ELSEVIER

Journal of Chromatography A, 905 (2001) 193–205

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Gas chromatography–mass spectrometry and chemometric resolution applied to the determination of essential oils in *Cortex Cinnamomi*

Fan Gong[a], Yi-Zeng Liang[b,*], Qing-Song Xu[c], Foo-Tim Chau[d]

[a]*College of Chemistry and Chemical Engineering, Institute of Chemometrics and Chemical Sensing Technology, Hunan University, Changsha 410082, PR China*

[b]*College of Chemistry and Chemical Engineering, Center of Modernization of Chinese Medicines, Central South University, Changsha 410083, PR China*

[c]*Department of Applied Mathematics, Hunan University, Changsha 410082, PR China*

[d]*Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, PR China*

Received 15 June 2000; received in revised form 11 September 2000; accepted 19 September 2000

Abstract

In this paper, a novel procedure for qualitative and quantitative analysis of the two-dimensional data obtained from GC–MS is investigated to determine chemical components of essential oils in *Cortex Cinnamomi* from four different producing areas. A new method named iterative optimization procedure (IOP) specially used to resolve embedded peaks is also developed. With the help of IOP and other chemometric techniques, such as heuristic evolving latent projections, evolving factor analysis, sub-window factor analysis and orthogonal projection resolution, and etc., the detection of the purity of chromatographic peaks can be first addressed, and then the overlapping peaks are resolved into the pure chromatogram and mass spectrum of each component. The similarity searches in the MS database are finally conducted to qualitatively determine the chemical components. The results obtained showed that the accuracy of qualitative and quantitative analysis could be greatly enhanced by chemometric resolution methods. The chemometric resolution techniques upon the two-dimensional data can be quite promising tools for the analysis of the complex samples like traditional Chinese medicine.

Keywords: Chemometrics; Oils; Essential oils; Cortex cinnamomi; Pharmaceutical analysis; Ethylxylene; Cineole; Capaene; Cinnamic acid; Cinnamyl acetate; Bisabolol; Methoxycinnamaldehyde; Cadinol; Eudesmol

1. Introduction

Traditional Chinese medicines (TCMs) have been used for thousands of years in China on account of the advantages of definite pharmacological activities and low toxicity [1]. However, it remains difficult to establish some simple, convenient and efficient techniques to determine the quality of TCMs as there always exist hundreds of chemical components in them. Surely, the determination of such complicate systems is not a trivial work and is a challenging problem to pharmacologists and analytical chemists.

*Corresponding author. Tel.: +86-731-882 2841; fax: +86-731-882-5637.

E-mail address: yzliang@public.cs.hn.cn (Y.-Z. Liang).

0021-9673/01/$ – see front matter
PII: S0021-9673(00)00976-6

Developing new instruments and relative techniques is urgently necessary to realize this intention. With the advent of hyphenated instruments in the latest decades, such as GC–MS, HPLC–MS, HPLC–DAD (diode array detection), HPCE–DAD, etc., two-dimensional data, which can provide information both in the chromatographic and spectral directions in one run, become available. Thus, the qualitative and quantitative identification of components can be performed not only with retention times, peak heights and areas but also UV and/or mass spectra. The chemometric resolution approaches addressing for the analysis of the two-dimensional data have been developed quickly to meet such a need for hyphenated instruments recently [2–12]. In general, the data from the hyphenated instruments have several advantages, such as non-negative spectra, non-negative and unimodal chromatographic profiles, the peak that elutes first will disappear first in chromatographic direction, or say simply ''first-in–first-out''. These features make it possible to resolve the chemical components involved in a complex analytical system in a successive way. Many evolving resolution techniques upon two-dimensional data, for example, heuristic evolving latent projections (HELP) [2–4], evolving factor analysis (EFA) [5–7], windows factor analysis (WFA) [8,9], sub-window factor analysis (SFA) [10,11] and orthogonal projection resolution (OPR) [12], have progressed rapidly in recent years. As a result, hyphenated instruments combined with chemometric resolution can make it possible to quantify one complicate ''black'' analytical system clearly and accurately. The methods have been successfully applied to many fields [13–16].

However, there often exists a special case called embedded peaks in chromatography, in which the assumption of ''first-in–first-out'' will not be held, in the complex analytical systems, such as the samples from traditional Chinese medicine. This will make difficulty for the above resolution techniques [2–12] to resolve them [17]. How to treat this particular elution pattern still remains in the air. A procedure developed in [18] claimed to resolve the embedded peaks in hyphenated chromatography. However, the crucial assumption for this procedure that the chromatographic peak of a pure component is locally symmetric might be too strong for all the chromatographic peaks to follow in real samples like TCMs.

Cortex Cinnamomi is a widely used TCM in China as it can contribute to treatments of some diseases [19,20]. The essential oil, which is the main efficacious constitution of *Cortex Cinnamomi*, always contains about one hundred chemical components and so form a very complex system. In general, the determination of the essential oil of *Cortex Cinnamomi* is usually conducted with GC [21–23], GC–MS [24] and HPLC [25]. However, as for above GC or HPLC methods, only few components, for instance, cinnamaldehyde, eugenol and so on, are determined and the identification of chromatographic peaks is just carried out by means of comparison of retention time with reference compounds. If there is not the standard substance of the component studied, it does not work at all. For the technique in [24], much more components are qualitatively and quantitatively analyzed, but the determination is performed only through the direct similarity searches in the MS database attached to the GC–MS instrument. There exist at least two serious problems for this approach. First, the background can not accurately corrected. Second, there are always overlapping peaks and even embedded peaks for this complicated system even under good separating conditions. Both of the above problems can possibly result in wrong similarity matches in the MS library, not to saying of obtaining very accurate qualitative and quantitative results.

In this paper, the essential oils in *Cortex Cinnamomi* from four different producing areas are detected with GC–MS, followed by the resolution of the two-dimensional data with OPR algorithm, HELP technique, SFA approach and the iterative optimization procedure (IOP) proposed in this study. Then, qualitative identification of these components is conducted with retention times and pure mass spectra. Finally, the quantitative analyses are carried out with the overall volume integration method [14–16]. Possible pharmacological activities of the essential oils in *Cortex Cinnamomi* from four different producing areas will be studied in our future work in order to further compare the differences of essential oils in chemical components and their pharmacological activities

2. Theory

The basic procedures of HELP, OPR, SFA and other chemometric resolution methods have been extensively documented in [2–12]. As for the HELP method, the chromatogram and spectrum are obtained by means of so-called full rank resolution technique after the determination of zero-component and selective regions of the target component. The OPR method first constructs a projection matrix from the zero-component region(s) followed by obtaining the concentration curve of the component investigated. The SFA approach directly extracts the spectrum from overlapping structures without first resolving the concentration profile in comparison with the subwindows which have only one eluting component in common. For the brevity of this paper, the details of these ways are not restated here. The following is just a brief description of the IOP applied to the resolution of embedded peaks.

Data analysis is performed on two-dimensional data, such as those produced by HPLC–DAD and/or GC–MS, etc. Four distinct embedded cases are shown in Fig. 1. The fourth case with no overlap at the peak maximum of the embedding peak can not be solved by the presented method. This situation is, however, easily detected [26]. The cases discussed in this paper are the first and third ones, which are the most common cases in practice. In fact, they are equivalent to each other in mathematical principle.

The data **X** discussed here are a matrix with m rows that correspond to spectra taken at regular time intervals and n columns which represent chromatograms measured at successive wavelengths. For a binary mixture, $m \times n$ matrix **X** can be expressed as two sums of the bilinear matrices:

$$\boldsymbol{X} = \boldsymbol{c}_1\boldsymbol{s}_1^{\,T} + \boldsymbol{c}_2\boldsymbol{s}_2^{\,T} \tag{1}$$

Here $\{\boldsymbol{c}_k = (c_{k1}, c_{k2}, \ldots, c_{km})^T, k = 1, 2\}$, $\{\boldsymbol{s}_k^T = (s_{k1} s_{k2}, \ldots, s_{kn})^T, k = 1, 2\}$ are the unknown chromatographic and spectra vectors of two components. On the other hand, the data might also be expressed as row vectors in time direction:

$$\boldsymbol{x}^T(t) = c_1(t)\boldsymbol{s}_1^{\,T} + c_2(t)\boldsymbol{s}_2^{\,T} \ (t = 1, 2, \ldots, m) \tag{2}$$

where $\boldsymbol{x}^T(t)$ is the spectral vector of mixture at time t, $c_1(t)$ and $c_2(t)$ are the values of chromatographic vectors $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ at time t.

These spectral vectors can be differentiated in the time direction:

$$\mathrm{d}\boldsymbol{x}^T(t)/\mathrm{d}t = \boldsymbol{s}_1^T\,\mathrm{d}c_1(t)/\mathrm{d}t + \boldsymbol{s}_2^T\,\mathrm{d}c_2(t)/\mathrm{d}t \tag{3}$$

Note that every derivative vector $\mathrm{d}\boldsymbol{x}^T(t)/\mathrm{d}t$ is still a combination of spectral vectors of pure components. Eq. (3) reveals a useful feature of derivative vectors that as t moves towards t^* (Fig. 2), the maximum point of the embedding peak, $\mathrm{d}c_1(t)/\mathrm{d}t$ varies from a

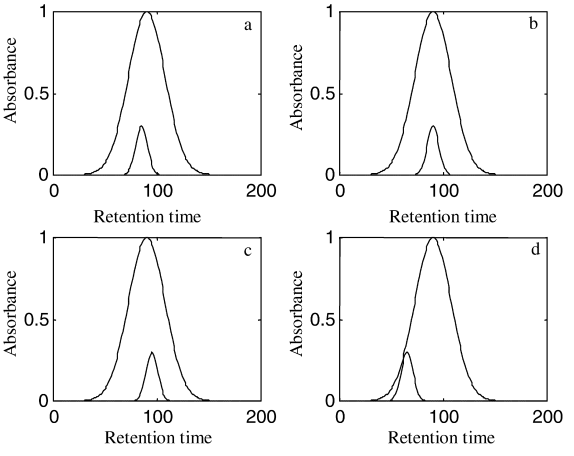

Fig. 1. Four different elution patterns for a chromatographic peak embedded in another one. The second and fourth patterns can not be resolved by the procedure proposed in this paper. The pattern a is equivalent to the pattern c mathematically.

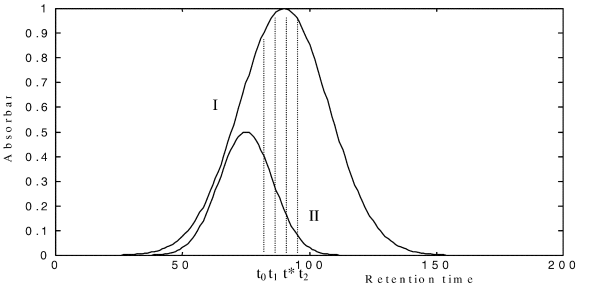

Fig. 2. The first elution pattern with the embedding (I) and embedded (II) peaks. t_0 is the retention time at which the spectrum of mixture reaches its maximum. t_1 is the retention time at which every element of vector $\mathrm{d}\boldsymbol{x}^T(t)/\mathrm{d}t$ is nonpositive. t^* is the maximum point of the embedding peak. t_2 is the retention time at which the rank of $\boldsymbol{H}(t)$ in Eq. (5) is equal to 1.

positive value to zero. During the same time, the mixture spectrum $d\boldsymbol{x}^T(t)/dt$ varies from combination of $\boldsymbol{s}_1^T$ and $\boldsymbol{s}_2^T$ to pure $\boldsymbol{s}_2^T$. The vector $d\boldsymbol{x}^T(t)/dt|_{t=t^*}$ is a vector parallel to $\boldsymbol{s}_2^T$.

The resolving of the embedded peaks can be fulfilled by the following iterative procedure:

(1) Initially, two retention time points (t_0, t_1) in the neighborhood of the embedding peak maximum are determined. First, t_0 should satisfy the following equation:

$$\boldsymbol{X}(t_0, j_0) = \max_{t,j} (\boldsymbol{X}(t, j)) \tag{4}$$

and t_1 is the retention time at which every element of vector $d\boldsymbol{x}^T(t)/dt$ is nonpositive, which is a point between t_0 and t^*, say $t_0 \le t_1 \le t^*$ (see Fig. 2). Obviously, t_1 is very close to t^*, and it can be thought as the first estimate of t^*.

(2) The retention time point t_2 corresponding to t_0 is then determined by the rank analysis of matrix $\boldsymbol{H}(t)$:

$$\boldsymbol{H}(t) = \begin{pmatrix} (d\boldsymbol{x}^T(t)/dt)_{t=t_1} \\ \boldsymbol{x}^T(t_0) - \boldsymbol{x}^T(t) \end{pmatrix} = \begin{pmatrix} c'_1(t_1) & c'_2(t_1) \\ c_1(t_0) - c_1(t) & c_2(t_0) - c_2(t) \end{pmatrix} \begin{pmatrix} \boldsymbol{s}_1^T \\ \boldsymbol{s}_2^T \end{pmatrix} \quad (t = t_1 + 1, \ldots, m) \tag{5}$$

In general, the rank of the matrix ($\boldsymbol{s}_1$, $\boldsymbol{s}_2$) is 2. If the rank of matrix $\boldsymbol{H}(t)$ becomes 1, then the rank of matrix:

$$\begin{pmatrix} c'_1(t_1) & c'_2(t_1) \\ c_1(t_0) - c_1(t) & c_2(t_0) - c_2(t) \end{pmatrix}$$

is 1, which suggests the corresponding elements of two rows of this matrix have same ratio. If $t_1 = t^*$ and the rank of $\boldsymbol{H}(t)$ is 1, then $c_1'(t_1) = 0$ and $c_1(t_0) = c_1(t)$.

The retention time point t_2 is selected that makes the rank of $\boldsymbol{H}(t)$ decrease from 2 to 1. Generally, $t_1 < t^*$, and one can see from Fig. 2 that $c_1'(t_1) > 0$, $c_2'(t_1) < 0$:

$$d\boldsymbol{x}^T(t)/dt|_{t=t_1} = c_1'(t_1)\boldsymbol{s}_1^T + c_2'(t_1)\boldsymbol{s}_2^T \tag{6}$$

while:

$$\boldsymbol{x}^T(t_0) - \boldsymbol{x}^T(t) = (c_1(t_0) - c_1(t))\boldsymbol{s}_1^T + (c_2(t_0) - c_2(t))\boldsymbol{s}_2^T \tag{7}$$

$c_2(t_0) - c_2(t)$ is always positive for $t = t_1 + 1, \ldots, m$. Thus, if the rank of matrix $\boldsymbol{H}(t_2)$ is 1, then $c_1(t_0) - c_1(t_2)$ must be negative, that is, $c_1(t_0) < c_1(t_2)$.

From above analysis, it is reasonable to determine the last estimate of t^* by the rank analysis of matrix $\boldsymbol{G}(t)$:

$$\boldsymbol{G}(t) = \begin{pmatrix} d\boldsymbol{x}^T(t)/dt \\ \boldsymbol{x}^T(t_1) - \boldsymbol{x}^T(t_2) \end{pmatrix} = \begin{pmatrix} c'_1(t) & c'_2(t) \\ c_1(t_1) - c_1(t_2) & c_2(t_1) - c_2(t_2) \end{pmatrix} \begin{pmatrix} \boldsymbol{s}_1^T \\ \boldsymbol{s}_2^T \end{pmatrix} \quad (t = t_1 + 1, \ldots, m) \tag{8}$$

Here the primes represent the differentiation. If $t^\times$ makes the rank of matrix $\boldsymbol{G}(t^\times)$ drop from 2 to 1, then, one takes $t^\times$ as the final estimate of t^*.

After the maximal point of the embedding peak is located, the pure spectrum of embedded component can directly be obtained from the derivative at the maximal retention time [26]. But derivative at the maximal point usually changes fast and this point is difficult to determine, if the chromatographic peak is sharp. Furthermore, the random noise greatly influence this determination on account of derivation transformation. Thus, the marker-object projection technique [29] is used here to determine the chromatogram of the embedding component. The spectrum of the embedding component is taken as a ''marker object'', since its spectrum, say $\boldsymbol{s}_1$, could be easily obtained in the selective region. Then, the chromatogram at each retention time is projected on this spectrum:

$$\boldsymbol{X}\boldsymbol{s}_1 = (\boldsymbol{c}_1\boldsymbol{s}_1^T + \boldsymbol{c}_2\boldsymbol{s}_2^T)\boldsymbol{s}_1 \tag{9}$$

When $\boldsymbol{s}_1$ is normalized, $\boldsymbol{s}_1^T\boldsymbol{s}_1 = 1$. Let $\boldsymbol{s}_2^T\boldsymbol{s}_1 = \alpha$ then Eq. (9) can be written as:

$$\boldsymbol{X}\boldsymbol{s}_1 = \boldsymbol{c}_1 + \alpha\boldsymbol{c}_2 = \boldsymbol{m} \tag{10}$$

That is,

$$\boldsymbol{c}_1 = \boldsymbol{m} - \alpha\boldsymbol{c}_2 \tag{11}$$

where $\boldsymbol{c}_1$ is a vector function of α. Optimization with maximum point of $\boldsymbol{c}_1$ as objective function can

provide the pure chromatogram of embedding component $\boldsymbol{c}_1$. In practice, the position of the maximum of $\boldsymbol{c}_1$ is not very sensitive to variable α, that is, different estimates of $\boldsymbol{c}_1$, could provide the maximum at the same retention time. Thus the average of these estimates is taken as the final result.

With estimates of $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ available, the pure spectra for the two components can consequently be obtained by least squares:

$$\boldsymbol{S}^T = (\boldsymbol{C}^T\boldsymbol{C})^{-1}\boldsymbol{C}^T\boldsymbol{X} \quad (12)$$

where $\boldsymbol{S}^T = (\boldsymbol{s}_1, \boldsymbol{s}_2)^T$, $\boldsymbol{C}^T = (\boldsymbol{c}_1, \boldsymbol{c}_2)^T$.

The iterative optimization procedure is as follows:

(1) Locate the selective region of the embedding component by HELP or other chemometric resolution methods and then obtain the pure spectrum $\boldsymbol{s}_1$.

(2) Estimate the chromatogram $\boldsymbol{c}_2$ of the embedded component by OPR.

(3) Locate t_0, t_1, t_2 by iterative rank analysis and get the estimate of maximum retention time $t^{\times}$ of the embedding peak.

(4) Determine the chromatogram $\boldsymbol{c}_1$ of the embedding component by marker-object optimization, and finally the pure spectrum of the embedded component by the least-squares method.

3. Experminental

3.1. Materials

Cortex Cinnamomi from four different producing areas (Zhaoqing, Guangdong province, China; Yulin, Guangxi province, China; Yunnan province, China and Vietnam) is purchased from a pharmaceutical store, and identified by a professor from Institute of Materia Medica, Hunan Academy of Traditional Chinese Medicine and Materia Medica, Changsha, Hunan, China.

3.2. Extraction of essential oil

Cortex Cinnamomi was dried for about 60 min under 40°C at first. About 200 g pre-weighed *Cortex Cinnamomi* was then swollen with over 1000 ml of distilled water in a standard extractor for extracting volatile oil and allowed to stand for 30 min at room temperature. Next, more (about 100 ml) distilled water was added. Then, the essential oil was prepared according to the standard extracting method for the volatile oil in TCMs in the Chinese Pharmacopoeia [27]. The yield of essential oils in *Cortex Cinnamomi* from four different producing areas is shown in Table 1.

3.3. Instruments

A Shimadzu GC-17A gas chromatograph and a Shimadzu QP-5000 mass spectrometer were employed in this study.

3.4. Detection of volatile oil

In the gas chromatographic system, an OV-17 capillary column (30 m×0.25 mm I.D.) was used. Column temperature was first set at 60°C, and then programmed from 60°C to 270°C at the rate 15°C min^{-1}. Inlet temperature was kept at 250°C. Helium carrier gas was used at a constant flow-rate of 1 ml min^{-1}. In the mass spectrometer, electron impact (EI^+) mass spectra were recorded at 70 eV ionization energy in full scan mode in the 20 to 350 mass range with 0.2 s $scan^{-1}$ velocity. The ionization source temperature was set at 230°C.

Table 1
Yield of essential oils and content of cinnamaldehyde and cinnamic acid of essential oils in *Cortex Cinnamomi* from different producing areas

	Zhaoqing	Yulin	Yunnan	Vietnam
Yield (%)	0.43	0.34	0.55	1.06
Cinnamaldehyde (g ml^{-1})	1.2339	0.9813	1.4773	1.5232
Cinnamic acid (g ml^{-1})	0.0335	0.0272	0.0026	0.0039

3.5. Data analysis

Data analyses were performed on a Pentium-based IBM compatible personal computer. All programs of the chemometric resolution methods were coded in MATLAB 5.1 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 62 000 compounds.

4. Results and discussion

4.1. Resolution of a simulated two-dimensional data

Curves I and II (solid lines) in Fig. 3a represent the simulated chromatograms of embedding and embedded components from two-dimensional data, respectively. Their corresponding spectra are shown in Fig. 3b (solid lines). From this figure, one can see that component I (embedding) elutes in first and elutes out secondly while component II (embedded) elutes in secondly and elutes out first. Thus, this two-dimensional data can not be resolved by means of commonly used chemometric resolution approaches. However, with the help of the IOP, the data can be uniquely resolved into pure chromatograms and spectra of two components. The resolved chromatograms and spectra of the embedding and embedded components (see dotted lines in Fig. 3a and b) are very close to real ones.

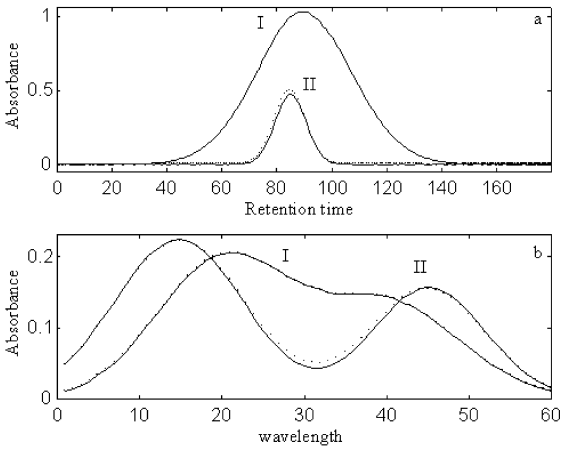

Fig. 3. Simulated and resolved chromatograms (a) and spectra (b) of the embedding peak I and the embedded peak II. The solid lines denote the simulated chromatograms and spectra while the dashed lines represent the resolved chromatograms and spectra by the IOP.

4.2. Resolution of real two-dimensional data from essential oils of Cortex Cinnamomi

Fig. 4 shows the real total ion chromatograms (TICs) of essential oils in *Cortex Cinnamomi* from four producing areas. Seen from these profiles, about

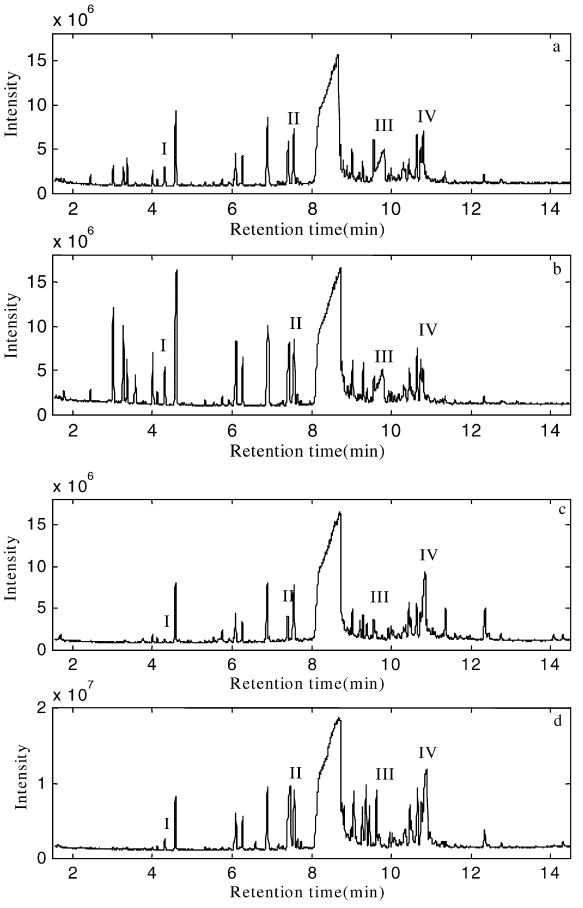

Fig. 4. Total ionic chromatograms of essential oils in *Cortex Cinnamomi* from four different producing areas. (a) Zhaoqing, Guangdong province, China.; (b) Yulin, Guangxi province, China.; (c) Yunnan province, China.; (d) Vietnam.

30, 32, 26 and 30 chromatographic peaks appear in Fig. 4a–d, respectively and parts of the peaks are overlapped with each other. However, the similarity indices (SIs) obtained from direct searching with the NIST MS database are quite low for many chromatographic peaks. On the other hand, the same component is possibly searched at different chromatographic scan points. All these indicate the great complexity of these analyzed systems, or saying concretely, most of the chromatographic peaks are overlapped. If these overlapping peaks are not resolved, the simple search with the database will definitely fail, since the mass spectrum of mixtures measured can never get a good matching index with that of a pure component in the NIST MS database. Furthermore, especially for the component with low content, it is also very difficult to be identified correctly with the database, since a two-dimensional data obtained by mass spectral measurement unavoidably contains peaks associated with column background and residual gases. Without background correction, both the resolution of the overlapped peaks and the identification of the component with low content are impossible. For commercial GC–MS system, background subtraction is usually performed as follows. First, a scan point, which only contains the background mass spectrum, is subjectively found. Next, the intensities of the same inter mass numbers appearing in the target and background spectra are subtracted, and so the practical target mass spectrum is obtained. Obviously, the practical target mass spectrum strongly depends on the selection of the background point. If this selection is wrong, different target mass spectra may be obtained. The genuine MS to be searched is surely confused with MS subtracted. As for HELP, EFA, OPR or SFA resolution methods, the local rank analysis of the zero-component regions, which contain no components eluting, before elution of the first chemical component starts and after the last chemical constituent has eluted, can together provide sufficient information for accurately correcting a drifting base line [2–4]. Much better background subtraction could be obtained. After background subtraction, the resolution of the overlapping peaks becomes possible. Now, four peak clusters named I (within 4.20–4.40 min), II (within 7.30–7.50 min), III (within 9.40–9.60 min) and IV (within 10.68–10.80 min) are taken as examples in Fig. 4 (1) to illustrate the data analysis procedure.

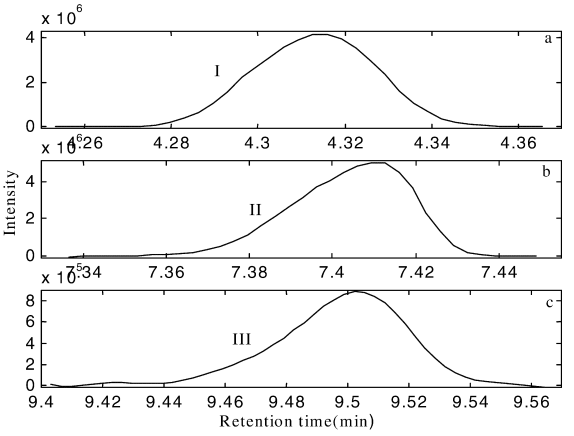

Fig. 5. Total ionic chromatograms of peak clusters I (a), II (b) and III(c) in Fig. 4a.

Fig. 5 shows the TIC curves of peak clusters I, II and III in Fig. 4a. Seen from Fig. 5, all these peak clusters look like one-component chromatographic peaks. Using library search of NIST MS database built in the instrument, *p*-cineole, capaene and cinnamic acid can be targeted for these three peaks shown in Fig. 5a–c, respectively. Therefore, all the peak clusters I, II and III can be easily regarded as one-component peaks in a classic analytical way. However, this conclusion is unfortunately wrong. If chemometric resolution methods, such as EFA, HELP, OPR, SFA and IOP, are applied to the corresponding two-dimensional data, two components named 4-ethyl-*o*-xylene and *p*-cineole in peak cluster I, one component named capaene in peak cluster II and two components named cinnamic acid and cinnamyl acetate in peak cluster III can be resolved. How the results become available will be described in detail in the following part.

Detection of peak purity of the two-dimensional data can be conducted by fixed size moving window evolving factor analysis (FSMWEFA) [28] or so-called eigenstructure tracking analysis [4]. In an FSWMEFA plot, the curve of logarithm values of eigenvalue higher than the noise level represents the appearance of components. If a studied system contains only one species, only one curve is higher than the noise level in its FSWMEFA plot. Other-

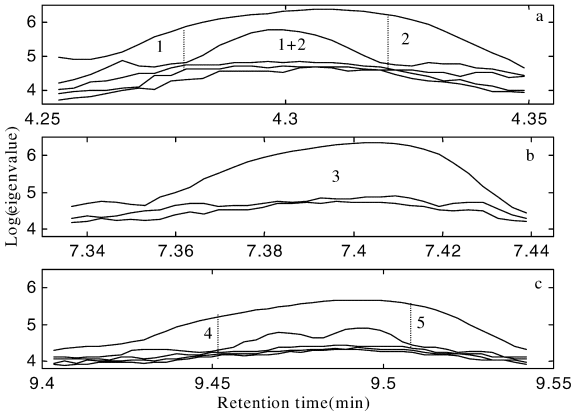

Fig. 6. Fixed size window method (FSWM) plots of peak clusters I (a), II (b) and III (c) in Fig. 4a. The horizontal dashed lines represent the noise level. The regions 1, 1+2, 2 in (a) denote the pure region of the component 1, overlapping region of the components 1 and 2, the pure region of the component 2, respectively. The region 3 in (b) denotes the pure region of the component 3. The regions 4, 5 in (c) denote the pure regions of components 4 and 5. In practice, the components 4 and 5 are the same component. The middle region between regions 4 and 5 in (c) is the overlapping region of the components 4 or 5 and another one whose pure region does not exist.

wise, there are at least two curves higher than the noise level. From the FSWMEFA plots of peak cluster I shown in Fig. 6a, there are obviously two curves higher than the noise level at peak region. It suggests that the peak cluster I is certainly not a pure one and the region of 1+2 is the overlapping region of the first and second components. The regions of 1 and 2 show the pure regions of the first and second components, respectively. While, in Fig. 6b only one curve is higher than the noise level and so peak cluster II may only contain one component. For peak cluster III, there are two curves higher than the noise level (see Fig. 6c) but of different shape from peak cluster I. Anyway, there may possibly be at least two components within this region. It is interesting to compare the mass spectra in the regions 4 and 5, one can find they are quite similar, just like the spectra from the same component. This suggests that the cluster III might be the embedded peaks mentioned above.

In order to further confirm the conclusion obtained from FSWMEFA, the evolving latent projection graphs (ELPGs) [2–4] are also plotted. The ELPG is actually a principal component projective plot. In the ELPG from the chromatographic direction, the straight-line section represents the pure selective region of one component, while the curving section denotes the overlapping region containing at least two components. The ELPGs of peak cluster I, II and III, are shown in Fig. 7. From Fig. 7a, one can see that peak cluster I is a two-component system. The two straight lines (marked by 1 and 2) indicate clearly the selective information from two different components, which strongly supports the result obtained from the FSMWEFA. While, in Fig. 7b one can only see one straight line (marked by 3) with some noise fluctuation. It suggests that there be only one component in peak cluster II as only the first principal component presents the signal and the second one may be just the random noises. The most interesting plot is shown in Fig. 7c. The first straight line (marked by 4) suggests selective information of a pure component, while the middle curve part indicates the mixture of two components. However,

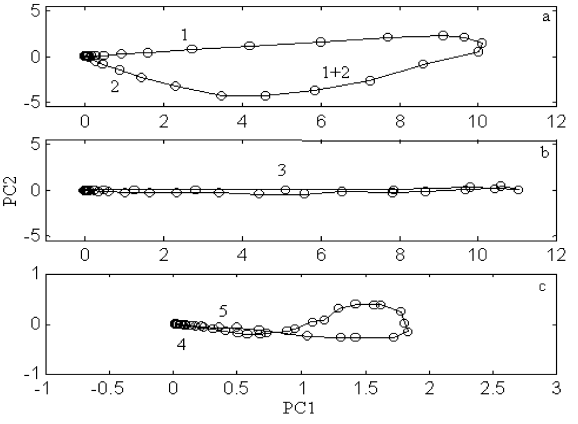

Fig. 7. Evolving latent projection graphs (ELPGs) of peak clusters I (a), II (b) and III (c) in Fig. 4a. The straight lines 1 and 2 in Fig. 7a denote the pure region of the component 1 and 2 while the curve 1+2 represents the overlapping region of the components 1 and 2. The straight line 3 with some noise fluctuation in Fig. 6b denotes the pure region of the component 3. The straight lines 4 and 5 in Fig. 6c denote the pure regions of components 4 and 5. In practice, the components 4 and 5 are the same component. The middle curve between straight lines 4 and 5 in Fig. 6c represents the overlapping region of the components 4 or 5 and another one whose pure region does not exist.

the last straight line (marked by 5) goes back to the first straight line again. It is a typical elution pattern of the embedded peaks [26]. Thus, the results obtained from FSMWEFA and ELPGs are consistent with each other.

Based on the information discussed above, the resolution of these three peak clusters can be conducted differently. For the peak cluster I, the simple least squares can be directly used to resolve the overlapping peaks, since one can easily pick up the two pure spectra from the selective information regions. But, for the peak cluster III, one have to use the IOP method described in the theoretic part. The resolved chromatograms and mass spectra are shown in Figs. 8 and 9, respectively. In order to further confirm the results obtained, one specially focuses one's attention on the peak cluster III. All the chromatograms at every m/z point are shown in Fig. 10. From this plot, one can easily see that they really contain some embedded peaks from some selective m/z points.

Fig. 11a shows the TIC of the peak cluster IV. Seen from Fig. 11a, this peak cluster might be more complicate than above peak clusters. Likewise, the two-dimensional data involved the peak cluster IV can be resolved into pure chromatograms and mass spectra. The results of the resolved chromatograms obtained are shown in Fig. 11b. Their corresponding spectra are mainly shown in Fig. 12. For the brevity

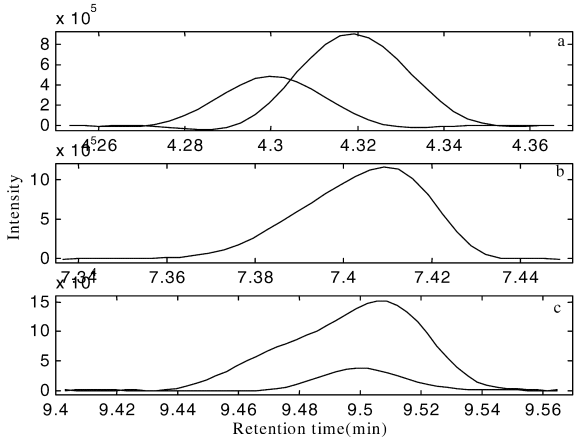

Fig. 8. Resolved Chromatograms of peak clusters I (a), II (b) and III (c) in Fig. 4a.

of the paper, the procedure about how to resolve this peak cluster has to be dropped. Anyway, the readers who are interested in the resolving procedure for overlapping peaks of multi-components can refer to the reference [16].

4.3. Qualitative determination of chemical components of essential oils in Cortex Cinnamomi

As the pure chromatographic curve and mass spectrum of each component have been resolved, the qualitative analyses can then be directly performed by means of similarity searches in the NIST mass library now. It shows that the components in peak cluster I, II and III are 4-ethyl-*o*-xylene and *p*-cineole, capaene, cinnamic acid and cinnamyl acetate, respectively. Their resolved mass spectra together with the standard spectrum of each component from the NIST MS library are given in Fig. 9. From the plot, one can see that the resolved mass spectra are almost the same as the corresponding standards. This also explains that the resolution results obtained are reliable.

As for the peak cluster IV, with the help of the same similarity searches, four of these resolved components, marked by 6, 7, 8, and 10 in Fig. 11b, are δ-cadinol, *o*-methoxycinnamaldehyde, bisabolol and eudesmol, respectively. While, the component marked by 9 can not be accurately determined, since there might be not the spectra of the corresponding component in the database. However, one can possibly deduce that its molecular formula may be $C_{15}H_{26}O$ with molecular mass 222. Their resolved chromatograms and mass spectra are showed in Figs. 11b and 12.

4.4. Quantitative analysis of chemical components of essential oils in Cortex Cinnamomi

As two-dimensional data from GC–MS can be resolved into pure chromatograms and mass spectra by means of chemometric resolution techniques, the peak area integration at every m/z point of each component can be easily calculated. Its sum, which is called the overall volume, is directly proportional to the content of the corresponding component [14–

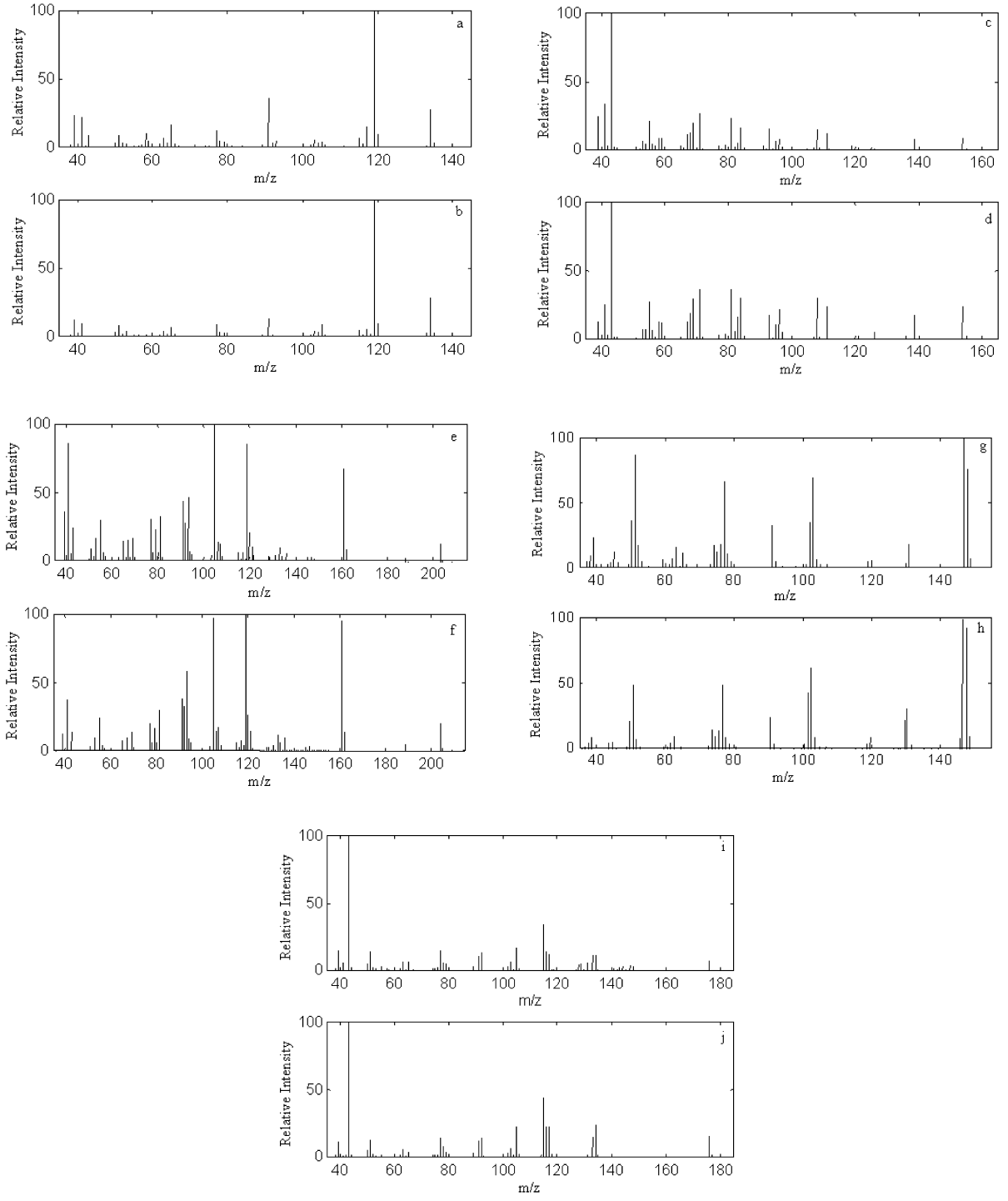

Fig. 9. Resolved mass spectra and their standard mass spectra of peak clusters I, II and III in Fig. 4a. Resolved (a) and standard (b) mass spectra of 4-ethyl-*o*-xylene; resolved (c) and standard (d) mass spectra of *p*-cineole; resolved (e) and standard (f) mass spectra of capaene, resolved (g) and standard (h) mass spectra of cinnamic acid; resolved (i) and standard (j) mass spectra of cinnamyl acetate.

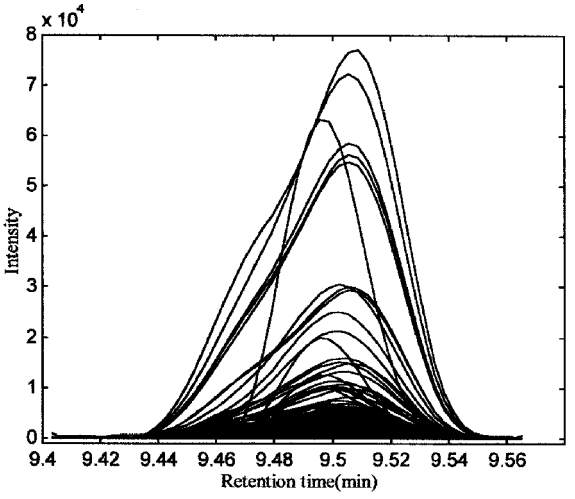

Fig. 10. All the ionic chromatograms of peak cluster III in Fig. 4a.

16]. This quantitative method is certainly superior to the common quantitative approach with which the quantitative analysis is usually conducted with the peak area integration based on the TIC. The determination of the content of all components is conducted by use of the overall volume technique in this paper. The results show that in the essential oils of *Cortex Cinnamomi* from Zhaoqing, Yulin, Yunan and Vietnam, there are 88, 94, 93 and 89 components separated, 61, 64, 61 and 58 of them determined representing about 93.62%, 93.39%, 92.03% and 92.59% of the total relative content, respectively.

Something should be pointed out here: the content of cinnamaldehyde is very high for every sample studied and it is possibly over the linear detection region of GC–MS (See Fig. 4). Therefore, the detection with diluted samples and low injection volume is carried out. On the other hand, the content of cinnamaldehyde and cinnamic acid is determined with their corresponding standard components (see Table 1).

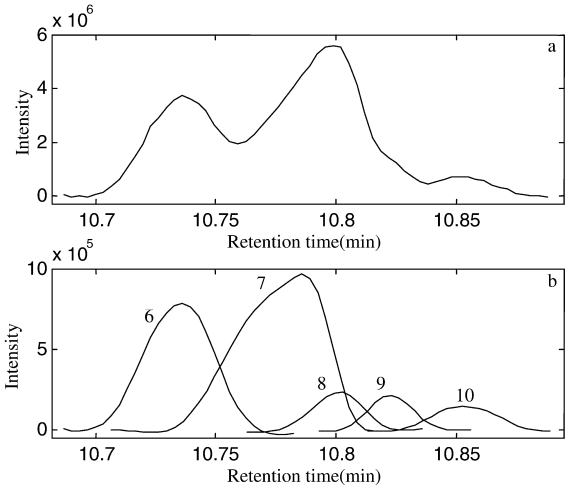

Fig. 11. Total (a) and resolved Chromatograms (b) of peak cluster IV in Fig. 4a.

5. Conclusion

From the above resolving process and results obtained by the use of chemometric resolution methods upon two-dimensional data together with the big mass spectral database, one can see that the real complex sample, like traditional Chinese medicine *Cortex Cinnamomi*, can be elegantly qualitatively and quantitatively analyzed. This shows that the hyphenated instruments combined with chemometric resolution methods will provide a completely new way for the quick and accurate analyses of real unknown complex systems. These methods, especially the IOP developed in this paper can not only greatly enhance the separation ability but also the qualitative identifying ability of the hyphenated chromatography, which shows the prosperous prospect for analysts to directly address very difficult problems in analytical chemistry.

Acknowledgements

This work is financially supported by the National Natural Science Foundation of China (grant No. 29875007) and the Union Laboratory of Asymmetric Synthesis of The Hong Kong Polytechnic University.

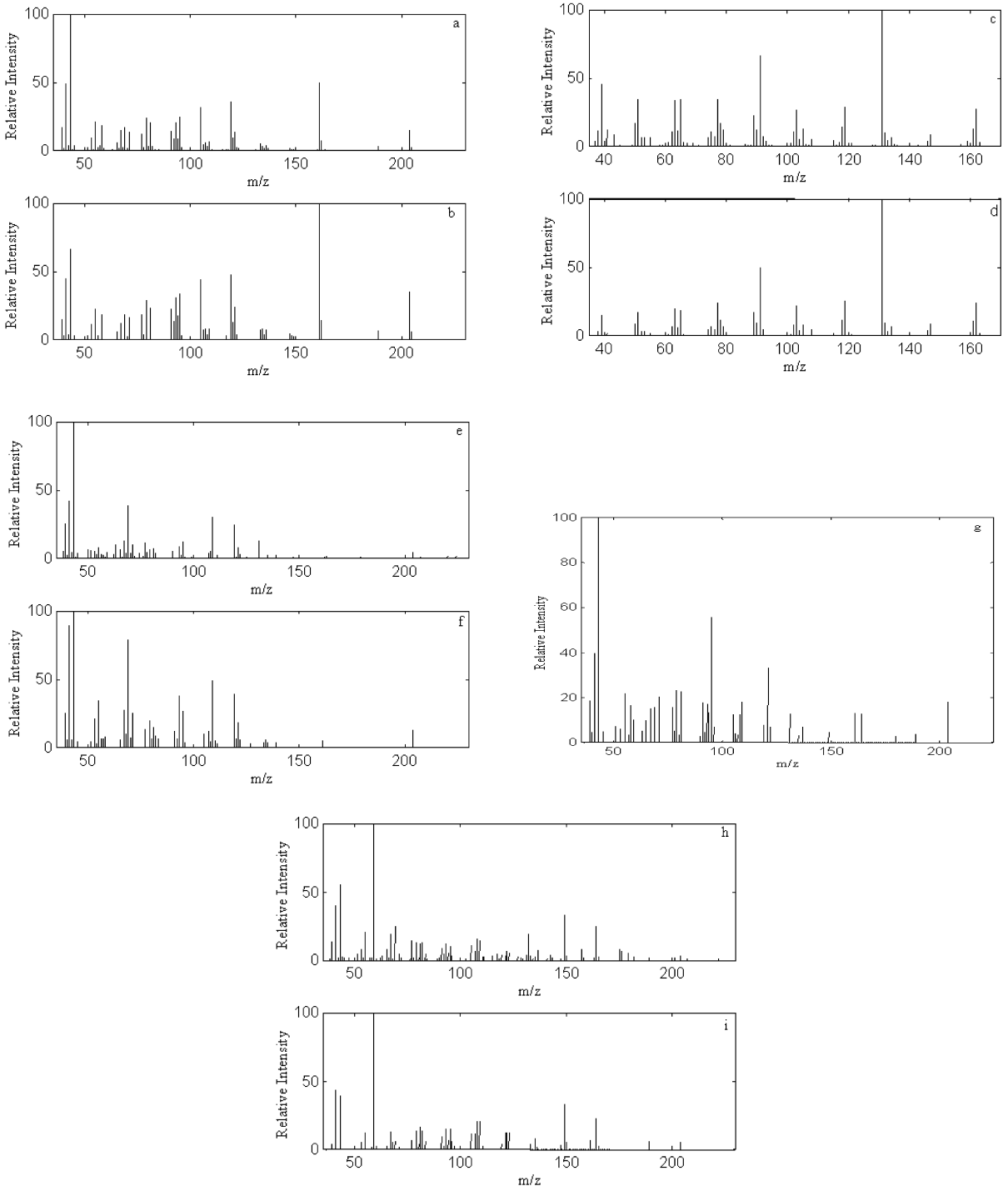

Fig. 12. Resolved mass spectra and their standard mass spectra of peak clusters IV in Fig. 4a. Resolved (a) and standard (b) mass spectra of δ-cadinol; resolved (c) and standard (d) mass spectra of *o*-methoxycinnamaldehyde; resolved (e) and standard (f) of bisabolol; resolved (g) mass spectrum of the fourth component (There is no corresponding standard mass spectrum); resolved (i) and standard (j) mass spectra of eudesmol; The fourth component cannot be determined.

References

[1] K.C. Wen, C.Y. Huang, F.L. Lu, J. Chromatogr. 631 (1993) 241.
[2] O.M. Kvalheim, Y.Z. Liang, Anal. Chem. 64 (1992) 936.
[3] Y.Z. Liang, O.M. Kvalheim, H.R. Keller, D.L. Massart, P. Kiechle, F. Erni, Anal. Chem. 64 (1992) 946.
[4] Y.Z. Liang, O.M. Kvalheim, A. Rahmani, R.G. Brereton, J. Chemom. 7 (1993) 15.
[5] M. Maeder, A.D. Zuberbuhler, Anal. Chim. Acta 181 (1986) 287.
[6] M. Maeder, Anal. Chem. 59 (1987) 527.
[7] M. Maeder, A. Zilian, Chemom. Intell. Lab. Syst. 3 (1988) 205.
[8] E.R. Malinowski, J. Chemom. 6 (1992) 29.
[9] W. Den, E.R. Malinowski, J. Chemom. 7 (1993) 89.
[10] R. Manne, H.L. Shen, Y.Z. Liang, Chemom. Intell. Lab. Syst. 45 (1999) 171.
[11] H.L. Shen, R. Manne, Q.S. Xu, D.Z. Chen, Y.Z. Liang, Chemom. Intell. Lab. Syst. 45 (1999) 323.
[12] Y.Z. Liang, O.M. Kvalheim, Anal. Chim. Acta 292 (1994) 5.
[13] Y.Z. Liang, in: White, Grey and Black Multicomponent Systems and Their Chemometric Algorithms, Hunan Publishing House of Science and Technology, Changsha, Hunan, 1996, pp. 177–206.
[14] F. Gong, Y.G. Peng, H. Cui, Y.Z. Liang, A.K.M. Leung, F.T. Chau, Acta Pharm. Sin. 34 (1999) 214.
[15] F. Gong, Y.G. Peng, H. Cui, Y.Z. Liang, A.K.M. Leung, F.T. Chau, Chem. J. Chin. Univ. 20 (1999) 199.
[16] H.L. Shen, Y.Z. Laing, R.Q. Yu, X.C. Li, X.X. Sun, Sci. China B 41 (1998) 21.
[17] R. Manne, Chemom. Intell. Lab. Syst. 27 (1995) 89.
[18] Y.Z. Liang, O.M. Kvalheim, Anal. Chim. Acta 276 (1993) 425.
[19] C.H. Xiao, in: Chemistry on Traditional Chinese Medicine, Shanghai Publishing House of Science and Technology, Shanghai, 1997, pp. 494–508.
[20] Z.Q. Lei, in: Science on Traditional Chinese Medicine, Shanghai Publishing House of Science and Technology, Shanghai, 1997, pp. 151–152.
[21] F.K. Chen, in: Determination of Efficacious Components in Commonly Used Traditional Chinese Medicine, Publishing House of People's Health, Beijing, 1997, pp. 227–236.
[22] W.G. Zhang, X.L. Li, Z. N Cai, Bull. Chin. Materia Medica 139 (1988) 38.
[23] W.N. Chen, Chin. Pharm. Bull. 22 (1987) 612.
[24] M.S. Zhu, S. Liu, R.J. Luo, Y.L. Bu, Acta Pharm. Sinica 31 (1996) 461.
[25] L.H. Zhao, in: Handbook of HPLC Applied to Determination of Chemical Components in Traditional Chinese Medicine, Chinese Publishing House of Medicinal Science and Technology, Beijing, 1994, pp. 318–322.
[26] Y.Z. Liang, R. Manne, O.M. Kvalheim, Chemom. Intell. Lab. Syst. 22 (1994) 229.
[27] Chinese Pharmacopoeia Committee, Chinese Pharmacopoeia - Appendix 62, Publishing House of People's Health, 1995.
[28] H.R. Keller, D.L. Massart, Anal. Chim. Acta 246 (1991) 379.
[29] O.M. Kvalheim, Chemom. Intell. Lab. Syst. 2 (1987) 283.