Quality assessment of Traditional Chinese Medicine using HPLC-PAD combined with Tchebichef image moments

Xue Wang, Bao Qiong Li, Min Li Xu, Jin Jin Liu, Hong Lin Zhai*

College of Chemistry & Chemical Engineering, Lanzhou University, Lanzhou, 730000, PR China

Abstract

Traditional Chinese medicines have gotten growing attention and the product quality also became uneven. In this paper, Tchebichef image moments coupled with high performance liquid chromatography with photodiode array detector were proposed to the similarity analysis of the fourteen batches *Pudilan Xiaoyan tablets* (Chinese medicine complex), which were based on the grayscale images of three dimensional (3D) fingerprint spectra. Hierarchical cluster analysis was used for displaying the classification results. As a comparison, traditional principal component analysis (PCA) method based on the chromatograms under single-fixed wavelength was carried out on the same samples. The analytical results indicated that the proposed approach could provide more accurate and reasonable results owing to its powerful invariant and multi-resolution capability of image moments.

Keywords: similarity analysis, Traditional Chinese Medicines, grayscale image, Tchebichef moment, principal component analysis

---

* Correspondence to: Tel.: +86 931 8912596; fax: +86 931 8912582; E-mail address: zhailh@163.com (H.L. Zhai).

1. Introduction

Traditional Chinese medicines (TCMs) have been used for thousands of years and are still in wide demand now. Not only the curative effects of the TCM for preventing and treating human diseases have been widely recognized, but also the medicine food homology has gotten more and more attention, which has been used as health-care products for cooking porridge and making tea in recent years [1]. Different from western medicine generally produced from chemical synthesis, TCM usually contains some herb ingredients that have certain treatment effects themselves. The synergic effect of the multi-ingredients so become the key of the TCM's therapeutic effects, which may induce the difficulty for quality control of TCM [2]. Additionally, the contents of pharmacological active components may be influenced by different growth conditions including climate, soil, cultivating conditions, harvest season as well as different processing technologies.

Pudilan Xiaoyan tablet (PXT), one of the common TCM prescriptions in daily life, is well known for its efficacy in antibacterial drugs, such as antiviral, anti-inflammatory, and it also has immune regulating function. Dandelion, Radix Scutellariae, Bunge Corydalis Herb and Radix Isatidis are the mainly ingredients of PXT, and each of them can be seen as an independent herbal medicine with unique effect. At present, many pharmaceutical manufactures were provided the PXT production license in China, which may induce the quality of the medicine uneven. Therefore, it is necessary to establish accurate and reliable analytical approaches for the quality assessment of TCM.

Traditionally, chemical markers or pharmacological active components were used to evaluate the quality of TCM [3]. However, in normal conditions, since the composition of the TCM is complex, the practice is not objective and accurate enough. Fingerprint analyses with the

characters of entirety and system were also widely applied as the measurements of the quality consistency and stability of TCM, and it has been highly recommended by both the Food and Drug Administration [4] and European Medicines Agency [5]. Moreover, State Food and Drug Administration of China suggested that all of herbal chromatograms should be evaluated by means of their similarities [6, 7]. Principal component analysis (PCA) was usually employed to the extraction of the most influential components in matrix measured for the similarity analysis [8, 9]. In recent years, with higher resolving power and reproducibility, high-performance liquid chromatography coupled with photodiode array detector (HPLC-PAD) has been widely used for assessment of TCM [10]. Under the multi-wavelength in PAD, the three-dimensional (3D) fingerprint spectra can reveal the quality characteristic of TCM completely [11]. Therefore, the similarity analyses of TCM based on the 3D fingerprint spectra should be more objective, comprehensive, and accuracy. However, the direct comparisons of 3D fingerprint spectra are hard work, and the similarity analyses based on the chromatographic spectra are often affected by many factors such as noise signals, the shifts of peaks and baseline.

Image moments as feature descriptors are powerful in many research fields and the applications mainly contained pattern recognition [12, 13], object classification [14], edge detection [15], and image retrieval [16, 17]. Owing to the invariance and multi-resolution properties, several image moment approaches were successfully applied to the quantitative and qualitative analyses of complex samples [18-20]. Wavelet moment combined with 3D HPLC-PAD fingerprint spectra was used to the similarity analysis of traditional Chinese medicine materials [21]. Compared with wavelet moments, discrete Tchebichef moments possess more powerful descriptive ability with multi-resolution property and minimum calculation error, and could obtain

more accurate results [20].

In this work, the discrete Tchebichef moment method was used for the similarity analysis of fourteen batches of *Pudilan Xiaoyan tablets* based on the digital grayscale images of 3D HPLC-PAD fingerprint spectra. Furthermore, PCA methods were performed on the same data matrix, and the obtained results were compared with that of Tchebichef moment method. Hierarchical cluster analysis was used for classifying the samples.

2. Material and methods

2.1 Reagents and samples

Fourteen batches of *Pudilan Xiaoyan tablet* (PXT) samples derived from different manufactories and procedures were purchased from different pharmacies (Lanzhou, China), and described in Table 1. HPLC-grade methanol used for HPLC determination was obtained from Biaoshiqi (Tianjin, China). Phosphoric acid was analytical reagent (Fuchen, Tianjin, China).

For each batch, ten tablets PXT were accurately weighed, ground into powder and extracted with 80 mL of 70% methanol in an ultrasonic water bath for 30 min. Additional methanol was added to adjust the volume, and then the solvent was filtered through a 0.22 μ m microfiltration membrane (Shanghai, China).

2.2 Chromatographic instrument and conditions

All HPLC analyses were performed with a Waters series HPLC system (Milford, MA, USA) equipped with a 2998 PAD, a 1525 Binary HPLC Pump, a column oven, and a manual injector. Different HPLC parameters including various columns, mobile phases, and column temperature were examined and compared. Finally, chromatographic separation was carried out at 30 °C using

a Waters C-18 column (4.6 mm × 250 mm, 5μm). The mobile phase consisted of 0.1% phosphoric acid aqueous solution (component A) and methanol (component B). The gradient elution condition (% A in B), with a flow rate of 1.0 mL min⁻¹, was set as follows: 0-40 min, 60%; 40-60 min, 20%; 60-80 min, 60%. The sample volume injected was 5μL. The DAD spectra were recorded from 200 to 400 nm.

The 3D HPLC-PAD fingerprint spectrum and the corresponding grayscale image of one sample are shown in Fig. 1.

3. Data and methodology

3.1 The calculation of Tchebichef moments

Tchebichef (also called Chebyshev) moments are a set of orthogonal moment functions based on the discrete Tchebichef polynomials that defined over the image coordinate space. Their implementation does not involve any numerical approximation, which makes Tchebichef moments superior to the conventional orthogonal moments such as Wavelet moments. For an image intensity function $f(x, y)$ that represents the intensity value at the position (x, y) in the image with size of $N \times M$, the Tchebichef moments with the order $(m+n)$ are defined as:

$$T_{nm} = \frac{1}{\rho(m, M) \rho(n, N)} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} t_m(x) t_n(y) f(x, y), \quad (1)$$

$$m = 0, 1, 2, \dots, M-1; \quad n = 0, 1, 2, \dots, N-1$$

The basis functions of Tchebichef moments are the discrete orthogonal Tchebichef polynomials satisfying the property of orthogonality in the discrete domain

$$\sum_{x=0}^{N-1} t_m(x) t_n(x) = \rho(n, N) \delta_{nm}, \quad 0 \leq m, n \leq N-1$$

where δ_{nm} is the Kronecher symbol, $\rho(n, N)$ is the square-norm and $t_n(x)$ is the discrete Tchebichef polynomial of degree n that has the following recurrence relation:

$$(n+1)t_{n+1}(x) - (2n+1)(2x-N+1)t_n(x) + n(N^2-n^2)t_{n-1}(x) = 0, \quad (2)$$

$$n = 1, 2, \dots, N-1.$$

In order to keeping the invariance of Tchebichef moments, Mukundan *et al.* [22] proposed normalizing the Tchebichef polynomial by $\beta(n, N) = N^n$ ($\beta(n, N)$ is a suitable constant which is independent of x), which leads to the modifications $\tilde{\rho}(n, N) = \rho(n, N) / (\beta(n, N))^2$ and $\tilde{t}_n(x) = t_n(x) / \beta(n, N)$. The Tchebichef moments then can be defined as:

$$f(x, y) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} T_{nm} \tilde{t}_m(x) \tilde{t}_n(y) \quad (3)$$

$$m = 0, 1, 2, \dots, M-1; n = 0, 1, 2, \dots, N-1$$

and the following normalized recurrence formulas are yielded:

$$\tilde{t}_n(x) = \frac{(2n-1)(2x-N+1)\tilde{t}_{n-1}(x)}{nN} - \frac{(n-1)[N^2-(n-1)^2]\tilde{t}_{n-2}(x)}{nN^2},$$

$$\tilde{t}_0(x) = 1,$$

$$\tilde{t}_1(x) = \frac{2x+1-N}{N},$$

$$\tilde{\rho}(n) = \left(\frac{2n-1}{2n+1} \right) \left(1 - \frac{n^2}{N^2} \right) \tilde{\rho}(n-1),$$

$$\tilde{\rho}(0) = N.$$

3.2 Principal component analysis

Principal component analysis (PCA) [23] was performed on the data matrix of the chromatographic fingerprints at single-wavelength (the wavelength of 270 nm was selected to assure the integrity of chromatography peaks information). Chromatography drifts, the common phenomena in the determination of HPLC-PAD, can affect the analysis results. In order to eliminate the influence, the Interval Correlation Optimised Shifting algorithm (*icoshift*) [24] was carried out to align chromatography peaks. Then, the similarity analyses with PCA method were performed on the data matrix that before and after alignment. At last the results were used to compare with that of our proposed method.

3.3 Hierarchical cluster analysis

Hierarchical clustering analysis (HCA) [25] is widely used for classification owing to its visual observation and simplify. Therefore, in our work, HCA was employed to present the classification of PXT samples, in which Euclidean distance and ward linkage were used to establish clustering tree.

All calculation programs were written in M-file based on MATLAB v7.0 (Mathworks Inc. USA), and carried out with a PC (CPU 3.40 GHz, RAM 16.0 GB).

4. Results and discussion

4.1 Tchebichef moments

Our previous studies indicate that image moment methods could well deal with the influence of noise and overlapped peaks [18-20]. Tchebichef moments have the better characters of invariance and multi-resolution. In addition, the orthogonal property in the domain of the image coordinate space has eliminated the numerical approximation and given minimal information redundancy. Thus, the discrete orthogonal Tchebichef moments are expected to perform better for the analysis of digital grayscale images.

Although Tchebichef moments have many advantages, the numerical instability still appears with increasing orders [26]. Therefore, we had made some experimental tests on the different order values of Tchebichef moments. Finally, the maximum orders ($n=5$, $m=2$) were employed to the similarity analysis in this work.

4.2 Similarity analyses with clustering trees

4.2.1 Tchebichef moment method

According to the calculated Tchebichef moments (a total of 18 moments), HCA was carried out to represent the result of similarity analysis, and the clustering tree was shown in Fig.2a.

4.2.2 Principal component analysis method

In HPLC analysis, chromatography drifts are a common situation (as shown in Fig.3a), and the aligned chromatograms by *icoshift* algorithm were shown in Fig.3b. Based on the aligned chromatograms, the clustering tree with the first two principal component described 92.5% of the variability is showed as Fig.2b. As the comparison, the top five principal components including 89.8% information of the raw data (unaligned chromatograms) were selected as the feature variable, and the clustering tree is shown as Fig.2c.

Although the principal components of PCA based on the two datasets (aligned and unaligned data) were interpreted the similar variable (about 90% of original variables), the classification results are different (shown in Fig.2b and Fig.2c). The more principal components of PCA could make the variable have more shifts pattern in the chromatograms [24]. The distance between the groups in Fig.2c was inconspicuous, which illustrated that there was no clear classification in Fig.2c. Thus, the retention time shifts will influence the optimizing and interpreting of dataset and it is necessary to be corrected.

4.2.3 Details comparison

In Fig.2a and Fig.2b, the samples can be regarded as dividing into three groups. The classification results of two methods were consistent in some samples, such as the four samples from *Jiren Pharmaceutical Company*. In addition, both in Fig.2a and Fig.2b, samples **5**, **6** and **12** in the second group and sample **7** and **8** as well as sample **10** and **11** in the third group were also classified consistent. The biggest difference was the classification for sample **9**, **13** and **14**. In

order to clarify their affiliation, the correlation coefficient among samples was calculated, and the fingerprint spectra (showed in Fig.4) as well as the grayscale images of 3D fingerprint spectra (showed in Fig.5) were represented.

Both in Fig.2a and Fig.2b, the second group was different with the samples **9** and **14**. Thus, we have calculated the correlation coefficient among samples **5**, **9**, **12** and **14**, and the values were listed in Table 2. In Table 2, the value between samples **5** and **12** was 0.9835 that greater than the values between samples **5** and **14**, samples **9** and **12** as well as samples **12** and **14** with the values of 0.9764, 0.9802 and 0.9684, respectively. Thus, that indicated the samples **5** and **12** were more similar for a class.

In Fig.2a, samples **7** and **8** were in same subgroup, but separated in Fig.2b, in which sample **8** was near sample **13**. For confirming their relation, the fingerprint spectra of the samples were plotted in Fig.4. As can be seen from Fig.4, the fingerprint spectrum of sample **13** was different from the other two especially when the retention time in 19 min, and that samples **7** and **8** were in complete agreement. Thus, according to the descriptions of the fingerprint spectra, the samples **7** and **8** were actually in one subgroup that exclusive sample **13**, which were agree well with the grouping represented in Fig.2a.

And in Fig.2a, Sample **13** was classified into one group with samples **9** and **14**. For further illustrating the similarity among these samples, the grayscale images of 3D fingerprint spectra were constructed in Fig.5 and sample **12** was selected as a reference. In Fig.5, the differences among these samples were mainly focus on the regions marked with the colored boxes labeled (a), (b), and (c), respectively. From Fig.5, it can be seen that only sample **12** was lacking of certain corresponding peaks compared with others, especially at the region of (a) and (c). Thus, samples **9**

and **14** should not be in same group with sample **12**, which means the classification of these samples in Fig.2b is unreasonable. Then compared with the grayscale images of samples **9**, **13** and **14** in Fig. 5, the regions of (a) and (b) were all consistent and there were some difference in region (c) that sample **9** was with more peaks and samples **13** and **14** were more consistent. Therefore, samples **13** and **14** were more similar than sample **9**, which were corresponding to the classification again in Fig.2a.

According to the comparison mentioned above, the moment method pays more attention to the comprehensive source of information, so the classification results based on our method were more accurate, which most probably owing to the characters of multi-resolution in Tchebichef moment method.

Additionally, from the clustering trees, the samples in the first manufacture were classified in one group, which indicated it was more stable than the other two manufactures. And from the whole Fig.2a, all products from different batches can be considered as classifying into one big group, which illustrated that the manufactories were followed the criterion of Chinese Pharmacopoeia. It was also because of this, the PCA with single wavelength probably ignored subtle differences that the grayscale images in Fig.5 were illustrated.

5. Conclusions

The chromatograms measured under a single-fixed wavelength could not reveal the systemic characteristic of complex sample completely. Therefore, traditional analytical methods are difficult to obtain the robust and accurate results. Here, the Tchebichef moment method that can describe the whole characters of 3D HPLC-DAD fingerprint spectra was successfully applied into the

similarity analyses of PXT. Our study demonstrated that the proposal method not only gave more accurate results for similarity analysis based on 3D HPLC-DAD fingerprint spectra, but also could be extended to other 3D spectra analyses.

Acknowledgement

The undergraduates Jun Dong Wang, Li Jie Yang, Yu Xin and Yuan Fang Xiong in College of Chemistry & Chemical Engineering at Lanzhou University took part in the experimental works. This study was supported by National Natural Science Foundation Committee of China (Grant No. 21275067) and National Science Foundation for Fostering Talents in Basic Research of the National Natural Science Foundation of China (Grant No. J1103307).

Reference

- [1] C. Ping, History and development of traditional Chinese medicine, Science Press, Beijing, 1999.
- [2] Y.Z. Liang, P. Xie, K. Chan, Quality control of herbal medicines, J. Chromatogr. B Analyt. Technol. Biomed. Life Sci., 812 (2004) 53-70.
- [3] J. Xie, W. Wang, Y. Zhang, Y. Bai, Q. Yang, Simultaneous analysis of glycyrrhizin, paeoniflorin, quercetin, ferulic acid, liquiritin, formononetin, benzoic acid and isoliquiritigenin in the Chinese proprietary medicine Xiao Yao Wan by HPLC, J. Pharm. Biomed. Anal., 45 (2007) 450-455.
- [4] FDA, FDA Guidance for Industry Botanical Drug Products (Draft Guidance), In: Administration, U.F.A.D, USA, (2000).
- [5] EMEA, Guidance on quality of herbal medicinal products, European Medicines Agency, London, (2006).
- [6] China, Requirements for Studying Fingerprints of Traditional Chianese Medicine Injection., IN: China, D.A.B.O. (Ed.), Beijing, China, (2002).
- [7] F. Gong, Y.-Z. Liang, P.-S. Xie, F.-T. Chau, Information theory applied to chromatographic fingerprint of herbal medicine for quality control, J. Chromatogr. A, 1002 (2003) 25-40.
- [8] Y. Ni, Y. Liu, S. Kokot, Two-dimensional fingerprinting approach for comparison of complex substances analysed by HPLC-UV and fluorescence detection, Analyst, 136 (2011) 550-559.
- [9] S.-k. Yan, W.-f. Xin, G.-a. Luo, Y.-m. Wang, Y.-y. Cheng, An approach to develop two-dimensional fingerprint for the quality control of Qingkailing injection by high-performance liquid chromatography with diode array detection, Journal of Chromatography A, 1090 (2005) 90-97.
- [10] B. Xie, T. Gong, M. Tang, D. Mi, X. Zhang, J. Liu, Z. Zhang, An approach based on HPLC-fingerprint and chemometrics to quality consistency evaluation of Liuwei Dihuang Pills

252 produced by different manufacturers, *J. Pharm. Biomed. Anal.*, 48 (2008) 1261-1266.

253 [11] H.L. Zhai, F.D. Hu, X.Y. Huang, J.H. Chen, The application of digital image recognition to the

254 analysis of two-dimensional fingerprints, *Anal. Chim. Acta*, 657 (2010) 131-135.

255 [12] S.O. Belkasim, M. Shridhar, M. Ahmadi, Pattern recognition with moment invariants - A

256 comparative study and new results, *Pattern Recognition*, 24 (1991) 1117-1138.

257 [13] J. Flusser, T. Suk, Pattern recognition by affine moment invariants, *Pattern Recognition*, 26 (1993)

258 167-174.

259 [14] C. Gope, N. Kehtarnavaz, G. Hillman, B. Würsig, An affine invariant curve matching method for

260 photo-identification of marine mammals, *Pattern Recognition*, 38 (2005) 125-132.

261 [15] S. Journal of Fluorescence Ghosal, R. Mehrotra, Orthogonal moment operators for subpixel edge

262 detection, *Pattern Recognition*, 26 (1993) 295-306.

263 [16] J.C. Patra, S. Sood, P.K. Meher, C. Bornand, Ieee, Content-based Image Retrieval Using

264 Orthogonal Moments Heuristically, in: 2008 Ieee International Conference on Systems, Man and

265 Cybernetics, 2008, pp. 512-517.

266 [17] G. Paschos, I. Radev, N. Prabakar, Image content-based retrieval using chromaticity moments, *Ieee*

267 *Transactions on Knowledge and Data Engineering*, 15 (2003) 1069-1072.

268 [18] H.L. Zhai, Z.J. Shan, R.N. Li, E. Yu, Simultaneous quantitative analysis of three compounds using

269 three-dimensional fluorescence spectra based on digital image techniques, *J Fluoresc*, 22 (2012)

270 1013-1019.

271 [19] H.L. Zhai, Y.Y. Zhai, P.Z. Li, Y.L. Tian, A simple approach to quantitative analysis using

272 three-dimensional spectra based on selected Zernike moments, *Analyst*, 138 (2013) 683-687.

273 [20] B.Q. Li, J. Chen, J.J. Li, X. Wang, H.L. Zhai, The application of a Tchebichef moment method to

the quantitative analysis of multiple compounds based on three-dimensional HPLC fingerprint spectra, *Analyst*, 140 (2015) 630-636.

[21] H.L. Zhai, B.Q. Li, Y.L. Tian, P.Z. Li, X.Y. Zhang, An application of wavelet moments to the similarity analysis of three-dimensional fingerprint spectra obtained by high-performance liquid chromatography coupled with diode array detector, *Food Chem.*, 145 (2014) 625-631.

[22] R. Mukundan, S.H. Ong, P.A. Lee, Image analysis by Tchebichef moments, *Ieee Transactions on Image Processing*, 10 (2001) 1357-1364.

[23] K.M. Pierce, J.L. Hope, K.J. Johnson, B.W. Wright, R.E. Synovec, Classification of gasoline data obtained by gas chromatography using a piecewise alignment algorithm combined with feature selection and principal component analysis, *J. Chromatogr. A*, 1096 (2005) 101-110.

[24] G. Tomasi, F. Savorani, S.B. Engelsens, icosift: An effective tool for the alignment of chromatographic data, *J. Chromatogr. A*, 1218 (2011) 7832-7840.

[25] P.R. Kannel, S. Lee, S.R. Kanel, S.P. Khan, Chemometric application in classification and assessment of monitoring locations of an urban river system, *Anal. Chim. Acta*, 582 (2007) 390-399.

[26] R. Mukundan, Some computational aspects of discrete orthonormal moments, *Ieee Transactions on Image Processing*, 13 (2004) 1055-1059.

294 **Table caption:**

295 Table 1 Summary of the tested samples

296 Table 2 the values of the correlation coefficient among samples

297

298 **Figure caption:**

299 Fig.1 The HPLC-DAD fingerprint spectrum of one sample (a) the three-dimensional spectrum (b)
300 the inverted digital grayscale image

301 Fig.2 The clustering tree of samples based on (a) Tchebichef moment method, (b) PCA for aligned
302 chromatograms, and (c) PCA for raw chromatograms

303 Fig.3 The fingerprint chromatograms of 14 samples (a) the raw chromatograms (b) the *icoshift*
304 aligned chromatograms

305 Fig.4 The comparison of the fingerprint spectra among the samples **7, 8** and **13**

306 Fig.5 The comparison of the grayscale images among the samples **9, 12, 13** and **14**

307

308

309

Table 1 Summary of the tested samples

Sample no.	Manufactory	Batch no.
1	Jiren Pharmaceutical Co., Ltd, Anhui, China	130801
2	Jiren Pharmaceutical Co., Ltd, Anhui, China	140106
3	Jiren Pharmaceutical Co., Ltd, Anhui, China	140108
4	Jiren Pharmaceutical Co., Ltd, Anhui, China	140112
5	Yunnan Baiyao Co., Ltd, Yunnan, China	ZKB1321
6	Yunnan Baiyao Co., Ltd, Yunnan, China	ZKB1326
7	Yunnan Baiyao Co., Ltd, Yunnan, China	ZAB1401
8	Yunnan Baiyao Co., Ltd, Yunnan, China	ZEC1401
9	Symbol Pharmaceutical Co., Ltd, Guangdong, China	20140101
10	Symbol Pharmaceutical Co., Ltd, Guangdong, China	20140104
11	Symbol Pharmaceutical Co., Ltd, Guangdong, China	20140105
12	Tianjin Pharmaceutical Co., Ltd, Tianjin, China	140413
13	Zhongtian Pharmaceutical Co., Ltd, Tianjin, China	ZT109294
14	Zhongding Pharmaceutical Co., Ltd, Jilin, China	20140101

310

311

Table 2 the values of the correlation coefficient among samples

Sample no.		R
5	12	0.9835
5	14	0.9764
9	12	0.9802
12	14	0.9684

312

Fig.1

Fig.2

Fig.3

Fig. 4

328 **Fig.5**

329

330

331