Available online at www.sciencedirect.com
SCIENCE DIRECT
Talanta 68 (2005) 108–115
Talanta
www.elsevier.com/locate/talanta

Identification of essential components of *Houttuynia cordata* by gas chromatography/mass spectrometry and the integrated chemometric approach

Cheng-Jian Xu [a], Yi-Zeng Liang [a,*], Foo-Tim Chau [b]

[a] *College of Chemistry and Chemical Engineering, Research Center for Modernization of Chinese Medicines, Central South University, Changsha 410083, PR China*
[b] *Chemometrics and Herbal Medicine Laboratory, Department of Applied Biology and Chemical Technology, The Hongkong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China*

Received 1 December 2004; received in revised form 11 March 2005; accepted 27 April 2005
Available online 25 May 2005

Abstract

Starting with Biller–Biemann's work [J.E. Biller, K. Biemann, Anal. Lett. 7 (1974) 515], various kinds of approaches have been proposed to extract GC/MS data to obtain pure components responses. In this paper, an integrated chemometric approach is proposed, which combine four sequential steps, data pretreatment, component perception, resolution and component identification, and then the proposed approach is manipulated to analyze the essential oils of a herbal medicine named *Houttuynia cordata* (HC). On the basis of the selective information obtained from both chromatograms and mass spectra, the proposed integrated chemometric approach can resolve the two-way GC/MS responses matrix into pure chromatograms and mass spectra without any model assumption on the peak shape. The resolution results obtained from HC samples demonstrate the performance of the proposed approach and indicate that it may be a promising one for analyzing complex chromatograms.

Keywords: Herbal medicine; Resolution; *Houttuynia cordata*; GC/MS data

1. Introduction

Gas chromatography hyphenated with mass spectrometry (GC/MS) is a commonly used method for characterization and identification of volatile organic compounds in complex mixtures. How to extract "pure" spectra from complex chromatograms has been a long-standing interesting topic for analytical researchers, because the acquired spectra are often "contaminated" with unresolved neighboring components arising from partial separation and also from GC septum and column bleed. Different methods have been proposed to extract pure mass spectra from complex GC/MS chromatograms. To our best knowledge, these methods can be roughly classified into four groups. The first one is mainly based on a maximization approach in mass fragmentogram profile, and it can be traced back to the Biller–Biemann algorithm [1]. In this algorithm, the mass spectrum of certain components can be reconstructed free from other interference by selecting only those mass to charge ratios and their relative intensities that are maximized at a certain scan time. Later, Dromey et al. [2,3] proposed the "model peak" method. Using the tabular peak models derived directly from the raw data, the spectra can be obtained free of other interference. Besides, Colby [4] improved the resolution of the Biller–Biemann maximization algorithm by using peak centroids. Moreover, Stein [5,6] improved the "model peak" approach and proposed an integrated method for spectrum extraction. The second group is based on analyzing the differential GC/MS data [7]. A typical method is called "backfolding" as developed by Pool et al. [8,9]. This type of methods is usually composed of two sequential steps with

* Corresponding author. Tel.: +86 731 8830831; fax: +86 731 8825637.
E-mail address: yizeng_liang@263.net (Y.-Z. Liang).

0039-9140/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2005.04.043

differentiation followed by recombination, and then background is eliminated and the chromatographic resolution is improved.

The remaining two groups are well developed by chemometricans, and the GC/MS data set is often considered as a two-way matrix. The third group is often started by a set of key spectra [10]. Then, the pure spectra and chromatogram can be obtained by alternating least squares (ALS) [11] or other transformation procedures. A number of methods have been established in this way, such as iterative target transformation factor analysis (ITTFA) [12], ALS, and elementary matrix transformation (IMEMT) [13] methods. The fourth group is usually based on local principal component analysis. The typical representations are evolving factor analysis (EFA) [14], window factor analysis (WFA) [15], heuristic evolving latent projections (HELP) [16–17] and subwindow factor analysis (SFA) [18]. These methods can guarantee to achieve the true solution when the assumption of sequential elution is met. Yet, human intervention is often needed during the resolution procedure.

In this paper, a new integrated approach is proposed to resolve complex GC/MS data set, and it is based on searching for selective information from both chromatograms and mass spectra directions. Moreover, the proposed approach is employed to analyze the essential oil of a Chinese herbal medicine called *Houttuynia cordata* (HC). HC has been used in traditional Chinese medicine (TCM) for treating infectious disease, refractory hemoptysis, malignant pleural effusion and nephrotic syndromes. Hayashi once reports its antileukemic activity and virucidal effects on human immunodeficiency virus (HIV) [19]. Now HC has been selected as one of the candidate medicines for treating sever acute respiratory syndrome (SARS) in China, and preliminary experimental results indicate HC extract kills SARS virus [20]. More information about the chemical constituents of the herb will certainly helps to understand its pharmacological activities within human body.

2. Theory

Suppose $\mathbf{X}_{m\times n}$ is a measured GC–MS data matrix of a multi-component mixture system, one wants to extract individual contribution of per component from the measurement matrix **X**. Assuming that a simple linear additive model of per responses holds, the measurement matrix $\mathbf{X}_{m\times n}$ can be decomposed into the product of two matrices as follows:

$$\mathbf{X}_{m\times n} = \mathbf{C}_{m\times k}\mathbf{S}^{\mathrm{T}}_{n\times k} + \mathbf{E} \quad (1)$$

Here, **C** represents the chromatographic profiles matrix, **S** signifies the mass spectra matrix of each component, respectively, and **E** denotes the measurement noise matrix. The superscript T denotes the matrix or vector transposition and k is number of component present in system. One would like to obtain the matrices **C** and **S** from the measurement matrix **X**. This kind of problem in mathematics is often called "the inverse problem of matrix". In general, it is impossible to obtain a unique solution under non-negative constraints on chromatographic profiles and mass spectra unless enough selective information is available.

Since the first studies on curve resolution, selectivity has been recognized as the corner-stone aspect to be considered for the recovery of the true solutions [11]. However, the selective information in GC/MS data set has not been utilized to great extent. The selectivity often lies in GC direction because of high separation ability in chromatography. The important point of using selectivity in chromatographic direction is that finding some informative time-windows. On the other hand, selective information is also possible existed in the mass spectra direction if the analyte gives a signal while co-eluted components do not.

In this paper, the selective information is extracted from the two directions of chromatographic and mass spectra space and, hence, the pure response can be easily obtained. The proposed approach used here can be expressed by the following four sequential steps.

2.1. Data pretreatment

Firstly, the raw data matrix **X** is divided into different submatrices according to the zero component regions present and then one can deal with each submatrices separately. Zero component regions are the regions, which no components elute in, and they can provide us the information of background and noise level [21]. Secondly, to eliminate the effect of background and chromatographic baseline shift in data, detection and removal of the background is often necessary. A linear background model is usually assumed over the region of scan intervals under consideration. Hence, the data are often corrected at each mass to charge ratio. If necessary, other methods, such as wavelet filter [22] or polynomial fitting [23] can also be employed. Moreover, in order to improve the detection ability and the quality of resolution of overlapped peaks with low signal-to-noise (SNR) ratio data obtained from GC/MS, smoothing technique [24] is sometimes necessary.

2.2. Component perception

After data pretreatment, one should find out that how many components are present in the subsystem under study, which is often called a chemical rank estimation problem [25–26]. The rank of the each submatrix is usually equal to the number of chemical components present in each subsystem as long as the mass spectra of each component are not exactly identical. Afterwards, one may get the elution information in the chromatographic direction by using fixed size moving window evolving factor analysis (FSMWEFA) [27]. FSMWEFA is essentially a moving principal component analysis (PCA) method, and it can tell us how many components elute in time

direction. Such elution information is important for our later resolution job.

2.3. Component resolution

If the chromatographic component meets the assumption of sequential elution that is the first appearing component will disappear firstly, then one can obtain its pure mass spectrum by SFA. The key step of using SFA is to find only one component of interest in two time-windows, named left subwindow and right subwindow. Then, the pure spectrum of the component can be obtained. The selection of two time-windows is based on the elution information achieved by FSMWEFA.

As shown in Fig. 1, component 2 lies in the two subwindows. Supposed that the submatrices corresponding to the left subwindow and the right subwindow are $\mathbf{A}_l$ and $\mathbf{A}_r$, respectively, one can decompose the submatrix $\mathbf{A}_l$ and $\mathbf{A}_r$ by singular value decomposition (SVD). Then, two groups of orthonormal base $\mathbf{E} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_g\}$ and $\mathbf{D} = \{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_h\}$ can be obtained. Here, g and h represent the number of components in the submatrix $\mathbf{A}_l$ and $\mathbf{A}_r$, respectively. The spectrum of the common component as denoted by $\mathbf{s}_c$, can be expressed by linear combination of $\mathbf{E}$ or $\mathbf{D}$ as

$$\mathbf{s}_c = \mathbf{E}\boldsymbol{\alpha} = \mathbf{D}\boldsymbol{\beta} \tag{2}$$

Here, the items of vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are linear combination coefficients. In reality, $\mathbf{E}\boldsymbol{\alpha}$ and $\mathbf{D}\boldsymbol{\beta}$ are not identical, and we instead search for vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ which minimize the squared norm $N = ||\mathbf{E}\boldsymbol{\alpha} - \mathbf{D}\boldsymbol{\beta}||^2$ under the conditions $\boldsymbol{\alpha}^T\boldsymbol{\alpha} = \boldsymbol{\beta}^T\boldsymbol{\beta} = 1$.

Since $\mathbf{E}^T\mathbf{E} = \mathbf{I}_g$ and $\mathbf{D}^T\mathbf{D} = \mathbf{I}_h$, one can get

$$\begin{aligned} N &= \boldsymbol{\alpha}^T\mathbf{E}^T\mathbf{E}\boldsymbol{\alpha} + \boldsymbol{\beta}^T\mathbf{D}^T\mathbf{D}\boldsymbol{\beta} - 2\boldsymbol{\alpha}^T\mathbf{E}^T\mathbf{D}\boldsymbol{\beta} \\ &= 2 - 2\boldsymbol{\alpha}^T\mathbf{E}^T\mathbf{D}\boldsymbol{\beta} \end{aligned} \tag{3}$$

It has been be shown that N is minimized if $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the left and right singular vectors associated with the largest singular value λ_1 of the matrix $\mathbf{E}^T\mathbf{D}$ [18]. Then, equation (3) can be changed into

$$N = 2(1 - \lambda_1) \tag{4}$$

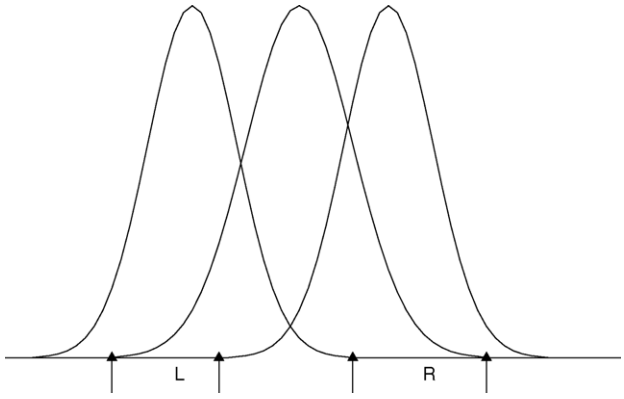

Fig. 1. A typical illustration of the subwindows of the middle peak. L: left subwindow of the peak and R: right subwindow.

Here, λ_i ($i = 1, 2, \ldots$) represent the singular values of the matrix $\mathbf{E}^T\mathbf{D}$, and $0 \le \lambda_i \le 1$. Then, the pure spectrum $\mathbf{s}_c$ can be obtained by equation (2), if the largest singular value λ_1 is close to one while the second singular value λ_2 is significantly less than one. Afterwards, the chromatographic profile of the pure component involved can be obtained by least square or other projection methods [28].

If the component does not meet the assumption of sequential elution or too strongly overlapped, that is to say, the selective information coming from chromatographic direction is not enough for resolution. One also can obtain selective information from the mass spectra direction at that time by finding out some m/z points where the analyte gives a signal and the neighboring one fails to do so. A recently proposed method named inner chromatogram projection (ICP) [29] was used here to find out selective points in mass spectra space. For the two-component embedded problem, the inner chromatographic profile and the spectrum of the outer component can be obtained easily [29]. Therefore, supposed that the normalized profiles of the inner one is $\mathbf{c}_{in}$, the $\mathbf{c}_j$ ($j = 1, \ldots, m$) is a normalized chromatographic profile under different mass to charge ratio. In this manner, one can obtain the existed selective point when f_j reaches the minimum [30]:

$$f_j = \mathbf{c}_j^T\mathbf{c}_{in} \tag{5}$$

After the pure chromatographic profile has been obtained, the pure mass spectra can be determined by the least square method.

$$\mathbf{S} = \mathbf{X}^T\mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1} \tag{6}$$

It should be noted that the ICP method has no assumption on elution order, but the selective information in mass spectra direction should be present here.

2.4. Component identification and quantitative analysis

After the mass spectra data sets of the pure components have been obtained by the above resolution step, they are usually identified via the similarity matching techniques against the standard mass spectra library. In general, the most widely used computer programs are Finnigan INCOS dot product [30] and probability based matching (PBM) algorithms [31]. Here, the dot product (cosine) method is utilized to evaluate the similarity of the standard and mass spectra under study.

$$\begin{aligned} &\mathrm{SIM} = (\mathbf{s}_{pure}, \mathbf{s}_i) = \mathbf{s}_{pure}^T\mathbf{s}_i = ||\mathbf{s}_{pure}||\,||\mathbf{s}_i||\cos\theta \\ &0 \le \theta \le \pi \end{aligned} \tag{7}$$

Here, $\mathbf{s}_{pure}$ denotes the normalized pure mass spectrum in standard library and $\mathbf{s}_i$ represents the normalized mass spectrum resolved by the above approach.

With the pure chromatographic curve and mass spectrum of each component obtained, the total two-way response of each component can be estimated by the outer product of the concentration vector and the spectrum vector of each pure

component. Then, one can determine the relative contents of each component according to their two-way response.

3. Experimental

Instruments: GC-17A Gas Chromatograph, QP-5000 Mass Spectrometer, Shimadzu.

Materials: *H. cordata* samples were collected from cultivated base of Sichuan by ourselves in September.

Extraction of volatile oil: The essential oil of the herb was prepared according to the standard extracting method in *Chinese Pharmacopoeia* (Chinese Pharmacopoeia Committee, Publishing House of People's Health, 2000, page appendix 64).

Detection of volatile oil: An OV-1 capillary column (30 m × 0.25 mm i.d.) was used. Column temperature was maintained at 50 °C for 6 min and programmed from 50 to 230 °C at 25 °C/min. Inlet temperature is kept at 280 °C. Helium carrier gas was used at a constant flow-rate of 0.7 ml/min.

Mass spectrometer: Electron impact (EI^+) mass spectra were recorded at 70 eV ionization energy in full scan mode in the 30–350 amu mass range with 0.2 s/scan velocity. The ionization source temperature was set at 280 °C.

Data analysis: Data analysis was performed on a Pentium 3 personal computer. All programs of our chemometric resolution methods were coded in MATLAB 5.3 for windows. The library searches and spectral matching of the resolved pure components were conducted on the National Institute of Standards and Technology (NIST) MS database containing about 107,000 compounds.

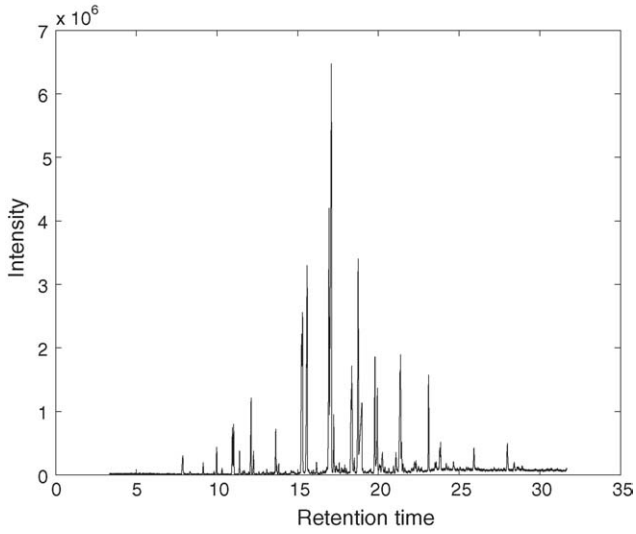

Fig. 2. The total ionic chromatogram (TIC) of the volatile oil from *Houttuynia cordata*.

4. Results and discussion

4.1. Qualitative analysis

The original HC samples that we collected from cultivated base are fresh grass, and they were immediately being extracted after collection. The total ion chromatogram (TIC) of the essential oil from HC is displayed in Fig. 2. With the presence of a large number of peaks, the figure clearly demonstrates that the volatile oil system to be studied is a very complex analytical system. Firstly, we divide the whole

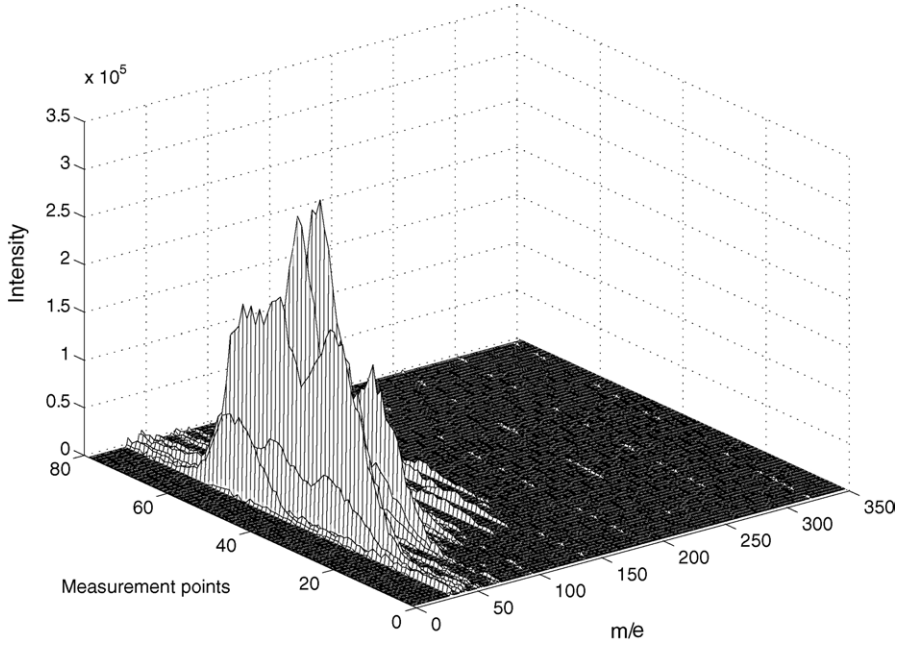

Fig. 3. The three-dimensional plot of the peak cluster A in retention time of 18.20–18.45 min.

matrix into different submatrices, and then deal with each submatrix, respectively. To illustrate how to extract the pure mass spectra efficiently, the chromatography segment in the range of 18.20–18.45 min called peak cluster A here, is taken as an example to demonstrate how our resolution algorithm works.

Our method as mentioned before employs the two-dimensional information from mass spectra and chromatograms to get the pure mass spectra. Therefore, there is no model constraint on the peak shape. Thus, the algorithm might be useful in practice. The three-dimensional plot of the cluster A is shown in Fig. 3. By looking in Fig. 3, it seems that there are three components co-eluting. The total ionic chromatogram of cluster A has been shown in Fig. 4 and a correction procedure has been implemented to reduce the background variation. The rankmap obtained by applying FSMWEFA on the cluster is depicted in Fig. 5. Here, the names of the components present in cluster A are named by their elution order. The rankmap tells that the local rank information in the elution sequence. If the local rank is one, only one component presents. But if the local rank is two, there are two components co-eluting. In this way, one can obtain the elution information of all components from the rankmap with the key assumption that the first appearing component disappears firstly also. Then, SFA can be used to extract the pure spectra directly. The key step of SFA is to find two subwindows where only one common components. For example, the pure spectra of component 2 can be obtained by selecting two subwindows in points 31–38 and 44–52, since in points 31–38 there are components 1 and 2 eluting, and in points 44–52 there are components 2 and 3 eluting. There is only one common component that is component 2. Therefore, the pure spectra of component 2 can be calculated by SFA. The reliability of spectra obtained by SFA can usually be guaranteed by observing singular values indicator. The first two singular values of matrix $\mathbf{E}^{\mathrm{T}}\mathbf{D}$, λ_1 and λ_2, are good indications to test the correctness of selecting subwindows. One can acquire the spectrum when the largest singular value λ_1 is close to 1 and the second singular value λ_2 is significantly less than 1, because there is only one common component in such case.

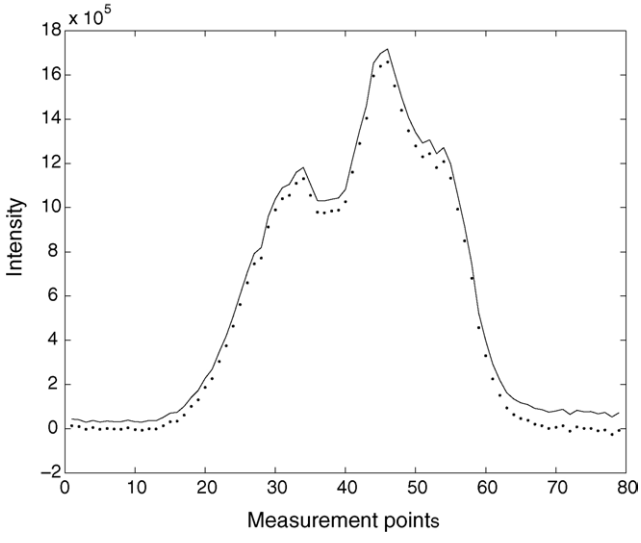

Fig. 4. Results of background correction pretreatment. Solid line: TIC of peak cluster A and dotted line: TIC of peak cluster A after background correction.

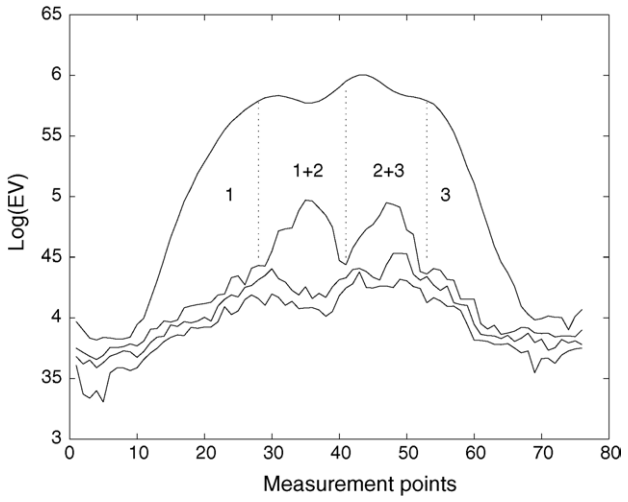

Fig. 5. The four evolving eigenvalues obtained using FSMWEFA with a window size of four for analyzing peak cluster A.

In the same way, the spectra of other components can also be obtained. Through similarity searches in the NIST MS library, the components found in cluster A may be identified. The three components, 1–3 (Fig. 6), are identified to be 2,7-dimethyl-2,6-octadien-1-ol, 1-decen-3-one, and [*S*]-3,7-dimethyl-1-octanol, respectively. The resolved and standard mass spectra of these three components are depicted in Figs. 6–8. Their pure chromatographic profiles can also be derived by the least square method and is shown in Fig. 9.

The pure response can also be obtained from the mass spectra direction. The first two components in peak cluster A are taken as an example to demonstrate how the ICP works. It is worthwhile to note that small signals may affect the resolution result owing to its low signal-to-noise ratio and should be eliminated. In practice, signals with values less than 5% of the maximum are set to be zero and eliminated. A higher threshold will not affect the resolution result as long as a relatively large selective signal exists. Suppose that the pure chromatographic profile of component 2 has already been obtained. Then, one can get the selective *m*/*e* point when f_j in equation (5) reaches the minimum. At these selective points, the pure chromatographic profile of component 1 can easily be deduced. Then, the pure spectra can be derived by equation (6). The resolved mass spectra thus obtained are shown in Figs. 6c and 7c. By observing Figs. 6 and 7, it can be seen that both SFA and ICP provide good results when the selective information is available.

The components in other overlapped peak clusters of Fig. 2 were resolved in a similar way. All the qualitative analysis

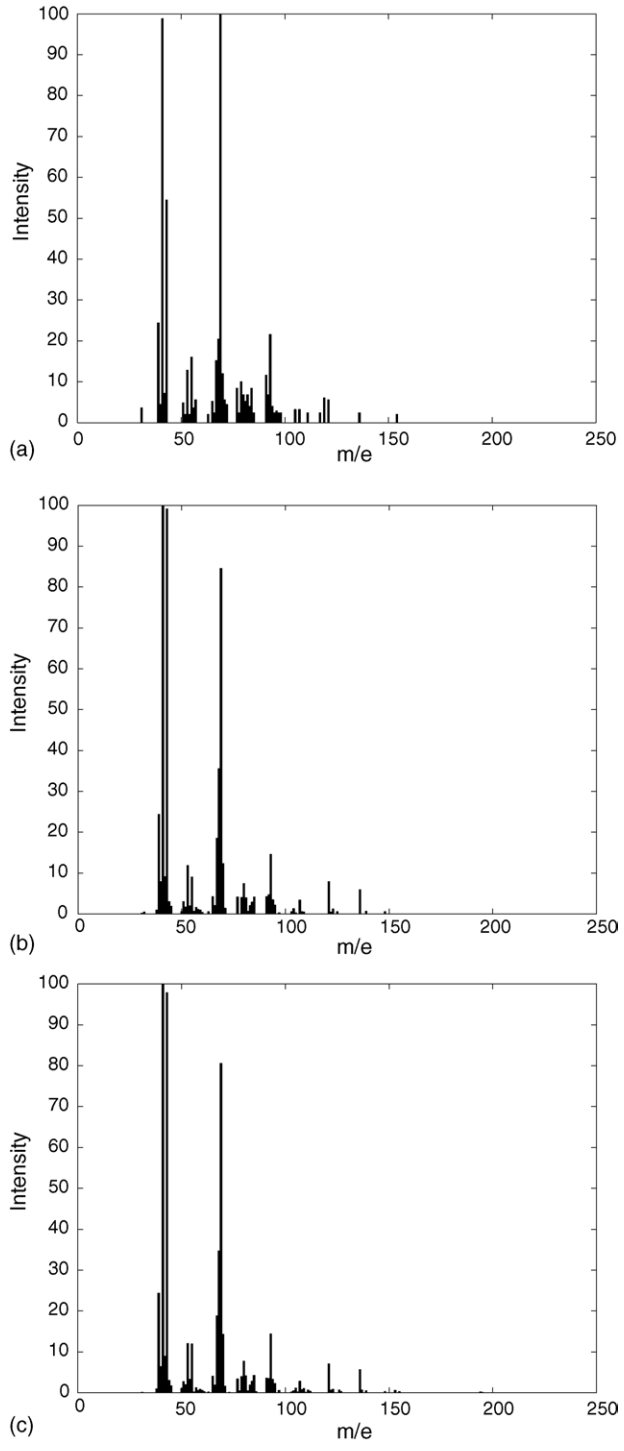

Fig. 6. (a) Standard mass spectrum of 2,7-dimethyl-2,6-octadien-1-ol from NIST library, (b) resolved mass spectrum by SFA, and (c) resolved mass spectrum by ICP.

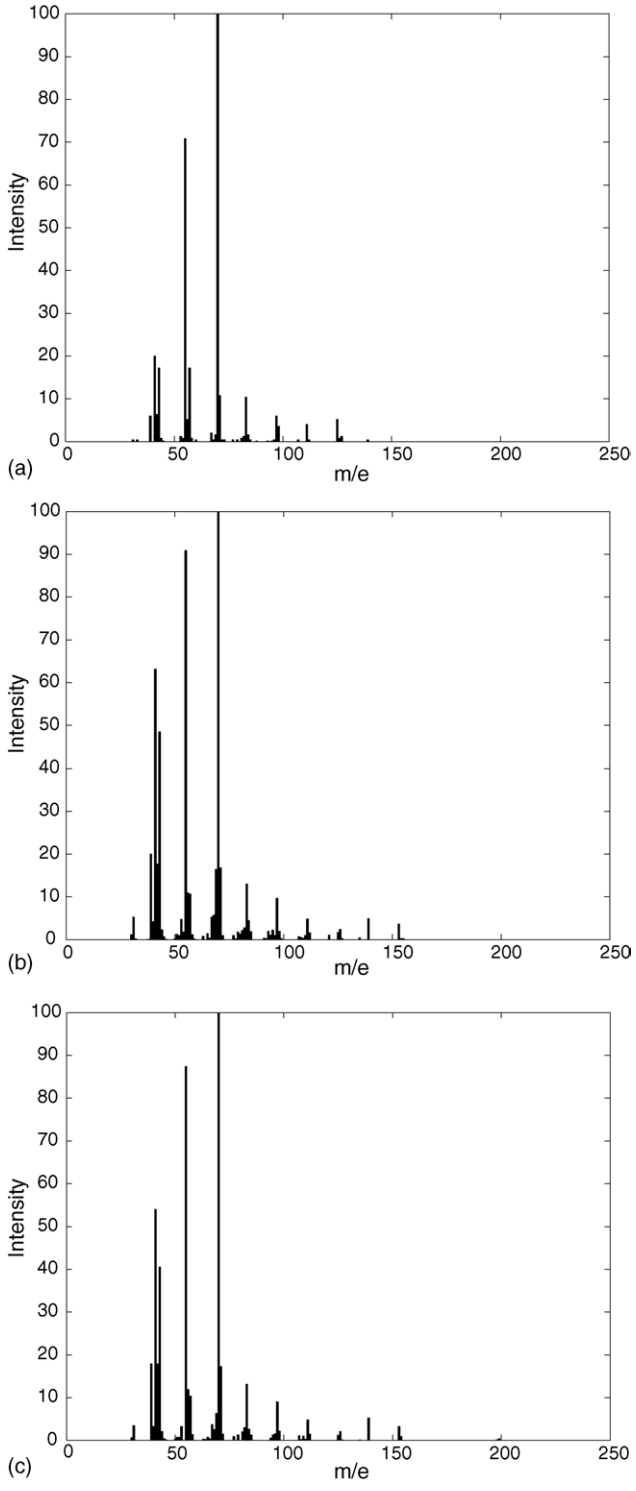

Fig. 7. (a) Standard mass spectrum of 1-decen-3-one from NIST library, (b) resolved mass spectrum by SFA, and (c) resolved mass spectrum by ICP.

results are listed in Table 1. As expected, there are still some components cannot be identified because of their low signal-to-noise ratios or limitation of the mass spectral database. The later component identification may resort to some structure elucidation system [32].

4.2. Quantitative analysis

With the pure chromatogram and mass spectrum of each component being identified, the total two-way response of each one can be obtained by the outer product of concentra-

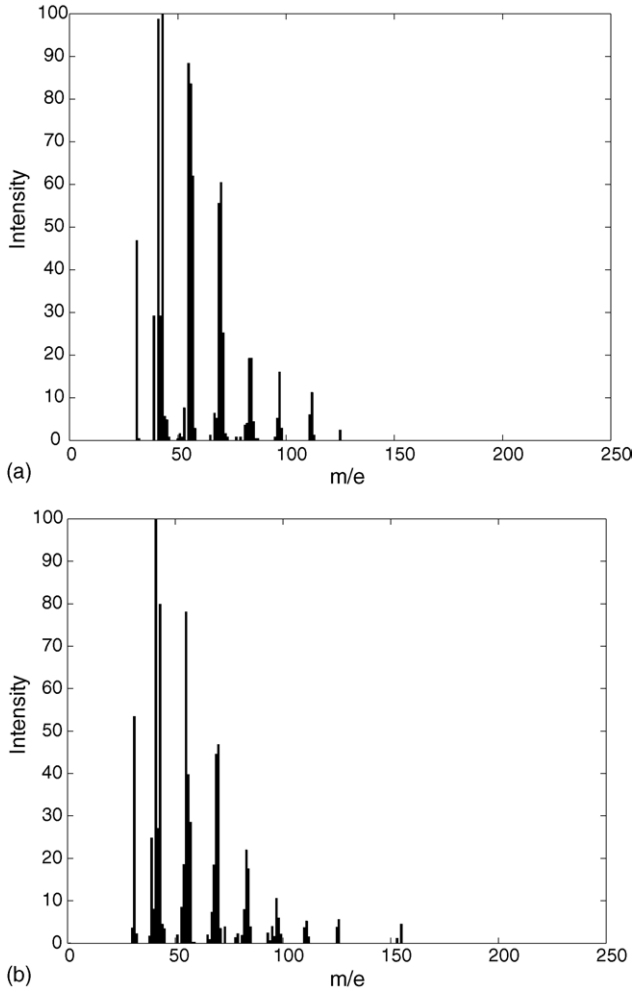

Fig. 8. (a) Standard mass spectrum of [*S*]-3,7-dimethyl-1-octanol from NIST library and (b) resolved mass spectrum obtained by SFA.

Table 1
Composition of the essential oil from *Houttuynia cordata*

Retention time (min)	Component identified	Molecular formula	Content (%)
7.88	3-Hexen-1-ol	$C_6H_{12}O$	0.81
9.15	Nonane	C_9H_{20}	0.30
9.98	2,6,6-Trimethyl-bicyclo[3.1.1]hept-2-ene	$C_{10}H_{16}$	0.79
10.31	Camphene	$C_{10}H_{16}$	0.14
10.95	Sabinen	$C_{10}H_{16}$	1.47
11.03	β-Pinene	$C_{10}H_{16}$	1.48
11.39	β-Myrcene	$C_{10}H_{16}$	0.62
12.11	3,7,7-Trimethyl-1,3,5-cycloheptatriene	$C_{10}H_{14}$	2.52
12.27	Limonene	$C_{10}H_{16}$	0.68
13.64	2-Decen-1-ol	$C_{10}H_{20}O$	1.24
13.70	Perillen	$C_{10}H_{14}O$	0.21
15.24	[*Z*]-2-Nonenal	$C_9H_{16}O$	3.50
15.30	[*Z*]-β-Terpineol	$C_{10}H_{18}O$	6.70
15.58	Capric aldehyde	$C_{10}H_{20}O$	8.90
16.95	1-Decanol	$C_{10}H_{22}O$	7.04
17.09	Methyl-*n*-nonylketone	$C_{11}H_{22}O$	22.95
17.23	Undecanal	$C_{11}H_{22}O$	1.36
18.31	2,7-Dimethyl-2,6-octadien-1-ol	$C_{10}H_{18}O$	1.92
18.35	1-Decen-3-one	$C_{10}H_{18}O$	2.51
18.38	[*S*]-3,7-Dimethyl-1-octanol	$C_{10}H_{22}O$	1.88
18.75	Lauraldehyde	$C_{12}H_{24}O$	7.28
18.98	Decanoic acid	$C_{10}H_{20}O_2$	6.33
19.78	1-Dodecanol	$C_{12}H_{26}O$	3.91
19.93	2-Tridecanone	$C_{13}H_{26}O$	2.62
21.33	Undecanoic acid	$C_{11}H_{22}O_2$	2.52
21.38	[*E*]-2-Tridecen-1-ol	$C_{13}H_{26}O$	1.16
23.11	3-Methylbutyl ester heptanoic acid	$C_{12}H_{24}O_2$	2.76
23.85	Hexadecanal	$C_{16}H_{32}O$	0.77
25.92	*n*-Hexadecanoic acid	$C_{16}H_{32}O_2$	0.84

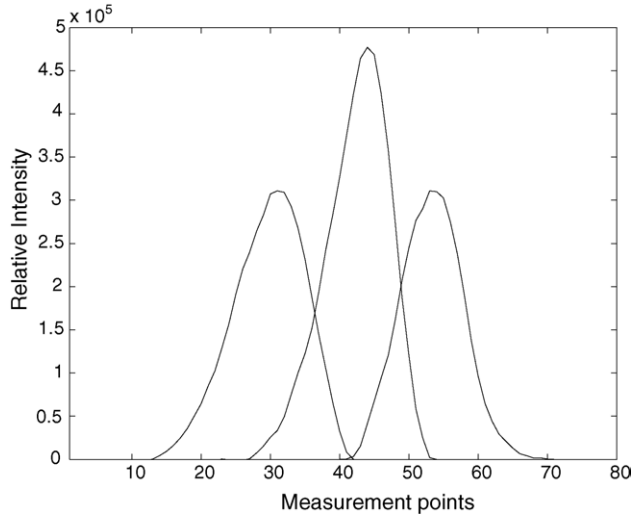

Fig. 9. The resolved chromatographic profiles in peak cluster A.

tion vector and spectra vector of each component. Table 1 lists the relative contents of the components being identified. It should be mentioned that the real concentration of each component can only be determined with when the real standard is available.

5. Conclusion

In this paper, an integrated approach has been proposed to resolve complex GC/MS data set and that of HC was used to illustrate how it works. The proposed integrated approach takes great advantage of main character of GC/MS data sets, and it make use of the selective information both from chromatographic and mass spectra direction. It seems a good way for extracting "pure" spectra for GC/MS data sets will help pharmacologists and analytical chemists to gain more information about the chemical nature of real systems with great complexity.

Acknowledgements

This work was financially supported by the National natural science foundation of the People's Republic of China (Grant No. 20235020 and 20175036) and the Area of Excellence grant by University Grants Council (UGC) of Hong Kong SAR (AoE/B-10/01).

References

[1] J.E. Biller, K. Biemann, Anal. Lett. 7 (1974) 515.
[2] R.G. Dromey, M.J. Stefik, T.C. Rindfleisch, A.M. Duffield, Anal. Chem. 48 (1976) 1368.
[3] W.F. Hargrove, D. Rosenthal, P.C. Cooley, Anal. Chem. 53 (1981) 539.
[4] B.N. Colby, J. Am. Soc. Mass Spectrom. 3 (1992) 558.
[5] S.E. Stein, J. Am. Soc. Mass Spectrom. 10 (1999) 770.
[6] Standard reference data program, National Institute of Standards and Technology, Gaithersburyg, MD, Standard Reference Database IA, Internet address: http://www.nist.gov/srd/nist1a.htm.
[7] A. Ghosh, R.J. Anderegg, Anal. Chem. 61 (1989) 2118.
[8] W.G. Pool, J.W. Leeuw, B. van de Graaf, J. Mass Spectrom. 31 (1996) 509.
[9] W.G. Pool, J.W. Leeuw, B. van de Graaf, J. Mass Spectrom. 32 (1997) 438.
[10] K.J. Schostack, E.R. Malinowski, Chemom. Intell. Lab. Syst. 6 (1989) 21.
[11] R. Tauler, A. Smilde, B. Kowalski, J. Chemom. 9 (1995) 31.
[12] P.J. Gemperline, J. Chem. Inf. Comp. Sci. 24 (1984) 206.
[13] R. Manne, B.V. Grande, Chemom. Intell. Lab. Syst. 50 (2000) 35.
[14] M. Maeder, Anal. Chem. 59 (1987) 527.
[15] E.R. Malinowski, J. Chemom. 6 (1992) 29.
[16] O.M. Kvalheim, Y.Z. Liang, Anal. Chem. 64 (1992) 936.
[17] Y.Z. Liang, O.M. Kvalheim, H.R. Keller, D.L. Massart, P. Kiechle, F. Erni, Anal. Chem. 64 (1992) 946.
[18] R. Manne, H.L. Shen, Y.Z. Liang, Chemom. Intell. Lab. Syst. 45 (1999) 171.
[19] K. Hayashi, M. Kamiya, T. Hayashi, Planta Med. 61 (1995) 237.
[20] Sci. Bull. 36 (2004) 4-4 (http://nr.stic.gov.tw/ejournal/SciNews/EJ05_v36n1.pdf).
[21] Y.Z. Liang, O.M. Kvalheim, A. Rahmani, R.G. Brereton, J. Chemom. 7 (1993) 15.
[22] H.L. Shen, J.H. Wang, Y.Z. Liang, K. Pettersson, M. Josefson, J. Gottfries, F. Lee, Chemom. Intell. Lab. Syst. 37 (1997) 261.
[23] P.J. Gemperline, J.H. Cho, B. Archer, J. Chemom. 13 (1999) 153.
[24] X.N. Li, Y.Z. Liang, F.T. Chau, Chemom. Intell. Lab. Syst. 63 (2002) 139.
[25] E.R. Malinowski, Factor Analysis in Chemistry, second ed., Wiley, New York, 1991.
[26] C.J. Xu, Y.Z. Liang, Y. Li, Y.P. Du, Analyst 128 (2003) 75.
[27] H.R. Keller, D.L. Massart, Anal. Chim. Acta 246 (1991) 379.
[28] C.J. Xu, J.H. Jiang, Y.Z. Liang, Analyst 124 (1999) 1471.
[29] Z.G. Wang, Z.P. Chen, F. Gong, H.L. Wu, R.Q. Yu, Analyst 127 (2002) 623.
[30] S. Sokolow, J. Karnofsky, P. Gustafson, The Finnigan Library Search Program: Finnigan Application Report 2, Finnigan Corp., San Jose, CA, March 1978.
[31] F.W. McLafferty, D.A. Stauffer, J. Chem. Inf. Comput. Sci. 25 (1985) 245.
[32] K. Varmuza, W. Werther, J. Chem. Inf. Comput. Sci. 36 (1996) 323.