Chemometrics and Intelligent Laboratory Systems 103 (2010) 83–89

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

An efficient method for determining the chemical rank of three-way fluorescence data arrays

Shaohui Yu [a,b], Yujun Zhang [a,*], Huanbo Wang [a], Xue Xiao [a]

[a] Key Laboratory of Environmental Optics and Technology, Anhui Institute of Optics and Fine Mechanics, The Chinese Academy of Sciences, Hefei 230031, China
[b] Mathematics Department, Hefei Normal University, Hefei 230061, China

ARTICLE INFO

Article history:
Received 21 January 2010
Received in revised form 2 April 2010
Accepted 9 April 2010

Keywords:
Chemical rank
Singular value decomposition
Monte Carlo simulation
Fourier transform
Frequency localization

ABSTRACT

An efficient method is proposed for determining the chemical rank of three-way fluorescence data arrays. At first, the original three-way fluorescence data arrays are preprocessed by Monte Carlo simulation and a new set of data arrays is generated. The new set of data arrays obtained does not only keep all the useful information, but the noises from the common background are largely removed, which results in the improvement of the signal to noise ratio of the data and is beneficial for the later frequency analysis. Then, we perform singular value decomposition over the new data and frequency analysis on the subsequent eigenvectors, with which it is very easy to distinguish the spectra from the noises. Furthermore, a new quantity *frequency localization* is introduced to quantify the frequency characteristics of the eigenvectors. With this quantity, we can easily and accurately select out the spectra from the mess of data. The feasibility of the method is verified by determining the chemical rank of two-component mixtures with simple calculation procedures and high efficiency. Finally, the efficiency of our method is further illustrated by comparison with the core consistency diagnostic (CORCONDIA) method in the analysis of mixtures with different concentration and different number of components.

1. Introduction

Estimation of the chemical rank, also known as the number of components in mixtures, is an important or even a primary issue in multi-component system analysis. In identifying and calibrating the multi-component systems, many approaches depend on the pre-estimation of the component number. Either overestimation or underestimation will lead to less efficiency or even failure of certain algorithms. Although some methods are not very sensitive to the component number, prerequisites must be met that estimation of the component number should be greater than the actual component number. Under this condition, the approach may work, but the run time, storage and number of iterations will increase with the overestimation of component number and this will inevitably affect the efficiency of the algorithm. Therefore, the accurate estimation of the number of components in multi-component systems is very important, especially in the field experiments.

It is well known that three-way fluorescence data arrays are constituted by a collection of two-way fluorescence excitation–emission matrices which are stacked together along the sample space. Without loss of generality, the three-way data arrays can be denoted by matrix $X_{I\times J\times K}$, where I denotes the dimension of excitation, J the dimension of emission and K the dimension of the sample (the number of samples measured) and the value of each matrix element is the fluorescence intensity of the corresponding point in the three dimensional space spanned by excitation, emission and sample.

In order to estimate the chemical rank of three-way data arrays, many methods have been proposed so far mainly as follows: (1) residual analysis [1]. This approach determines the correctness of the selected number of factors according to the residuals between the value of the PARAFAC algorithm [2] and the three-way data arrays. (2) CORCONDIA [3]. Based on the equivalence of the PARAFAC model [2] and the constrained Tucker3 model [4], this method defines a core consistency function to estimate the component number. (3) ADD-ONE-UP [5]. This approach starts unfolding $X_{I\times J\times K}$ along the excitation and emission space, then performs singular value decomposition (SVD) and reconstructs $X_{I\times J\times K}$ by different eigenvectors. The component number is determined by the residual of PARAFAC algorithm and the three-way data array. (4) PNVOP [6]. The three-way data array $X_{I\times J\times K}$ is unfolded and an orthogonal projection matrix is constructed. Then the component number is estimated by F-test. (5) TMSC [7]. The three-way data array is preprocessed by unfolding $X_{I\times J\times K}$ to form two two-way matrices with dimensions $I\times JK$ and $J\times IK$ (here JK and IK mean $J\times K$ and $I\times K$, respectively). After singular value decomposition, the component number is determined by the residual of orthogonal projection matrix which is formed by the first i columns of eigenvectors. (6) Pseudo-sample extraction and the projection technique [8]. This method firstly produces two pseudo-samples by

* Corresponding author.
E-mail address: yjzhang@aiofm.ac.cn (Y. Zhang).

0169-7439/$ – see front matter © 2010 Published by Elsevier B.V.
doi:10.1016/j.chemolab.2010.04.006

DTLD algorithm [9] and continues with singular value decomposition, then estimates the component number by the residual of projection space. (7) SPPH [10]. This method is carried out by unfolding $X_{I\times J\times K}$, producing two-way arrays and performing singular value decomposition. Then the component number is determined by the residual resulting from different factors. (8) LTMC [11]. There are two matrices $R1$ and $R2$ produced in this method. $R1$ is formed by the sum of $X_{..k}$ ($k=1, \cdots, K$) and $R2$ is produced by multiplying each $X_{..k}(k=1, \cdots, K)$ with random non-zero number and adding them together. As the subspace spanned by the corresponding spectra eigenvectors of N components is the same but orthogonal with the subspace spanned by noise eigenvectors, the component number can be estimated by subspace projection method together with Monte Carlo simulation. (9) Canonical correlation technique [12]. This approach estimates the number of component by the fact that the subspace spanned by different spectra eigenvectors is the same and orthogonal with the noise eigenvectors subspace. The difference from that in Reference [11] is the introduction of a canonical correlation technique. (10) Morphological approach [13]. A morphological function is constructed by using the eigenvectors of singular value decomposition and the component number is estimated by the morphological function in Reference [13].

Since the eigenvectors from the singular value decomposition of the original three-way data arrays contain much more information than the eigenvalues, estimation of component number from eigenvectors is a good starting point and in fact, it has been taken by many methods. It can be seen that the above methods for estimation of the chemical rank of three-way fluorescence data arrays fall into two groups. In the first group, the component number is determined by comparing the residual between three-way data arrays and numerical solution from PARAFAC algorithm. This kind of method which is influenced by signal to noise ratio, requires a lot of computation because of the multiple executions of PARAFAC algorithms. In the second group, the chemical rank is determined by projection space which is spanned by eigenvectors after singular value decomposition. Though these kinds of approaches take advantage of the frequency characteristics of spectra and noises to certain extent, they are possibly also influenced by signal to noise ratio. In a word, the above methods work well only on certain three-way fluorescence data arrays due to their intrinsic limitations. In practice, at present, the joint application of several different approaches is usually adopted to evaluate the component number of a multi-component system. So, it is of great significance to explore ways of accurate estimation of component number by taking into full consideration all possible differences between the spectra and the noises. Rossi et al. [14] have estimated the rank of excitation–emission arrays by frequency analysis of eigenvectors. They investigate the ratio of the amplitude sum in one specific frequency range over the amplitude sum in the whole frequency range of each eigenvector. Since it requires to set an artificial threshold and repeatedly try different frequency ranges, this method has not been widely adopted and developed.

Inspired by frequency analysis of eigenvectors, we propose a new scheme to estimate the chemical rank of three-way fluorescence data arrays from multi-component systems. Firstly, in order to eliminate the noise effect as possible as we can, the original measured data are preprocessed by subtracting any two data arrays with Monte Carlo simulation to produce new sample data. Then singular value decomposition is applied to the new three-way data arrays and frequency analysis is performed on the subsequent eigenvectors. By inspecting the frequency spectra, we can already accurately select out the spectra. Finally, a new quantity *frequency localization* is defined for the eigenvectors to quantify their localization degree in frequency space and the component number is accurately determined by comparison of frequency localization of different eigenvectors.

The rest of the paper is organized as follows: The theoretical concepts and steps of the method including Monte Carlo preprocessing, singular value decomposition, time-frequency transform and frequency localization are introduced in Section 2. In Section 3, the feasibility of the method is testified by analyzing the three-way fluorescence data arrays of a two-component system with and without Monte Carlo preprocessing and by comparing the results from our method and the popular CORCONDIA. At last, a comprehensive summary is given in Section 4.

2. Theory

2.1. The preprocessing of the three-way data arrays

In measurement of the fluorescence of mixtures, we obtain a collection of two-way excitation–emission matrices $X_{..k}(k=1, \cdots, K)$. The obtained data contain specral signals and noises and the noises mainly arise from two origins, with one from the common background, such as those from the instruments used in the measurements and water in the solution, etc., and the other from the random noises. Generally, the noises from the common background are much smoother than the random noises and have similarities to spectral signals, thus the common background can greatly affect the estimation of the chemical rank. In order to suppress the effects of the noises in the analysis, the obtained data matrices $X_{..k}$ are preprocessed by Monte Carlo simulation before singular value decomposition is performed, i.e., two data arrays are randomly extracted from $X_{..k}(k=1, \cdots, K)$ and then one is subtracted from the other to produce a new two-way data array. This process is repeated for N times and we obtain N new two-way data arrays:

$$Y_{..n} = X_{..i} - X_{..j} (n = 1, \cdots, N), \tag{1}$$

with $i, j \in 1, \cdots, K$, $i \neq j$ and $K \leq N \leq C_K^2$. In this way, we can produce a new set of "samples" as many as $C_K^2 = K(K-1)/2$. However, we do not need so many new "samples" and only part of them are enough. For example, we can choose $N=K$. Note that, in the construction of the new data arrays $Y_{..n}$, we should make sure that each of the original K samples is chosen at least for once so that all the measured data are processed without any lose of them. The new samples $Y_{..n}(n=1, \cdots, N)$ are different from the common pseudo-samples [9] in chemometrics in that the new samples keep all the useful information contained in the initial samples, but the known and unknown common background is removed, which results in the better satisfaction of the trilinearity of the data matrix which is the basis of singular value decomposition that will be discussed in the next subsection. In addition, the subtraction of the data arrays may bring the increase of the random noises but it does not change the basic features of them, such as randomness and high-frequency oscillation and thus it will not affect the frequency analysis in this work. Since the noises from the common background are largely eliminated in the subtraction, the signal to noise ratio of the data is improved.

2.2. Singular value decomposition

The matrix $X_{..k}^T X_{..k}$ for each sample k can be decomposed into three matrices U, Λ, V with SVD and since $X_{..k}^T X_{..k}$ is symmetrical and positive semi-definite, we have $U=V$ and thus

$$X_{..k}^T X_{..k} = U \Lambda U^T, \tag{2}$$

where U is an orthogonal matrix, the superscript T of U denotes transposition, and Λ is a diagonal matrix with the elements $\Lambda_{11} \geq \Lambda_{22} \geq \cdots \geq \Lambda_{jj} \geq 0$. Note that for a symmetrical and positive semi-definite matrix, the SVD of it is equivalent to its eigenvalue decomposition. So following the routine in literatures (see, e.g., Reference [13]), we call U here as the eigenvectors and Λ_{ii} as the eigenvalues. If $X_{..k}$ is free of noise, the number of non-zero elements in

Λ is exactly the component number. But actually, noises are always mixed in $X_{..k}$, thus all the values of the elements in Λ are greater than zero. This means that it is hard to distinguish the spectra from the noises only by eigenvalues Λ_{ii} in matrix Λ. On the contrary, it is a good idea to turn to analyze the eigenvectors since they contain much more information than the eigenvalues.

2.3. Time-frequency transform of eigenvectors

In order to obtain the exact component number of a multi-component system, firstly, we unfold $Y_{..n}(n=1,\cdots,N)$ along the sample space to form an augmented matrix $[Y_{..1},\cdots,Y_{..N}]$ and then perform singular value decomposition over it,

$$[U,S,V]=svd\left([Y_{..1},\cdots,Y_{..N}]^T[Y_{..1},\cdots,Y_{..N}]\right). \tag{3}$$

Since the matrix $[Y_{..1},\cdots,Y_{..N}]^T[Y_{..1},\cdots,Y_{..N}]$ of the new samples is also symmetrical, we have $U=V$ in Eq. (3).

Then we construct a continuous function $U(t)$ from each eigenvector $\{U_i\}(i=1,\cdots,N)$ by artificially setting the time interval between any two nearest elements in this eigenvector as 1 s and do Fourier transform over the function $U(t)$,

$$A(u)=\frac{1}{T}\sum_{t=0}^{T-1}U(t)\exp\left(\frac{-i2\pi ut}{T}\right) \tag{4}$$

where t is the artificial time point, T is the total row number of U and u is the frequency.

In the frequency space, the value of eigenvector $A(u)$ is complex. Its amplitude as a function of frequency can be computed by

$$Z(u)=|A(u)|=\sqrt{(A_{\text{real}}(u))^2+\left(A_{\text{imag}}(u)\right)^2}. \tag{5}$$

Since in this work singular value decomposition is performed on the new samples produced by Monte Carlo preprocessing, it will be easier to distinguish between spectra and noise from the amplitude $Z(u)$. In comparison, the method introduced in Reference [14] is greatly influenced by artificial intervention owing to the requirement of setting a threshold value and a frequency range which have to be tried repeatedly for several times. Furthermore, the threshold value is highly dependent on the selected frequency range. In doing so, a different threshold value or a different frequency range may result in a different conclusion, thus it is hard to control.

2.4. Frequency localization

As noises are high-frequency vectors, they are uniform and highly oscillating functions of frequency. On the contrary, the spectra are low-frequency ones which are much smoother than the noises and finite values are localized at a very small range around some specific frequencies. So following the idea of introducing inverse participation ration (IPR) to quantify the localization of states in real space in solid state physics [15], we can also define a similar quantity *frequency localization* FL(Z) to describe the localization of an eigenvector $Z(u)$ in the frequency space. Firstly, $Z(u)$ is normalized,

$$Z(u)=Z(u)/\left(\sum_{u=1}^{T}Z(u)\right), \tag{6}$$

where u, like t, has the same total number T. Thus the new $Z(u)$ at each frequency u has value between 0 and 1. Then, frequency localization can be defined as follows:

$$FL(Z)=\sum_{u=1}^{T}(Z(u))^2. \tag{7}$$

From this definition, we know that $1/T\le \text{FL}(Z)\le 1$ for each $Z(u)$, where the maximum value of 1 means the most localized state since $Z(u)$ is not zero only at one specific frequency, while the minimum value $1/T$ means the most extended state and $Z(u)$ is equal at all frequencies. Thus a more localized eigenvector has larger frequency localization. Since the spectra appear only in a very small range around some specific frequencies while noises are much more irregularly distributed in the whole frequency space, as will be seen later, the spectra has a much larger FL(Z) than the noises. Therefore, we can determine whether one eigenvector is from a signal or from a noise only with the aid of FL(Z) and it is not necessary to inspect how the eigenvector $Z(u)$ is distributed in the frequency space. Correspondingly, the component number can be determined accurately.

3. Experiment results and discussions

As universal concerns in the monitoring of priority pollutants, cresol and phenol have low concentration in water and high similarity in spectra (see Fig. 1), so the qualitative and quantitative analysis of cresol and phenol in water is always a difficult but very important problem. In this work, cresol and phenol are taken as experimental reagents to verify the feasibility of the proposed method.

Ten mixtures of cresol and phenol with different concentration are prepared and the concentration of each mixture is listed in Table 1. Each mixture is scanned by fluorescence spectrophotometer with excitation wavelength EX 220–290 nm and emission wavelength EM 291–400 nm. In the measurement, the intervals are taken as 2.0 nm for the excitation wavelength and 1.0 nm for the emission wavelength. The scanning speed is 12,000 nm/min and the measured data form three-way data arrays $X_{ijk}(i=1,\cdots,110;j=1,\cdots,36;k=1,\cdots,10)$. In many previous methods, generally, singular value decomposition is performed directly on X_{ijk}, and then the diagonal elements of Λ is arranged in descending order and the component number is usually estimated according to the relative magnitude of these eigenvalues since generally noises have smaller eigenvalues than the spectra. In order to see more clearly the difference between the cases with and

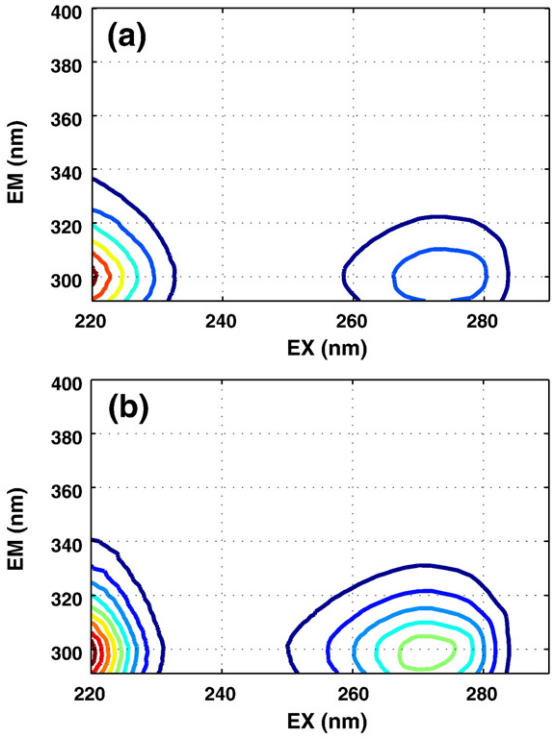

Fig. 1. The three-dimension fluorescence spectra of: (a) cresol and (b) phenol. EX means the excitation wavelength and EM means the emission wavelength.

Table 1
The two-component mixtures of cresol and phenol with different concentrations (Unit: mg/l).

	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10
Cresol	0	0.05	0.08	0.1	0.2	0.3	0.5	0.8	1	0.05
Phenol	1	0.8	0.5	0.3	0.2	0.1	0.08	0.05	0	1

without Monte Carlo preprocessing, at first, we also perform SVD over the original data matrix X_{ijk} and do Fourier transform over the eigenvectors U to obtain their frequency spectra. The results for the first six eigenvectors are plotted in Fig. 2.

From Fig. 2, we can see that it is very difficult to distinguish the spectra from the noises. For example, Fig. 2(a) and (b) are from the spectra, but they do not show any better frequency characteristics than the other eigenvectors, namely, the noises. It indicates that the contrast of the spectra over the noises in the frequency space is very bad.

Though the frequency spectrum is intuitive when it is visualized in a figure, we have to inspect them one by one to extract the spectra and exclude all the noises. When there are many components in the mixtures or when the data matrix is huge, it is not convenient to directly determine the component number from the frequency spectra. Therefore, the frequency spectrum needs to be quantified and it will be very convenient if we can describe each eigenvector with only one quantity. To this aim, frequency localization is calculated and the information about all the eigenvectors is shown in Fig. 3. It can be seen that the frequency localizations of the first seven eigenvectors are significantly greater than the others, but the first two corresponding to the spectra eigenvectors, especially the second one, have no obvious difference from the other five noises. Clearly, the common background and the random noises in the original fluorescence data severely affect the estimation of chemical rank.

In comparison, if singular value decomposition is imposed on the new data arrays (in fact it is the matrix $[Y_{\cdot\cdot 1}, \cdots, Y_{\cdot\cdot N}]$) produced by

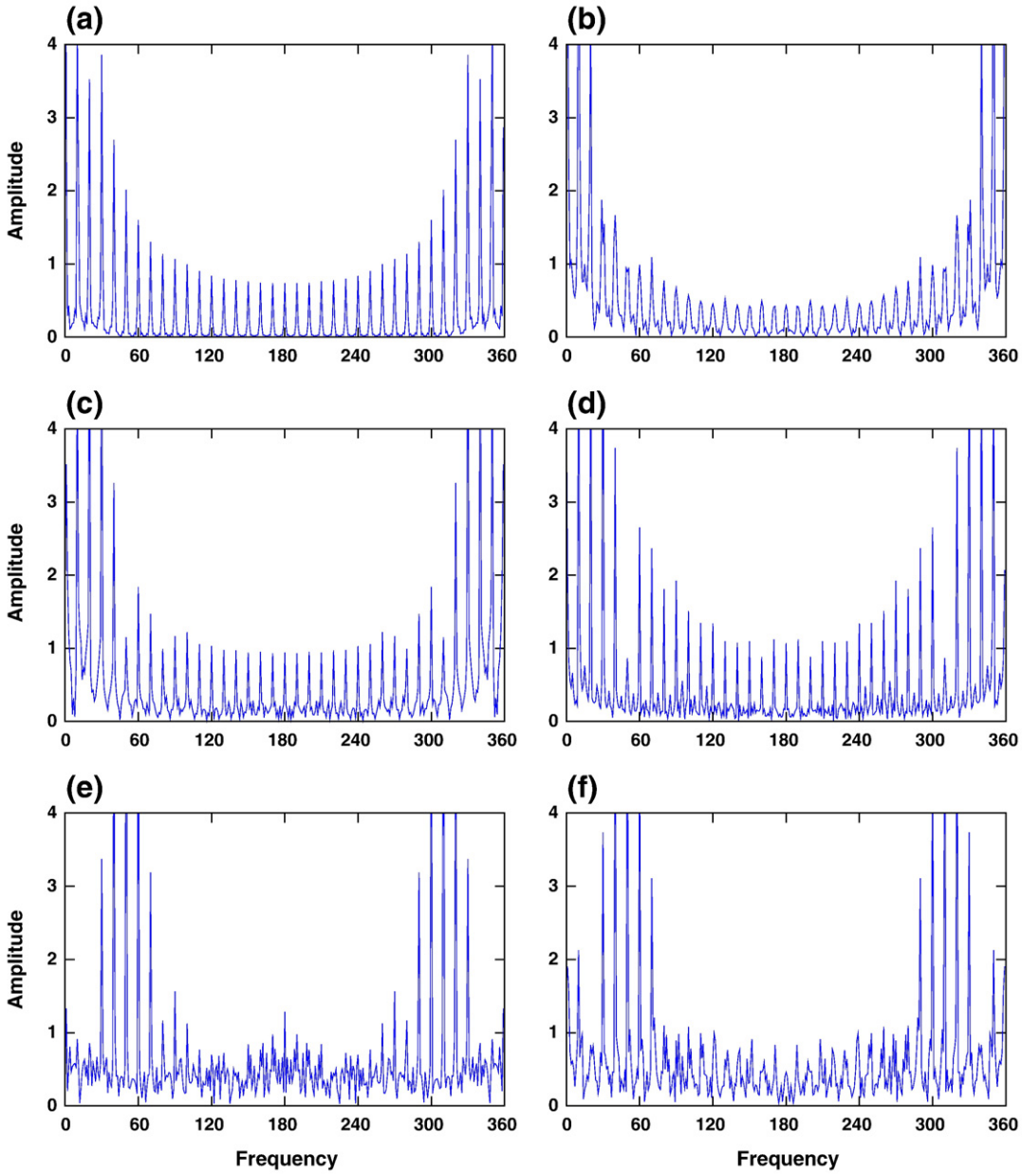

Fig. 2. Frequency spectra of the first six eigenvectors after singular value decomposition for the case with no Monte Carlo preprocessing over the original data.

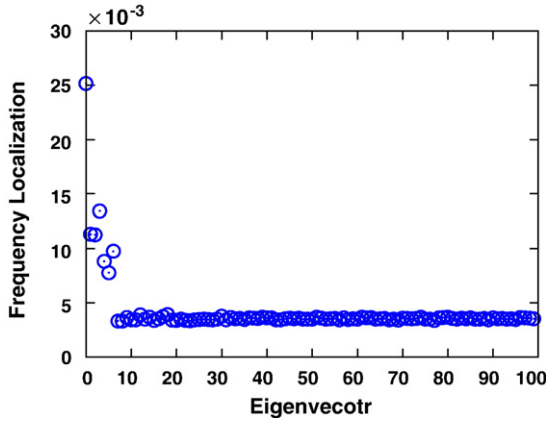

Fig. 3. Frequency localization for the case with no Monte Carlo preprocessing over the original data.

Monte Carlo preprocessing and with the common background removed, the spectra of the eigenvectors corresponding to the first six largest eigenvalues are obviously different from those in Fig. 2 and we can easily tell that Fig. 4(a) and (b) are from the spectra, while Fig. 4(c)–(f) are from the noises. In contrast, the eigenvectors shown in Fig. 4(c)–(f) are much more oscillative than those in Fig. 2(c)–(f). This is because, on one hand, the effect of the common background has been excluded and on the other hand, the subtraction of the data arrays may increase the randomness (frequency) of the random noises.

It has been shown that frequency localization is a good quantity to measure the localization degree of eigenvectors in the frequency space. The frequency localizations of the eigenvectors with Monte Carlo preprocessing are shown in Fig. 5. From this figure, it can be seen that the localization (about 7.0×10^{-3}) of the first two eigenvectors are far greater than the other eigenvectors, so the first two eigenvectors are the corresponding spectra. The localization of all the latter eigenvectors is around 3.5×10^{-3}, which fully demonstrates the common characteristics of the noises, namely, random frequency

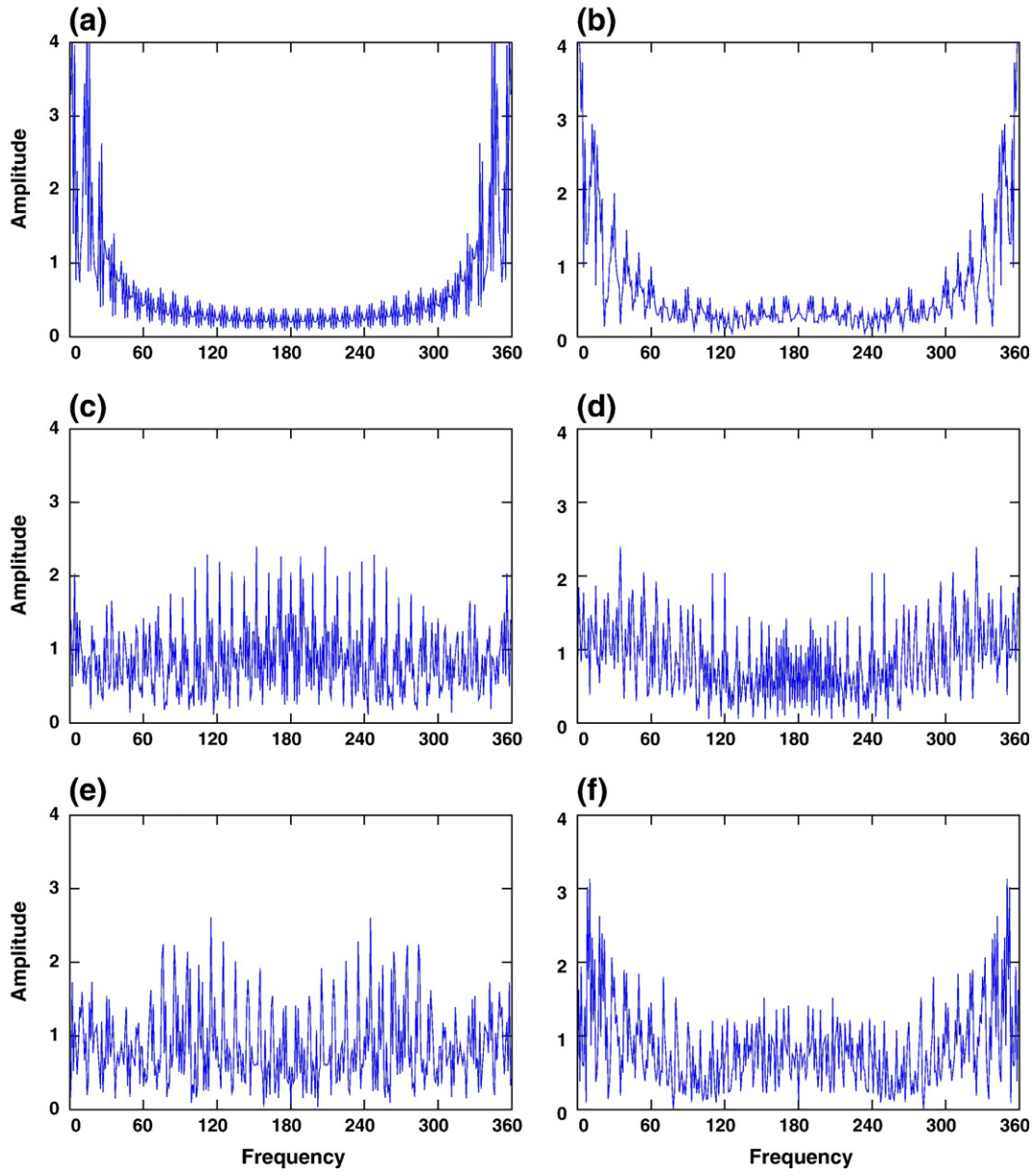

Fig. 4. Frequency spectra of the first six eigenvectors of $Y_{..n}$ after singular value decomposition for the case with Monte Carlo preprocessing over the original data.

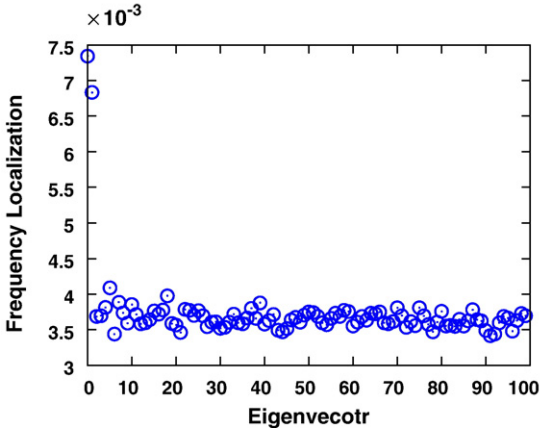

Fig. 5. Frequency localization for the case with Monte Carlo preprocessing over the original data.

Table 2
The three-component mixtures of cresol, thymol and phenol with different concentrations (Unit: mg/l).

	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10
Cresol	0	0.05	0.08	0.1	0.2	0.3	0.5	0.8	1	0.05
Thymol	0.3	0.5	0.2	0.8	0.08	1.0	0	0.1	0.05	0.4
Phenol	1	0.8	0.5	0.3	0.2	0.1	0.08	0.05	0	1

distribution. Thereby, the component number of the three-way fluorescence data arrays studied in this experiment is easily obtained as 2 by the frequency localization. This is simpler and more direct than by visualizing the frequency spectrum one by one, especially for mixtures in field experiments with probably many components. In this example, we see that we do not need to set up an artificial threshold. However, if it is in a case with very difficult data, a suitable decision criterion is still suggested. The component number can be decided as the number of the eigenvectors whose frequency localization is larger than the value $FL_{min}+a\times(FL_{max}-FL_{min})$, where FL_{min} denotes the minimum of the frequency localization and FL_{max} the maximum of the frequency localization. The parameter a is tunable and 0.5 is suggested generally. Here note that, what shown in Figs. 4 and 5 are the results for $N=10$ chosen in Eq. (1). In fact, we have tried many different ways to select out the original data matrices for producing the new data arrays for a fixed N and we have tried different N which is between 10 and $45(=C_{10}^2)$. In all these tests we get the same conclusion.

To further illustrate the effectiveness of the method, we have compared our method and the currently popular CORCONDIA [3] in the analysis of mixtures with different concentration and different number of components including two or three components. We find that, in all the cases where CORCONDIA succeeds, our method can also get the correct component number. In addition, we find several cases where CORCONDIA fails while our method can still get the correct results. In the following, we show such an example.

Ten three-component mixtures of cresol, thymol and phenol with different concentrations are prepared and the concentration of each mixture is listed in Table 2. Cresol, thymol and phenol are selected in this experiment since they have highly similar excitation and emission spectra (See Fig. 6) and are difficult to distinguish. The three-way fluorescence data arrays of the ten mixtures are measured under the same conditions and analyzed with the same procedure as in the previous experiment. The frequency localizations of the first thirty eigenvectors are showed in Fig. 7. From Fig. 7, we can easily tell that the component number of the ten mixtures is 3, as the first three localizations are obviously far greater than the others.

We have also calculated the chemical rank by the CORCONDIA method. However, a wrong number of 2 is obtained by this method. Analysis shows that the main reason that CORCONDIA fails in this example is that the spectra of two components (cresol and phenol) in the mixtures are severely overlapped, which are very difficult to deal with in CORCONDIA especially for the low concentration. However, our method can still easily select out all the spectra signals. From the comparisons that have been made, we find that the efficiency of our method is at least comparable to CORCONDIA and even better in some situations. Of course, one difference should be noted that, our method is based on frequency analysis while CORCONDIA is not. Furthermore, our method is much simpler than CORCONDIA due to the requirement of multiple executions of PARAFAC algorithms which are generally very time consuming in CORCONDIA.

4. Conclusion

An efficient method based on frequency analysis for determining the component number of multi-component system is proposed. Compared with other similar methods that are based on frequency analysis, the advantages of this approach are the following: 1. Before the singular value decomposition, the original three-way fluorescence data arrays are preprocessed by Monte Carlo simulation. The obtained new three-way fluorescence data arrays keep all the useful information and in the meantime the effects from the common background are removed and the signal to noise ratio of frequency characteristics is greatly improved. 2. A new quantity frequency localization FL(Z) is

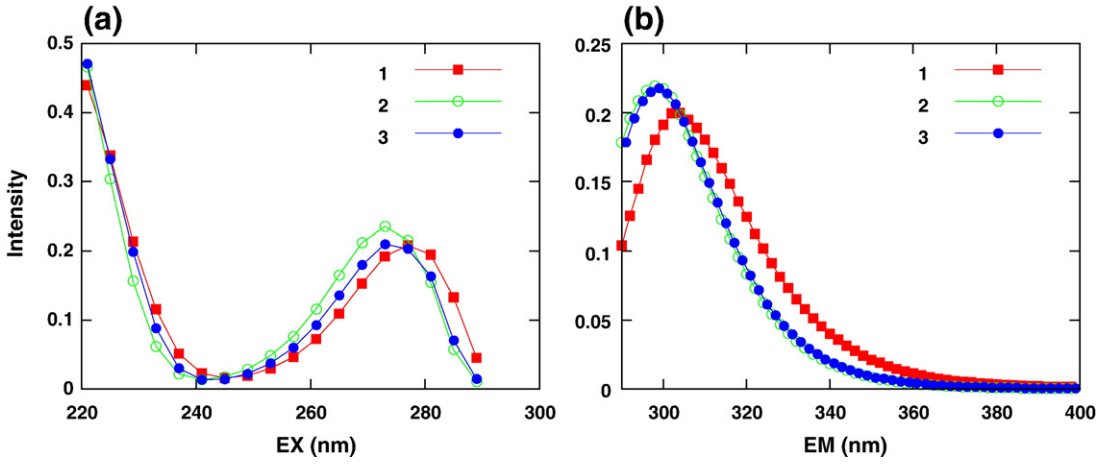

Fig. 6. The excitation and emission spectra for the three components in the three-component system: (a) Excitation spectra; (b) Emission spectra. EX means the excitation wavelength and EM means the emission wavelength. "1" stands for thymol, "2" for cresol and "3" for phenol.

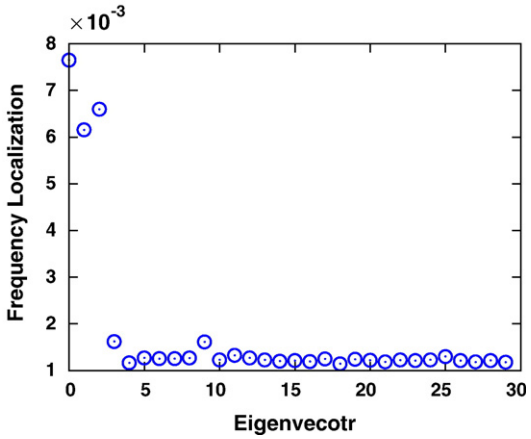

Fig. 7. Frequency localization for the three-component system with Monte Carlo preprocessing.

introduced for describing the frequency characteristics of the eigenvectors after Fourier transform. Generally, the frequency localization of spectra is far larger than that of the noises. Thus, in most cases, it is not necessary to set an artificial threshold for frequency analysis as done in other methods. By investigation of the FL(Z), the spectra and noises can be easily distinguished by the relative magnitude of the frequency localization of the eigenvectors. This will not only make the results more reliable but also more intuitive and simpler. Nevertheless, we note that, in cases of very difficult data, a suitable decision criterion is still necessary. Although the feasibility of the method is shown in detail only by examples of the three-way fluorescence data arrays of the two-component and three-component mixtures, it is applicable to any kind of three-way data arrays and even multi-way data arrays. In addition, the method is not sensitive to the signal to noise ratio in the original data owing to the adoption of Monte Carlo preprocessing of the original data arrays. Although the proposed method is very successful in determining the number of components of mixtures in laboratory and has advantages over other methods that are based on frequency analysis, for sake of safety, the application of a combination of several different methods is still recommended in practical problems due to the complications involved in these problems such as spectral overlapping, diversity of noises and the fluorescence quenching, etc. Finally, it must be mentioned that like any other methods that are based on frequency analysis, the proposed method works best when the spectra are broad band and the noises are narrow band, which is generally the case in most situations.

Acknowledgements

The work was supported by National 863 Program of China under grant no 2007AA061502, the Natural Science Foundation of Anhui Province under grant no 090415215.

References

[1] S.R. Durell, C.H. Lee, R.T. Ross, E.L. Gross, Archives of Biochemistry and Biophysics 278 (1990) 148–160.
[2] R. Harshman, UCLA Working Papers in Phonetics 16 (1970) 1–84.
[3] R. Bro, Multi-way analysis in the food industry, PhD. Thesis, 1998.
[4] L.R. Tucker, Psychometrika 31 (1966) 279–311.
[5] Z.P. Chen, Z. Liu, Y.Z. Cao, R.Q. Yu, Analytica Chimica Acta 444 (2001) 295–307.
[6] H.P. Xie, J.H. Jiang, G.L. Shen, R.Q. Yu, Computers & Chemistry 26 (2002) 183–190.
[7] H.P. Xie, J.H. Jiang, N. Long, G.L. Shen, H.L. Wu, R.Q. Yu, Chemometrics and Intelligent Laboratory Systems 66 (2003) 101–115.
[8] L.Q. Hu, H.L. Wu, J.H. Jiang, Y.J. Ding, A.L. Xia, R.Q. Yu, Analytical and Bioanalytical Chemistry 384 (2006) 1493–1500.
[9] E. Sanchez, B.R. Kowalski, Journal of Chemometrics 4 (1990) 29–45.
[10] A.L. Xia, H.L. Wu, Y. Zhang, S.H. Zhu, Q.J. Han, R.Q. Yu, Analytica Chimica Acta 598 (2007) 1–11.
[11] L.Q. Hu, H.L. Wu, J.H. Jiang, Q.J. Han, A.L. Xia, R.Q. Yu, Talanta 71 (2007) 373–380.
[12] X.M. Tu, D.S. Burdick, D.W. Millican, L.B. McGown, Analytical Chemistry 61 (1989) 2219–2224.
[13] J.H. Wang, J.H. Jiang, J.F. Xiong, L. Yang, Y.Z. Liang, R.Q. Yu, Journal of Chemometrics 12 (1998) 95–104.
[14] T.M. Rossi, I.M. Warner, Analytical Chemistry 58 (1986) 810–815.
[15] D.A. Drabold, P.A. Fedders, Stefan Klemm, O.F. Sankey, Physical Review Letters 67 (1991) 2179;
D.A. Drabold, P.A. Fedders, Physical Review B 60 (1999) R721;
D.A. Drabold, Journal of Non-Crystalline Solids 266 (2000) 211;
R. Atta-Fynn, P. Biswas, P. Ordejon, D.A. Drabold, Physical Review B 69 (2004) 085207.