Chemometrics and Intelligent Laboratory Systems 140 (2015) 141–148

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Application of Fourier transform to preprocessing chromatographic fingerprints of traditional Chinese medicine

Long Jiao [a,b,*], Zhiwei Xue [c], Gangfeng Wang [c], Xiaofei Wang [a], Hua Li [b]

[a] College of Chemistry and Chemical Engineering, Xi'an Shiyou University, Xi'an 710065, China
[b] College of Chemistry and Materials Science, Northwest University, Xi'an 710069, China
[c] No. 203 Research Institute of Nuclear Industry, Xianyang 712000, China

ARTICLE INFO

Article history:
Received 19 December 2013
Received in revised form 13 June 2014
Accepted 15 November 2014
Available online 22 November 2014

Keywords:
Chromatographic fingerprints
Retention time shift
Fourier transform
Correlation optimized warping
Principal components analysis

ABSTRACT

Retention time shift is a significant impediment against analyzing chromatographic fingerprints of Traditional Chinese Medicine (TCM) with similarity estimation and pattern recognition methods. This work aims at developing a fingerprint preprocessing method that can eliminate the negative effect of retention time shift on similarity estimation and pattern recognition analysis in the quality evaluation of TCM. Both simulated chromatograms and experimental chromatographic fingerprints of *Radix Puerariae thomsonii* (RPT) were investigated in order to evaluate the practicality of the proposed method. Fourier transform (FT) was used to process these chromatographic signals. The magnitude data of these signals were calculated. In the obtained magnitude data, the negative effect of retention time shift was successfully excluded. Thus, the quality of the investigated RPT samples can be correctly evaluated by using the magnitude data as the input variables of similarity estimation and pattern recognition analysis. It is demonstrated that FT is a practicable and promising method for preprocessing chromatographic fingerprints of TCM prior to similarity estimation and pattern recognition analysis.

1. Introduction

Chromatographic fingerprint has been commonly applied to the quality control of Traditional Chinese Medicine (TCM) [1,2]. It is generally believed that regarding the chromatogram of TCM as its fingerprints is reasonable and practicable because the chromatogram can reflect the "chemical integrities" of TCM [1,3–5]. The quality evaluation of TCM is usually achieved by analyzing its chromatographic fingerprints with chemometric methods, such as similarity estimation, principal component analysis (PCA) and cluster analysis [1,4–13]. However, the retention time shift, which is inevitable in the chromatography [4,5], is a significant impediment against analyzing chromatographic fingerprints of TCM with similarity estimation and pattern recognition methods [11]. When retention time shift occurs, it is difficult to get correct evaluation results [4]. Consequently, it is necessary to overcome the problem of retention time shift prior to evaluating the quality of TCM according to its fingerprint [14–16].

Several warping methods have been proposed and applied to the correction of retention time shift in chromatographic fingerprints of TCM [5,17–19]. However, chromatographic profiles will be inevitably changed in the warping procedure. When using warping methods to align the chromatographic fingerprints, there is the risk of unreasonably changing the chemical information in the fingerprints. If the fingerprints are unreasonably warped, the quality of TCM might be incorrectly evaluated. Thus, it is worthy to develop a method which will not change the chromatographic profiles while solving the problem of retention time shift.

Fourier transform (FT) is a mathematical transformation employed to transform signals between time (or spatial) domain and frequency domain, which has many applications in physics, chemistry and engineering [20–22]. In the FT procedure, no artificial information is introduced into the processed signal. Thus, the chemical information in the fingerprints will not be changed when using FT to process the chromatographic fingerprints of TCM. This is the significant advantage of FT in contrast to warping methods. Therefore, a method based on FT is proposed in this work for solving the problem of retention time shift in the fingerprint analysis of TCM. This method is a fingerprint preprocessing method which can be used to eliminate the negative effect of retention time shift on similarity estimation and pattern recognition analysis in the quality evaluation. Simulated chromatograms and the chromatographic fingerprints of *Radix Puerariae thomsonii* (RPT) were investigated by using the method.

2. Theory

2.1. Fourier transform

Fourier transform [20–22] is a conventional signal processing method. It has been widely used in analytical chemistry [20]. The Fourier

* Corresponding author. Tel.: +86 29 88383172.
E-mail address: mop@xsyu.edu.cn (L. Jiao).

http://dx.doi.org/10.1016/j.chemolab.2014.11.004
0169-7439/

transform, which breaks down a signal into constituent sinusoids of different frequencies, defines a relationship between a signal in the time domain and its representation in the frequency domain. The transformed frequency domain data only correlate with the profiles of the original time domain signal. It is shift-invariant for time domain signal and, hence, can be used to solve the problem of retention time shift.

The Fourier transform is defined as

$$X(f) = \int_{-\infty}^{\infty} s(t)e^{-jft}dt \tag{1}$$

where f is frequency variables, $X(f)$ is usually called the frequency domain representation of the original signal $s(t)$ and $X(f)$ is a complex-valued function that is periodic in f with period 2π. In Matlab, a discrete Fourier transform algorithm [22] is provided to conduct Fourier transform:

$$X(f) = \sum_{t=1}^{N} s(t)\omega_N^{(t-1)(f-1)}, \quad \omega_N = e^{(-2\pi i)/N} \tag{2}$$

The output of Fourier transform is a complex number, which cannot be used as input variables of similarity estimation and pattern recognition analysis. Thus, the magnitude of transformed data was calculated and used as input variables for subsequent similarity estimation and pattern recognition analysis.

2.2. Principal component analysis

PCA [22,23] is a technique that can reduce the dimensionality of a data matrix by extracting the relevant information from the original data into fewer new variables, called principal components. It is a conventional method for unsupervised classification requiring no priori knowledge of a data set. The PCA can be described as

$$\boldsymbol{X} = \boldsymbol{S}\boldsymbol{L}^T + \boldsymbol{E} \tag{3}$$

Table 1
Details about *Radix Pueraria Thomsonii* samples.

Samples	Products (group)	Plant origins (province)	Pharmaceutical stores
A1 A2 A3	A	Sichuan(1)	Xi'an Po Chi Tang Chinese Medicine supermarket, Xi'an
B1 B2 B3	B	Sichuan(2)	Lanzhou Hui Ren Tang pharmaceutical company, Lanzhou
C1 C2 C3	C	Anhui	Xi'an Yi Kang pharmaceutical supermarket, Xi'an
D1 D2 D3	D	Shanxi	Jiangsu Pharmaceutical company, Nanjing
E1 E2 E3	E	Guangdong	Guangzhou Guangming ginseng antler company, Guangzhou
F1 F2 F3	F	Zhejiang	Xi'an traditional Chinese medicine company, Xi'an
G1 G2 G3	G	Henan	Zhengzhou Tong Ren Pharmaceutical store,
H1 H2 H3	H	Jiangsu	Pharmacy of Bengbu hospital, Bengbu
I1 I2 I3	I	Jiangxi	Xi'an civil pharmaceutical store, Xi'an

in which $\boldsymbol{X}$ is the original data matrix, $\boldsymbol{S}$ and $\boldsymbol{L}$ are called loading matrix and score matrix, respectively, and $\boldsymbol{E}$ is the residual matrix. Loading matrix contains principal component coefficients, and score matrix is the representation of $\boldsymbol{X}$ in the principal component space. One of the main advantages of using PCA for classification is that the results can be presented graphically in the score plot, a scatter plot of the score matrix, also known as principle component projection plot. In the score plot, the projection points of similar samples will be clustered together.

3. Experimental

3.1. Materials and reagents

RPT products were purchased from nine different pharmaceutical stores. The summary of these products is listed in Table 1. The standard RPT sample was purchased from National Institute for the Control of Pharmaceutical and Biological Products (China).

Methanol (HPLC grade, Kermel Co. Ltd., Tianjin, China), acetonitrile (HPLC grade, Hanbang Co. Ltd., Jiangsu, China) and double distilled water were used for preparation of mobile phase.

3.2. Instruments

All the chromatograms were collected from Shimadzu LC-10ATvp high-performance liquid chromatography equipped with a Shimadzu SPD-10A vp diode array detector and the workstation software of CLASS-VP (Ver.6.1). The producer of the HPLC system is Shimadzu Co. Ltd. (Japan).

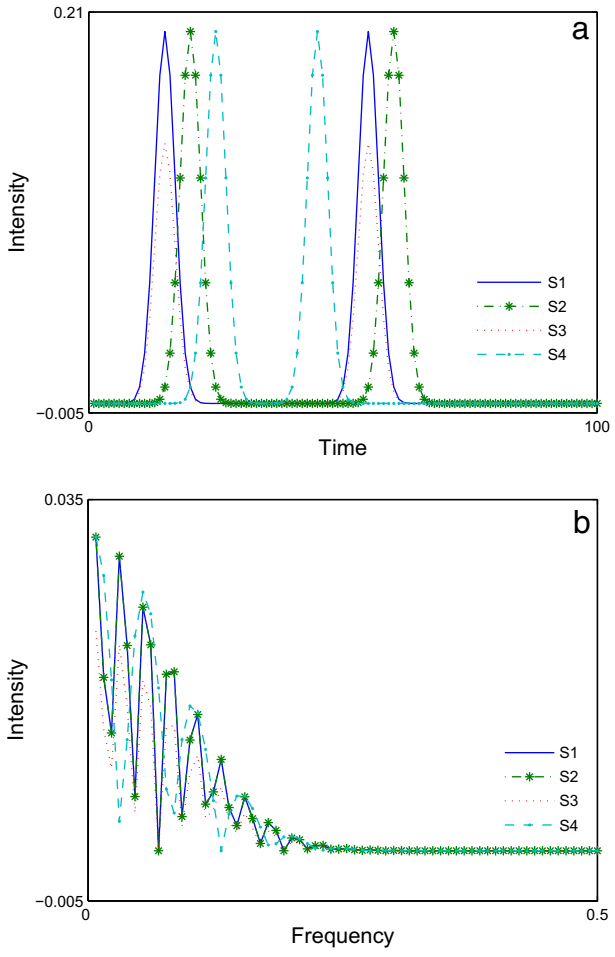

Fig. 1. (a) Simulated chromatograms and (b) magnitude data obtained from global FT.