Identification of automobile transmission fluid using hyperspectral imaging technology

Jiang Lulu[1], Yu Xinjie[2], He Yong[3*]

(1. *Zhejiang Technology Institute of Economy, Hangzhou 310018, China*;
2. *Ningbo Institute of Technology, Zhejiang University, Ningbo 315100, China*;
3. *College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China*)

Abstract: An identification method combining sparse representation with principal component analysis (PCA) was proposed for discriminating varieties of transmission fluid of automobile by using hyperspectral imaging technology. Principal component analysis was applied to obtain the characteristic information in the 874-1 733 nm spectra. For each transmission fluid variety, 80 samples were randomly selected as the training set, and 20 samples as the testing set. The eigenvectors of all training samples form the matrix were used for the sparse representation, and the problem of transmission fluid types classification was transformed into one to solve a sample expressed by the overall training sample matrix through optimization under the 11 norm. The results demonstrate that the accuracy of the algorithm that was composed of sparse representation and principal component analysis (PCA) was 93%. The accuracy is higher than those of PCA-LDA (Linear Discriminant Analysis) and PCA- LS-SVM (Least Squares Support Vector Machine). Therefore, the proposed method provides a better approach for the identification of transmission fluid types.

Keywords: transmission fluid, hyperspectral image, sparse representation, principal component analysis, identification

DOI: 10.3965/j.ijabe.20140704.009

Citation: Jiang L L, Yu X J, He Y. Identification of automobile transmission fluid using hyperspectral imaging technology. Int J Agric & Biol Eng, 2014; 7(4): 81－85.

1 Introduction

Auto gearbox is a very important part of a car, instead of using conventional gear oils or lubricants, it uses a dedicated automotive transmission fluid. Only the oil of prescribed specifications can be used, and other fluids are not allowed for supplement or substitute. Otherwise, the service life of the automatic transmission will be shortened[1]. There is a variety of transmission fluid in market. Users usually estimate the fluid quality by color, fluidity, bubbles, viscosity and smell. These methods are too subjective to accurately judge the quality. Another method is to check the state of auto gearbox after a type of transmission fluid is used, but this method harms the transmission. Therefore, a new non-destructive approach to identify the species of transmission fluid is in need. Hyperspectral imaging technology is a new type of non-destructive testing technology. Combining the benefits of computer vision and spectroscopic detection, it can record external features and internal quality better. Studies have shown that hyperspectral imaging technique has been widely used in non-destructive testing of agricultural and livestock products as well as fruits. Taking advantage of hyperspectral image, researches have been performed on, for example, apple's soluble solid content and hardness determination, weight prediction model and minor

Received date: 2013-09-08 **Accepted date:** 2014-07-10

Bibliorgaphies: Jiang Lulu, PhD, Professor, research interests: rapid detection technique of automobile based on spectroscopy. Email: abbyu111@sina.com. **Yu Xinjie,** Associate Professor, research interests: nondestructive inspection of agricultural products. Email: xjyu1979@163.com.

*** Corresponding author: He Yong,** PhD, Professor, research interests: nondestructive inspection of agricultural products. Mailing Address: 866 Yuhangtang Road, Hangzhou, Zhejiang 310058, P. R. China. Tel/Fax: +86-571-88982143. Email: yhe@zju.edu.cn.

damage detection, snowflake's characteristics, citrus' rust feature, rapeseed classification[2-6], and long jujube's nondestructive detection of sugar content[7]. Because of a large number of mechanical devices applied in agriculture, transmission fluid is very important to agricultural production. Sparse representation theory is also known as compressed sensing theory. It is an optimization method based on l1 norm minimization. In recent years, it has been widely used in the field of pattern recognition[15]. Sparse representation is an important outcome of the research in the field of pattern recognition in recent years. Many researches were done using sparse representation such as rice varieties identification[8], face recognition[9], and vehicle recognition in the traffic scene[10].

In this paper, we used principal component analysis[11] to extract the inherent spectral information, and then used sparse representation to classify transmission fluids.

2 Materials and methods

2.1 Instrument and equipment

The schematic of the hyperspectral imaging system provided by Isuzu Optics company of Taiwan is shown in Figure 1. The system mainly consisted of a light source unit Schott DCR Ⅲ, a hyperspectral imaging spectrometer V10E-QE, an OLE-23 lens, an electronically controlled displacement desk PSA200-11-X, a controller SC300A, a computer, camera obscura and other components. Schott DCR III is a halogen line light source of 150W DC power and 50Hz of light flashing AC output frequency. The spectral range of the hyperspectral measurement is 874-1 733 nm, a total of 256 bands.

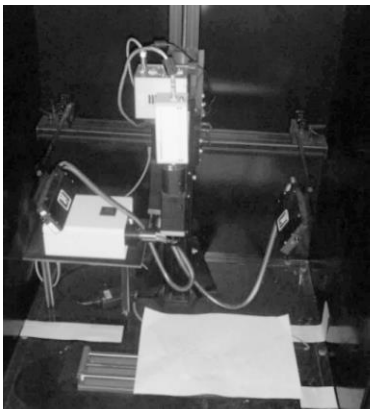


Figure 1 Hyperspectral imaging system

2.2 Materials and hyperspectral data acquisition

The five transmission fluids purchased from the market were: Shell ATFIII of Zhejiang Shell Chemical Petroleum Limited, Mobil ATF220 of ExxonMobil, Castrol ATF-MV of Castrol, Shanghai Volkswagen G055025A2 of Shanghai Volkswagen Automotive Co., Ltd, Shanghai GM DEXTRON III of Shanghai General Motors Co. Ltd. Six different production batches and dates for each transmission fluid were used. For each transmission fluid, 40 samples were randomly selected. Samples were held in transparent containers with a diameter of 60 mm and height of 14 mm.

Hyperspectral data acquisition was based on spectral cube software (Spectral Imaging Ltd., Finland). To acquire clear image without distortion, the exposure time was set to 30 ms and the moving speed of the electronically controlled displacement desk was set to 1.58 mm/s before hyperspectral image acquisition. Eight sample containers of the same variety were placed on the mobile station at a time to acquire the hyperspectral image data.

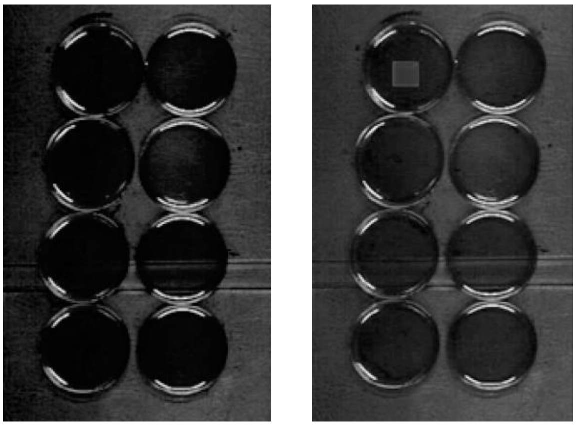
2.3 Pretreatment of hyperspectral data

In the collection process, the intensity distribution of the light source in each band was uneven and there was dark current noise, which signaled that the collected images needed to be corrected. The correction was calculated as follows:

$$R = \frac{I - B}{W - B} \tag{1}$$

where, B is a full black calibration image; W is the full white calibration image; I is the collected image, and R is the relative image.

The corrected hyperspectral image is shown in Figure 2a. The area of interest shown in Figure 2b (red region) was extracted from the corrected hyperspectral images. The size was 20×20, and the average reflectance value of this region was calculated. From the total of 100 transmission fluid samples, 80 were randomly selected as the training sample set, and 20 as the testing set. After removing the noises, the spectral reflectance values between 950 and 1 725 nm (a total of 230 bands) were selected. Figure 3 is the original spectral curve of a part of the samples. Because the original values of the spectrum were impacted by scattering and offset, we linearized the original values of the spectrum using multiplicative scatter correction (MSC). Figure 4 is the spectral curve of a part of the original samples after MSC. The hyperspectral data analysis software ENVI4.8 (ITT, USA) was used. The hyperspectral data pre-processing software was Unscramble v9.7 (CAMO Software, Inc., Woodbridge, NJ).



(a) (b)

Figure 2 (a) (b) Pre-processing of hyperspectral image

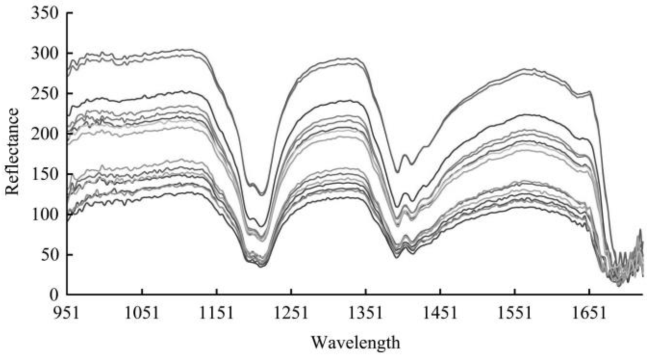


Figure 3 Reflectance curves of the original spectrum

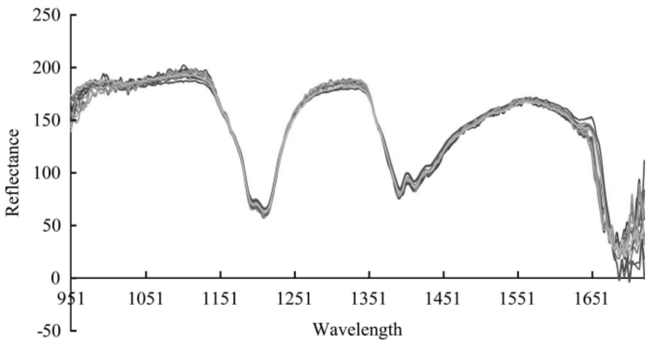


Figure 4 Reflectance curves of the spectrum after MSC processing

2.4 Sparse representation method

There were five kinds of transmission fluids to be classified, each class having 80 training samples. The training samples can be expressed as:

$$A=[f_1^1 f_2^1 \cdots f_{80}^1 \cdots f_1^5 \cdots f_{80}^5]=[A^1 A^2 A^3 A^4 A^5] \tag{2}$$

where, A^i is the vector space of the i-th. The number of rows of the matrix A represents the number of sample feature parameters (10 principal components) and the number of columns is the total number of training samples.

The vector y from the object class of the k-th will be represented by the linear formula consisting of the training matrix A:

$$y=Ax=A[x_1^1 x_2^1 \cdots x_{80}^1 x_1^2 \cdots x_{80}^2 \cdots x_1^5 \cdots x_{80}^5]' \tag{3}$$

Where the element x^i_j is the projection coefficient that y projected on the j-th training which belong to the i-th of class A.

Because y belongs to the i-th sample, the projection coefficients are zero except for those associated with the i-th class. Actually, the class of testing vector y is unknown. Base on the theory of sparse representation, if the column vector x is spare enough, formula (2) can be expressed as:

$$\hat{x}=\arg\min\|x\|_1 \quad subiect\ \ to\ \ Ax=y \tag{4}$$

Hyperspectral data acquisition process includes noise, then it is difficult for y to accurately represent A in linear, so the above expression can be rewritten as:

$$\hat{x}=\arg\min\|x\|_1 \ subject\ \ to\ \|Ax-y\|_2 \le \varepsilon \tag{5}$$

where, $\hat{x}$ is an approximate solution of x, and ε is an error threshold.

For each class i, δ_i is a relevant coefficient vector between $\hat{x}$ and the i-th, therefore, we can use $\hat{y}_i$ to approximate y if y belongs to the i-th. The smaller the distance between y and $\hat{y}_i$ is, the higher the reliability is that $\hat{y}_i$ belongs to the i-th. Therefore, the formula for identifying category can be defined as follows:

$$\min_i r_i(y)=\min_i(\|y-A\delta_i(\hat{x})\|_2) \tag{6}$$

Get the smallest residual $r_i(y)$, then y can be recognized as the i-th class.

$$l(y)=\arg\min_i r_i(y) \tag{7}$$

where, $l(y)$ is the symbol of y [12].

3 Results and discussion

The raw hyperspectral data of transmission fluid were processed as follows: (1) we used the PCA dimensionality reduction algorithm by MATLAB2009a and extracted a total of 12 principal components; (2) we used the classification algorithm based on sparse representation by MATLAB2009a, and the l1 norm minimization method by the MATLAB package written by Koh et al. [13]; (3) we used LDA by MATLAB2009a; and (4) we used LS-SVM by the MATLAB package written by Suykens et al. [14].

The PC specification is shown as follows: OS, Windows 7 32 bit; CPU, Intel Core i3 M 380 2.53 GHz; Memory, 2 GB.

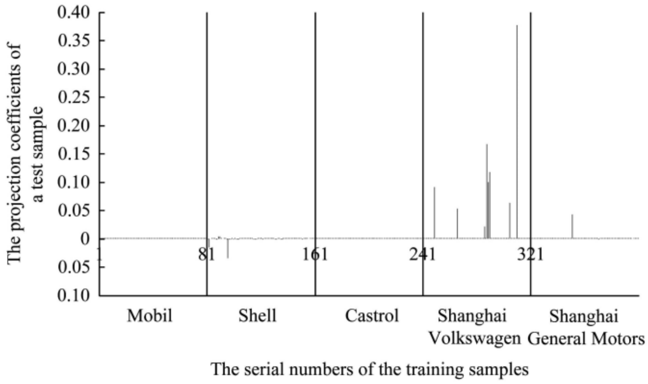
Principal component analysis (PCA) was used to reduce the dimensionality of hyperspectral data. A total of 12 principal components were extracted whose contribution rate reached 100%. The 12 principal components therefore were adequate to express the hyperspectral data of transmission fluids.

The following is an instance based on sparse representation classification algorithm proposed in the paper. In this example, the numbering sequence of the 400 training samples is shown in Table 1. A test sample of the Shanghai Volkswagen brand was randomly selected as *y*. Sparse representation was used to identify the category of *y*.

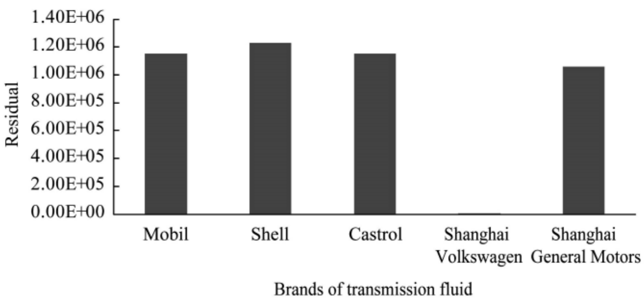
Table 1 Sample code in training sample matrix A

Sample Type	Mobil (the first class)	Shell (the second class)	Castrol (the third class)	Shanghai Volkswagen (the forth class)	Shanghai General Motors (the fifth class)
number	1-80	81-160	161-240	241-320	321-400

As shown in Figure 5a, the coefficients that the test samples y projected to the range of Shanghai Volkswagen were the highest while they were smaller or zero in other classes. Formulas 6 and 7 were used to calculate the residual of estimated and actual values. The residuals shown in Figure 5b were the smallest in the range of Shanghai Volkswagen. Therefore, the test sample y was correctly classified as Shanghai Volkswagen.



a. The projection coefficients of a test sample



b. The result of the residual

Figure 5 (a) (b) Example of sparse representation classification results

Table 2 shows the average accuracy of the classification of the test samples of the five transmission fluids using PCA-LDA, PCA-LS-SVM and PCA-SR, which were 87%, 86% and 93%, respectively. The classification accuracy of SR was significantly higher than those of the other two classification methods. This is because LDA and LS-SVM are based on the Euclidean distance between the training and test samples to distinguish the types of the testing samples. In SR, the sparse characteristics of the transmission fluids' data are calculated, thus the obtained residual is smaller, thereby improves the rate of correct classification.

Table 2 Using PCA to reduce dimensionality and three classification methods

Variety	Classification Results					
	PCA-LDA		PCA-LS-SVM		PCA-SR	
	Number of correct classification	Precision rate/%	Number of correct classification	Precision rate/%	Number of correct classification	Precision rate/%
Mobil	17	85	18	90	18	90
Shell	18	90	14	70	17	85
Castrol	18	90	19	95	18	90
Shanghai Volkswagen	19	95	15	75	20	100
Shanghai GM	15	75	20	100	20	100
Average		87		86		93

4 Conclusions

An identification method combined with principal component analysis and sparse representation - PCA-SR is proposed in this paper. Principal component analysis was first used to extract characteristic variables from hyperspectral data, and SR was then used for classification. The results demonstrated that the average identification accuracy of the proposed method was 93%, which was significantly higher than those of other methods. Therefore, the proposed method provides a promising way for effective identification of transmission fluid varieties.

Acknowledgments

The authors appreciate the financial support to the study by 863 National High-Tech Research and Development Plan (Project No: 2013AA102301), National Natural Science Foundation of China (Project No: 31201446), Zhejiang Provincial Science & Technology Innovation Team Project and Ningbo Natural Science Foundation of China (Project No: 201301A6101002).

[References]

[1] Yao M. Identification of the automatic transmission fluid replacement. Journal of Car Driver, 2004.

[2] Zou W, Fang H, Liu F, Zhou K Y, Bao Y D, He Y. Identification of rapeseed varieties based on hyperspectral imagery. Journal of Zhejiang University (Agriculture & Life Sciences), 2011; 37(2): 175–180. (In Chinese with English abstract)

[3] Cheng G S, Guo J X, Shi Z, Amuti R, Kang Y X. Prediction of the Weight of Xinjiang Fuji Apple by Hyperspectral Imaging Techniques. Journal of Xinjiang Agricultural University, 2011; 34(3): 249–252. (In Chinese with English abstract)

[4] Cai J R, Wang J H, Chen Q S, Zhao J W. Detection of rust in citrus by hyperspectral imaging technology and band ratio algorithm. Transaction of the CSAE, 2009; 25(1): 127–131. (In Chinese with English abstract)

[5] Hong T S, Qiao J, Wang N, Ngadi, M O, Zhao Z X, Li Z. Non-destructive inspection of Chinese pear quality based on hyperspectral imaging technique. Transactions of the CSAE, 2007; 23(2): 151–155. (In Chinese with English abstract)

[6] Xing J, Baerdemaeker J D. Bruise detection on 'Jonagold' apples using hyperspectral imaging. Postharvest Biology and Technology, 2005; 37: 152- 162.

[7] Xu S, He J G, Yi D, He X G. Nondestructive detection of sugar content in long jujude based on hyperspectral imaging technique. Food and Machinery, 2012; 28(6): 168–170. (In Chinese with English abstract)

[8] Yang S Q, Ning J F, He D J. Identification of varieties of rice based on sparse representation. Transaction of the CSAE, 2011; 27(3): 191–195. (In Chinese with English abstract)

[9] Wright J, Yang A Y, Ganesh A, Sastry S S, Ma Y. Robust face recognition via sparse representation. IEEE T. Pattern Anal, 2009; 31(2): 210-227.

[10] Li X Z, Wu J, Cui Z M, Chen J M. Sparse representation method of vehicle recognition in complex traffic scenes. Journal of Image and Graphics, 2012; 17(3): 387–392.

[11] Absdi, H. Williams L J. Principal component analysis. WIRES Computational Statistics, 2010; 2: 433–459.

[12] Yang S Q, Ning J F, He D J. Identification of varieties of rice based on sparse representation. Transaction of the CSAE, 2011; 27(3): 191–195. (In Chinese with English abstract)

[13] Koh K, Kim S J, Boyd S. Simple MATLAB solver for l1-regularized least squares problems. http://www.stanford.edu/~ boyd/l1_ls/ Accessed on [2008-05-15].

[14] Pelckmans K, Suykens J A K, Gestel T V, Brabanter J D, Lukas L, Hamers B, et al. Least squares-support vector machines. http://www.esat.kuleuven.be/sista/lssvmlab/ Accessed on [2011-08-16].

[15] Bruckstein A, Donoho D, Elad M. From sparse solutions of systems of equations to sparse modeling of signals and images. SIAM Review, 2009; 51(1): 34–81.