absolute deviation of prediction results of each sample when the full spectral variables were used as inputs for LS-SVM. When threshold error was set as ± 0.2 , the prediction rate was only 80%, even when the threshold error was set as ± 0.5 , three samples still exceeded this limit. Therefore, The BVB-LS-SVM model not only significantly improved prediction accuracy, but also compressed the raw spectral data greatly.

3.6. Elimination of Uninformative Variables in BVB

In order to improve prediction accuracy further, uninformative variable elimination by PLS (UVE-PLS) was employed to remove uninformative variables of BVB. In the UVE-PLS method, how to estimate the cutoff is a very important issue. In the previous research, they used artificial random variables, added to the data set to calculate the cutoff [11], and this way was experiential. In order to overcome this defect, in this work, SA was employed to search for the optimal cutoff of UVE-PLS; the fitness function of SA was the same as Equation (6). The number of principal components (PCs) for PLS was also according to the accumulative reliabilities of PCs, here, the number of PCs was six.

Figure 6. Result of optimal cutoff by SA. (a) Optimal cutoff. (b) Best function value.

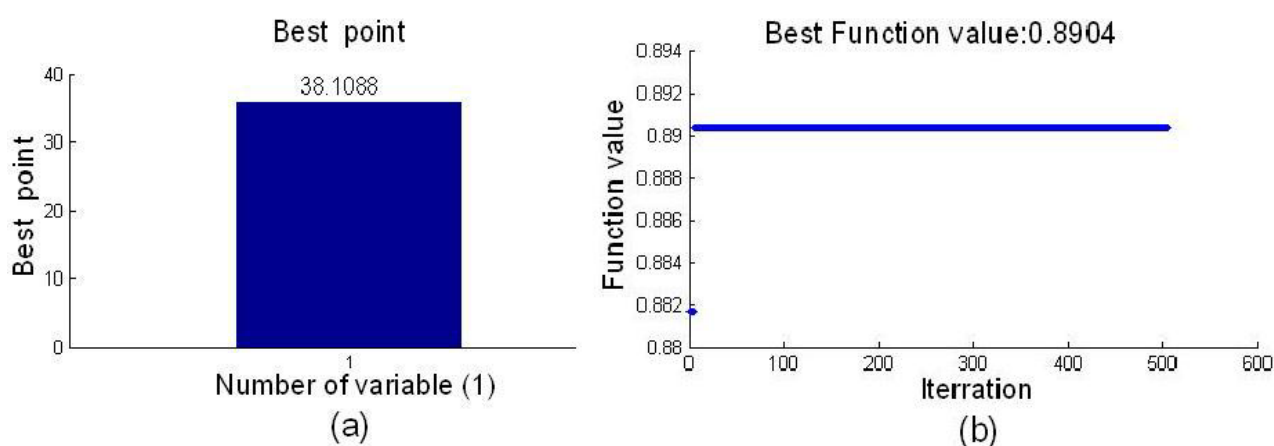


Figure 6 shows the result of the optimal cutoff and the best fitness function, and the optimal cutoff was 38.1088, and the best fitness function value of 0.8904 was obtained by SA. The stability of each variable in the BVB was shown in Figure 7. In the Figure 7, the dotted lines show the cutoff threshold, the vertical bar in Figure 7 indicated the stability range. Variables whose stability lies within the dotted lines were eliminated, and the variables whose stability lies out of the dotted lines were retained. Finally, only five variables in the BVB were retained.

The retained five variables constituted a new set of variables for LS-SVM (BVB-RV-LS-SVM). The performance of BVB-RV-LS-SVM was evaluated by the same 90 unknown samples mentioned above, and R of 0.9950 and RMSEP of 0.0829 were obtained, respectively. The absolute deviation of prediction results of each sample is shown in Figure 8. As one could see from Figure 8, when the threshold error was set as ± 0.3 , no sample was mistakenly discriminated, if the threshold error was ± 0.2 , there were only two misclassified samples, these were samples 62 and 65, and the total discrimination rate reached 97.78%. Therefore, compared with BVB-LS-SVM, the BVB-RV-LS-SVM model not only improved the prediction accuracy greatly, but also produced a more parsimonious model.

Figure 7. Stability distribution of each variable, and the two red dotted lines indicate the lower and upper cutoff.

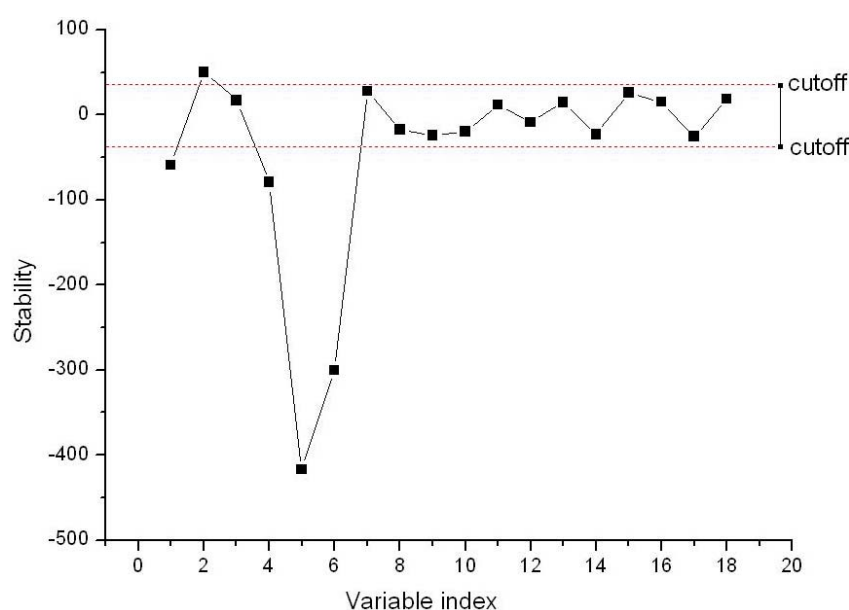
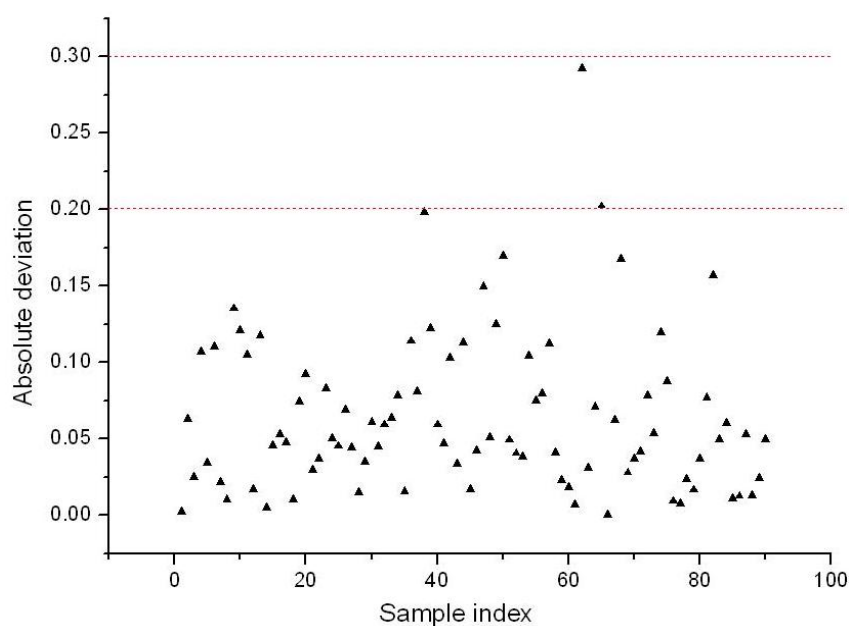


Figure 8. The absolute deviation value of prediction results of BVB-RV-LS-SVM.



3.7. Comparison of Different Calibration Models

In this section, the method proposed in this work was compared with the conventional methods:

Method I: LS-SVM model with certain LVs as the inputs [5]. Partial least squares discriminant analysis (PLS-DA) was also applied as a way to extract the Latent variables (LVs) which were determined by the lowest value of predicted residual error sum of squares (PRESS). Then, selected LVs were employed as the input sets of LS-SVM according to their explained variance.

Method II: Partial least squares regression (PLSR) [23].

Method III: LS-SVM with optimal wavelengths. SA was employed to search for the optimal wavelength in the full spectrum; this method is equivalent to a genetic algorithm (GA) searching for the optimal wavelength from the full spectrum.

Method IV: LS-SVM with certain wavelengths. Certain wavelengths were selected from the full spectrum by UVE-PLS, and optimal cutoff threshold of UVE-PLS was searched by SA.

Method V: The method proposed in this work (BVB-RV-LS-SVM).

Table 1. The discrimination results of calibration and validation sets by different calibration models.

Parameters	Method I	Method II	Method III	Method IV	Method V
PCs/LVs/Sw/Sv ^a	6	6	389	151	5
Calibration Set					
R^b	0.9845	0.9539	0.9864	0.9844	0.9943
RMSEC ^c	0.1433	0.2540	0.1351	0.1472	0.0878
Validation Set					
R^b	0.9511	0.9256	0.9844	0.9950	0.9950
RMSEP ^c	0.2718	0.3194	0.1472	0.0829	0.0827

^a PCs/LVs/Sw/Sv: Principle components/latent variables/Selected wavelengths/Selected variables.

^b R : correlation coefficient. ^c RMSEC/RMSEP: root mean square error of calibration or prediction.

Compared with all the models above, the model presented in this work achieved satisfactory prediction results (seen in Table 1). It was clear that using BVB-RV-LS-SVM, with only five variables, the prediction result could be slightly better than that of method IV using 151 variables (wavelengths). For the calibration set, the highest correlation coefficient R of 0.9943 and lowest RMSEC of 0.0878 were obtained by BVB-RV-LS-SVM. This result demonstrated that good best performance was achieved; the results also indicated that these five mined variables might be the best diagnostic information for the discrimination of varieties of lubricants. Furthermore, R for calibration and validation set, RMSEP and RESEC were very close, which indicated that almost no overfitting happened, and a good stability and generalization was achieved by this new model. Therefore, according to the result of calibration and validation set, the five variables as feature information could represent the main variance information, and these variables could be applied instead of the whole original spectra to discriminate the varieties of lubricants. Moreover, these mined variables could be used as diagnostic information for the development of lubricant variety discrimination apparatus.

4. Conclusions

The variety discrimination of lubricants was successfully performed by Vis-SWNIR spectroscopy with a hybrid method combination of WPT, UVE-PLS, SA and LS-SVM. Using the WPT, UVE_PLS and SA, the raw spectra data set was greatly compressed, and only five variables were mined. Then, these variables were used as input set for LS-SVM to build a recognition model, and good performance was achieved. The overall results indicated that the method combining WPT, UVE-PLS and SA was a powerful way of compressing the spectral data set and selection of diagnostic information.

Acknowledgments

This study was supported by 863 National High-Tech Research and Development Plan (2011AA100705) and Zhejiang Provincial Science & Technology Innovation Team Project (2009R50001).

References

1. Zhao, J.W.; Cheng, W.; Feng, Y.H. Determination of the content of wear metal in lubrication oil by ICP-AES. *Spectrosc. Spectr. Anal.* **2004**, *24*, 733-736.
2. Wei, H.J.; Yu, H.L.; Guan, D.L.; Sun, P.T.; Wang, H.Z. MOA Spectral analysis of additive element contents in lube oil. *Spectrosc. Spectr. Anal.* **2006**, *26*, 738-740.
3. Borin, A.; Poppi, R.J. Application of mid infrared spectroscopy and iPLS for the quantification of contaminants in lubricating oil. *Vibrat. Spectrosc.* **2005**, *37*, 27-32.
4. He, Y.; Feng, S.J.; Deng, X.F.; Li, X.L. Study on lossless discrimination of varieties of yogurt using the Visible/NIR-spectroscopy. *Food Res. Int.* **2006**, *39*, 645-650.
5. Liu, F.; He, Y.; Wang, L. Determination of effective wavelengths for discrimination of fruit vinegars using near infrared spectroscopy and multivariate analysis. *Anal. Chim. Acta* **2008**, *615*, 10-17.
6. Li, X.L.; He, Y.; Fang, H. Non-destructive discrimination of Chinese bayberry varieties using Vis/NIR spectroscopy. *J. Food Eng.* **2007**, *81*, 357-363.
7. Li, X.L.; He, Y. Discriminating varieties of tea plant based on Vis/NIR spectral characteristics and using artificial neural networks. *Biosyst. Eng.* **2008**, *99*, 313-321.
8. Shao, X.G.; Wang, G.; Chen, D.; Su, Q.D. A method for near-infrared spectral calibration of complex plant samples with wavelet transform and elimination of uninformative variables. *Anal. Bioanal. Chem.* **2004**, *378*, 1382-1387.
9. Mallat, S.B. Wavelet Packet and Local Cosine Base. In *A Wavelet Tour of Signal Processing*, 2nd ed.; China Machine Press: Beijing, China, 2006; Volume 8, Chapter 8, pp. 465-481.
10. Chen, D.; Shao, X.G.; Hu, B.; Su, Q.D. Simultaneous wavelength selection and outlier detection in multivariate regression of near-infrared spectra. *Anal. Sci.* **2005**, *21*, 161-166.
11. Center, V.; Massart, S.; de Noord, O.E. de Jong, S.; Vandeginste, B.M.; Sterna, C. Elimination of uninformative variables for multivariate calibration. *Anal. Chem.* **1996**, *68*, 3851-3858.
12. Put, R.; Heyden, Y.V. Review on modelling aspects in reversed-phase liquid chromatographic quantitative structure-retention relationships. *Anal. Chim. Acta* **2007**, *602*, 164-172.

13. Polanski, J.; Gieleciak, R. Comparative Molecular Surface Analysis (CoMSA) for modeling dye-fiber affinities of the Azo and Anthraquinone dyes. *J. Chem. Inform. Model.* **2003**, *43*, 1754-1762.
14. Kirkpatrick, S.; Gelatt, C.D., Jr.; Vecchi, M.P. Optimization by simulated annealing. *Science* **1983**, *220*, 671-680.
15. Hörchner, U.; Kalivas, J.H. Further investigation on a comparative study of simulated annealing and genetic algorithm for wavelength selection. *Anal. Chim. Acta* **1995**, *311*, 1-13.
16. Suman, B. Study of simulated annealing based algorithms for multiobjective optimization of a constrained problem. *Comput. Chem. Eng.* **2004**, *28*, 1849-1871.
17. Geng, X.T.; Xu, J.; Xiao, J.H.; Pan, L.Q. A simple simulated annealing algorithm for the maximum clique problem. *Inform. Sci.* **2007**, *177*, 5064-5071.
18. Salinas, S.V.; Chang, C.W.; Liew, S.C. Multiparameter retrieval of water optical properties from above-water remote-sensing reflectance using the simulated annealing algorithm. *Appl. Opt.* **2007**, *46*, 2727-2742.
19. Mallat, S. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Patt. Anal. Mach. Intell.* **1989**, *11*, 674-693.
20. Cocchi, M.; Seeber, R.; Ulrici, A. WPTER: Wavelet packet transform for efficient pattern recognition of signals. *Chemometr. Intell. Lab. Syst.* **2001**, *57*, 97-119.
21. Ghazanfari, M.; Alizadeh, S.; Fathian, M.; Koulouriotis, D.E. Comparing simulated annealing and genetic algorithm in learning FCM. *Appl. Math. Comput.* **2007**, *192*, 56-68.
22. Liu, F.; He, Y.; Wang, L. Comparison of calibrations for the determination of soluble solids content and pH of rice vinegars using visible and short-wave near infrared spectroscopy. *Anal. Chim. Acta* **2008**, *610*, 196-204.
23. Geladi, P.; Kowalski B.R. Partial least-squares regression: A tutorial. *Anal. Chim. Acta* **1986**, *185*, 1-17.
24. Vapnik, V.N. Controlling the Generalization Ability of Learning Process. In *The Nature of Statistical Learning Theory*, 2ed ed.; Springer-Verlag: New York, NY, USA, 1995; Volume 3, Chapter 4, pp. 99-101.
25. Suykens, J.A.K.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293-300.
26. Liu, F.; He, Y. Use of visible and near infrared spectroscopy and least squares-support vector machine to determine soluble solids content and pH of cola beverage. *J. Agr. Food Chem.* **2007**, *55*, 8883-8888.