Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 151 (2015) 34–39

Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

journal homepage: www.elsevier.com/locate/saa

Detection of adulteration in acetonitrile using near infrared spectroscopy coupled with pattern recognition techniques

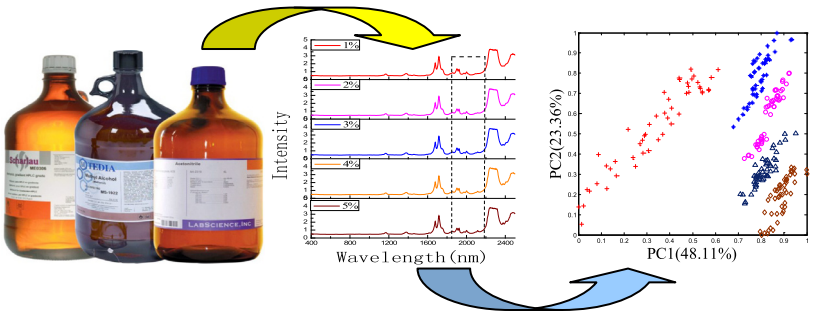
Le-Qian Hu *, Chun-Ling Yin, Zhi-Peng Zeng

School of Chemistry and Chemical Engineering, Henan University of Technology, Zhengzhou 450001, China

HIGHLIGHTS

- NIR and chemometrics were used to determinate the adulteration of acetonitrile.
- NIR spectroscopy data were subjected to analysis by PCA and SVM model.
- SVM has more dominance in determining the impurities of acetonitrile than PCA.
- Acetonitrile with 1% impurities can be recognized by NIR coupled with chemometrics.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 19 January 2015
Received in revised form 19 June 2015
Accepted 22 June 2015
Available online 23 June 2015

Keywords:
Adulteration
Acetonitrile
NIR
Support vector machine
Principal component analysis

ABSTRACT

In this paper, near infrared spectroscopy (NIR) in cooperation with the pattern recognition techniques were used to determine the type of neat acetonitrile and the adulteration in acetonitrile. NIR spectra were collected between 400 nm and 2498 nm. The experimental data were first subjected to analysis of principal component analysis (PCA) to reveal significant differences and potential patterns between samples. Then support vector machine (SVM) were applied to develop classification models and the best parameter combination was selected by grid search. Under the best parameter combination, the classification accuracy rates of three types of neat acetonitrile reached 87.5%, and 100% for the adulteration with different concentration levels. The results showed that NIR spectroscopy combined with SVM could be utilized for determining the potential adulterants including water, ethanol, isopropyl alcohol, acrylonitrile, methanol, and by-products associated with the production of acetonitrile.

1. Introduction

Acetonitrile is one of the most popular organic solvent [1–2]. With it increasing widespread use in industry, especially in pharmaceutical, chromatography solvents and DNA testing, people pay more attention about the quality of acetonitrile. Since the production of acetonitrile is as a by-product from the manufacture of acrylonitrile, the crude acetonitrile has some impurities, such as acrylonitrile, hydrocyanic acid, alcohol, methanol, isopropanol and water. Although there are many advanced high-purity refined acetonitrile process, it is difficult to remove all the impurity in commodity acetonitrile. And the shortage of acetonitrile has driven its price up and thus the adulteration may provide financial benefit for those who practice it [3]. However, few reports concerned about the issue of how to detect the impurities in acetonitrile. For this reason, a feasible method of this issue was presented in this paper.

Near infrared spectroscopy is a fast, easy and non-destructive technique that it is popular used for quality control testing [4–6]. Chen et al. has developed a method to distinguish the impure acetonitrile with near infrared spectroscopy [3]. Acetonitrile and

* Corresponding author.
E-mail address: leqianhu@yeah.net (L.-Q. Hu).

http://dx.doi.org/10.1016/j.saa.2015.06.067
1386-1425/

each of its impurities, such as water, acrylonitrile, alcohol, methanol and isopropanol, could be distinguished according to the spectral feature without using chemometrics. The impurities might be identified in acetonitrile with visual evaluation of the differences between NIR spectra. This approach might be improper when the spectral differences are subtle. Allen et al. and Harrison et al. also applied NIR and mid-infrared technology to study the characteristic of acetonitrile [7–8] in a similar manner.

In this research, we propose the approach relied on the application of support vector machine (SVM), known as a reliable classification technique based on statistical learning theory and Vapnik–Chervonenkis dimension [9–10]. The support vector machine conceptually implements the following idea: input vectors are non-linearly mapped to a very high-dimension feature space. In this feature space a linear decision surface is constructed. Special properties of the decision surface ensure high generalization ability of the learning machine [9]. Recently, SVM has been successfully used in various fields such as, identification of teas [5], classification of gasoline [11], classification of diesel pool streams [12] and feature selection [13]. Therefore, this paper aims at studying the capabilities of SVM to derive accurate and robust calibration models for predicting the types and amount of impurities in acetonitrile from NIR spectra data. As a comparison, unsupervised pattern recognition methods, principal component analysis (PCA) was also used to classify the data set in this research.

2. Materials and experiments

2.1. Materials

Three types of neat acetonitrile in this experiment were obtained from Yuwang (HPLC grade), Kermel (AR grade) and XiLong (AR grade) Chemical Group, respectively. Five types of adulteration chemicals in this experiment are ethanol, isopropyl alcohol, acrylonitrile, methanol and water. Ethanol (AR grade) and isopropyl alcohol (AR grade) was obtained from Tianli Chemical. Acrylonitrile (AR grade) was obtained from FuChen Chemical group. Methanol (HPLC grade) was from shield. All the water used in this experiment was double distilled.

2.2. Sample preparation and detection of neat acetonitrile

In total, 150 neat acetonitrile samples sourced from YuWang ($n = 50$), Kermel ($n = 50$), XiLong ($n = 50$) were used in this research. All the samples were measured using a FOSS XDS NIR Rapid Content Analyzer with standard sample cell. The temperature and humidity were kept a steady level in laboratory. The NIR spectra were collected in the range from 400 nm to 2498 nm and the data were measured in 2 nm intervals, which resulted in 1050 variables.

2.3. Sample preparation and detection of adulterated acetonitrile

Five types of adulteration chemicals, ethanol, isopropyl alcohol, acrylonitrile, methanol and water, were added to the YuWang acetonitrile to prepare the adulterated acetonitrile samples, respectively. Each foreign substance with volume percentages of 1%, 2%, 3%, 4% and 5% were blended into 50 different production batches of YuWang acetonitrile samples. Then each concentration level would produce 50 adulteration samples. Total 250 YuWang adulteration acetonitrile samples were constructed. Two other neat acetonitrile, Kermel and XiLong Chemical were also spiked at levels in the range of 1–5% of each adulterant with the same manner. The collection method of NIR spectra was same as neat acetonitrile.

3. Methods

3.1. Basic theory of principal component analysis (PCA)

PCA is a method which transforms the original data matrix into one composed of orthogonal variables called principal components (PCs). Each PC is a linear combination of the original data, and there are as many PCs extracted from the data matrix as there are original variables. In the same time, each PC accounts for consecutively decreasing the amount of data variance, which results in the compression of significant data into just a few PC variables [15]. PCA is a way of identifying patterns in the data and expressing them in such a way that similarities and differences can be seen, reducing the dimensionality without losing too much information.

Mathematically, each PC can be described as a linear combination of the original variables where the importance of each of these variables is given by the loadings. This linear combination also provides values for each object, called scores [12]. The data set **X** ($n \times p$), consisting of n objects and p variables is decomposed into the product of two smaller matrices:

$$X_{n\times p} = U_{n\times p}V^{T}_{q\times p} + E_{n\times p}$$

where **U** is a matrix of object scores of the principal components, **V** is the matrix of loadings (weights) of the variables on the principal components, and **E** is the matrix of the residuals (noise). The superscript **T** indicates the transpose of the matrix, q is a scalar indicating the number of principal components that describe most of the data variance. In this study, the objects are represented by the samples number and variables are represented by the absorbance intensity of the specific wavelength of NIR spectrum.

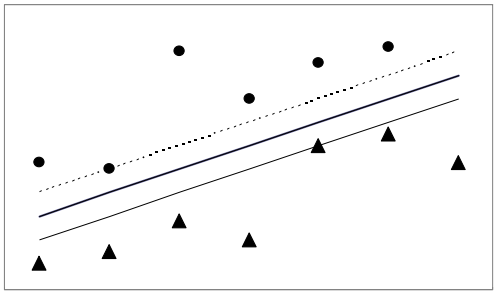


Fig. 1. The solid line is the optimal hyperplane.

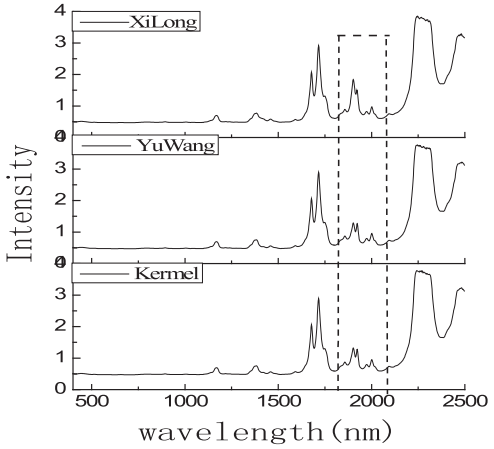


Fig. 2. The NIR spectra comparison of three types of acetonitrile.