Journal of Chromatography, 506 (1990) 223–243
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 1741

Resolution of unresolved peaks containing unknown components by high-performance liquid chromatography with multi-wavelength detection

ICHIRO SAKUMA*
Faculty of Science and Engineering, Tokyo Denki University, Hatoyama, Hiki-gun, Saitama 350-03 (Japan)
NOBUHARU TAKAI
Institute of Industrial Science, University of Tokyo, 22-1 Roppongi, 7-chome, Minato-ku, Tokyo 106 (Japan)
TAKEYOSHI DOHI
Faculty of Engineering, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113 (Japan)
YASUHIRO FUKUI
Faculty of Science and Engineering, Tokyo Denki University, Hatoyama, Hiki-gun, Saitama 350-03 (Japan)
and
AKIYUKI OHKUBO
Faculty of Medicine, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113 (Japan)

SUMMARY

A method for the resolution of unresolved peaks obtained by high-performance liquid chromatography with multi-wavelength detection was developed. The method estimates the elution profiles and absorption spectrum of a component eluting at the rising edge or trailing edge of the unresolved peak and estimates the relative intensity of the derived three-dimensional chromatogram of one component by rank annihilation.

Artifical unresolved peaks and actual unresolved three-component peaks were resolved by the developed method. The results showed that the method can estimate peak area with errors of less than about 10% when the resolution R_s of the components is greater than about 0.4. The accuracy of estimation is considered to be superior to that of the method based on principal component analysis followed by multiple regression analysis, especially if the elution profiles of components are distorted from a Gaussian shape such as with tailing, where the estimation of elution profiles by principal components analysis seems erroneous.

INTRODUCTION

Several methods that resolve unresolved peaks in chromatograms only by data analysis of three-dimensional chromatograms obtained by high performance liquid chromatography (HPLC) with multi-wavelength detection such as with diode-array

0021-9673/90/$03.50

detectors have been reported. In these methods, the multivariate analysis technique is applied to the analysis of an unresolved peak in a three-dimensional chromatogram[1–7]. In one of these methods, the spectrum of a component that elutes at the rising edge of an unresolved peak and that of a component at the trailing edge are estimated by extrapolation of observed spectra, and the elution profiles of two components are estimated by means of multiple regression analysis (curve fitting)[4]. However, the application of this method is limited to the resolution of two-component unresolved peaks.

For the resolution of unresolved peaks containing more than three components, principal component analysis (factor analysis) is used, where the data matrices obtained are decomposed into the orthonormal vectors calculated by principal component analysis, and the elution profile or spectrum of each component present is estimated as a linear combination of these orthonormal vectors (the first step). Generally, only the elution profile or absorption spectrum is estimated. Thereafter, the remaining spectra or elution profiles are estimated by curve fitting of the estimated value to the observed data matrix (the second step)[1–3,6,7]. However, in these methods, errors in the first estimation step have direct influences on the estimation results of the second step. Consequently, the precision of the qualitative and quantitative estimation of each component's spectrum and elution profile may deteriorate.

In this study, a new method was developed that estimates both the elution profile and the absorption spectrum of each component in an unresolved peak and estimates the shapes and intensities of the three-dimensional chromatograms of components in the unresolved peak. The performance of the developed method in unresolved peak resolution was evaluated and compared with that of the conventional peak resolution method.

EXPERIMENTAL

Estimation of elution profile

Details of the theoretical aspects of the estimation method using principal component analysis are available in several references[8–10]. A brief description of algorithm is presented here. A multi-wavelength absorption detector can simultaneously monitor several chromatograms at N different wavelengths λ_i ($i = 1, 2, \ldots, N$). Let d_{ij} be the measured absorbance of an unresolved peak at wavelength λ_i and time t_j ($j = 1, 2, \ldots, M$); three-dimensional chromatographic data can be expressed as an $N \times M$ matrix as follows:

$$\mathrm{D} = \begin{pmatrix} d_{11} & d_{12} & \ldots & d_{1M} \\ d_{21} & d_{22} & \ldots & d_{2M} \\ \vdots & \vdots & & \vdots \\ d_{N1} & d_{N2} & \ldots & d_{NM} \end{pmatrix} \tag{1}$$

The ith-row vector is an observed chromatogram at wavelength λ_i and the jth-column vector is an observed spectrum at time t_j.

Here we assume that an unresolved peak consists of L components, and the

spectrum S_k and elution profile (chromatogram) C_k are expressed as the following vectors:

$$S_1 = \begin{pmatrix} s_{11} \\ s_{12} \\ \vdots \\ s_{1N} \end{pmatrix}, S_2 = \begin{pmatrix} s_{21} \\ s_{22} \\ \vdots \\ s_{2N} \end{pmatrix}, \ldots, S_L = \begin{pmatrix} s_{L1} \\ s_{L2} \\ \vdots \\ s_{LN} \end{pmatrix}$$

$$C_1 = \begin{pmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1M} \end{pmatrix}, C_2 = \begin{pmatrix} c_{21} \\ c_{22} \\ \vdots \\ c_{2M} \end{pmatrix}, \ldots, C_L = \begin{pmatrix} c_{L1} \\ c_{L2} \\ \vdots \\ c_{LM} \end{pmatrix}$$

where s_{ki} is the relative intensity of the spectrum of the kth component at wavelength λ_i and c_{kj} is the relative intensity of the elution profiles of the kth component at time t_j.

If the detector output is proportional to the concentration of a sample, and the principle of superposition is valid for both the absorption spectra and elution profiles of sample mixtures, the observed data matrix D can be expressed as

$$D = \sum_{k=1}^{L} e_k S_k C_k^{\mathrm{T}} + R \tag{4}$$

where the superscript $^{\mathrm{T}}$ denotes the transposed matrix (vector), e_k is a value proportional to the concentration of the kth component and R is the noise matrix. If we neglect R, the rank of the matrix D is equal to the number of components L, as the spectra and elution profiles of different components are linearly independent. Hence the data matrix D can be decomposed by L sets of orthonormal vectors u_k and $v_k (k = 1, 2, \ldots, L)$ as follows:

$$D = \sum_{k=1}^{L} \xi_k u_k v_k^{\mathrm{T}} \tag{5}$$

The ξ_k are coefficients in the linear combination. As u_k and v_k are orthonormal,

$$Dv_k = \xi_k u_k \tag{6}$$

$$u_k^{\mathrm{T}} \mathrm{D} = \xi_k v_k^{\mathrm{T}}$$

It follows that:

$$D^{\mathrm{T}} D v_k = \xi_k^2 v_k \tag{7}$$

$$D D^{\mathrm{T}} u_k = \xi_k^2 u_k$$

These relationships mean that u_k and v_k are the eigenvectors of second moment matrix D^TD or DD^T *and* ξ_k^2 is its kth eigenvalue. In other words, u_k and v_k ar principal components of multivariate data D. In a real data matrix, because of the existence of noise the rank of the data matrix D is greater than the number of components L. However, eigenvalues of the matrix D^TD contain the information about the number of component in the unresolved peak. By analysing the eigenvalues, we can estimate the number of components, as described by Malinowski[11,12].

On the other hand, comparing eqns. 4 and 5, we can see that the elution profile and spectrum of the kth component can be expressed as a linear combination of eigenvectors u_k and v_k:

$$S_k = \sum_{l=1}^{L} x_{kl}\, u_l \tag{8}$$

$$C_k = \sum_{l=1}^{L} y_{kl}\, v_l \tag{9}$$

where the x_{kl} and y_{kl} ($k = 1, 2, \ldots, L$; $l = 1, 2, \ldots, L$) are the coefficients in the linear combination that must be determined for each component k.

As v_k can be calculated from a given data matrix D using eqn. 7, and the elution profile of each component is expressed as eqn. 9, we can estimate the elution profile of each component by estimating the coefficients y_{kl} in eqn. 9.

In the estimation of y_{kl}, two natural constraints and one evaluation function are assumed:

(1) the elution profile of each components is not negative, which can be expressed as

$$c_{kl} \geqslant 0 \tag{10}$$

(2) the spectrum of each component is not negative; after estimating the elution profiles of all the components in the unresolved peak, we can estimate the spectrum (spectra multiplied by the relative concentration) of each component S_k, which can be expressed as

$$s_{ki} \geqslant 0 \tag{11}$$

(3) under the above two constraints, the elution profiles have an ideal shape as a chromatographic peak. To express unimodality of a chromatographic peak, the area to norm ratio is generally used[2,6]. However, in this study, to express other properties of chromatographic shape (smoothness and stability of the baseline in addition to unimodality), the following function expressing the entropy of time derivatives of elution profiles are adopted[13–15]:

$$H = -\sum_{l=1}^{L}\sum_{j=1}^{M} P_{lj} \log (P_{lj}) \tag{12}$$

$$P_{lj} = | c'_{lj} | / (\sum_{j=1}^{M} | c'_{lj} |) \quad (13)$$

where c'_{lj} is the first or second derivative of the lth component's elution profiles at time t_j. By minimizing eqn. 12, ideal elution profiles of components are estimated under the constraints of eqns. 10 and 11. In order to avoid ambiguity in the relative intensities of the elution profiles, the area of each elution profile is set to unity in the actual estimation process.

Estimation of spectrum

By the same discussion as for the spectra, the observed spectrum at time t_j can be expressed as a linear combination of eigenvectors $\boldsymbol{u}_k$ in eqn. 7. Let $\boldsymbol{d}_j$ be the jth column vector of the data matrix expressing the spectrum at time t_j; $\boldsymbol{d}_j$ can be also expressed as a linear combination of $\boldsymbol{u}_k$ as follows:

$$\boldsymbol{d}_j = \sum_{l=1}^{L} w_{jl} \boldsymbol{u}_l \quad (14)$$

To discuss changes in the coefficients w_{jl} as time advances, let us consider the two-component case. Fig 1a shows a schematic diagram of three-dimensional chromatogram of a two-component unresolved peak. Using eqn. 14, each observed spectrum must be located on the plane determined by two eigenvectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. Considering eqn. 14, the coefficients w_{jl} are considered to be coordinates in the plane determined by $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. They also express the direction of the observed spectrum in this plane as Fig. 1b shows. We can plot the trajectory of the observed spectra (Fig. 1b). Consider that the spectra at the rising edge of unresolved peak resemble the spectrum of the component which elutes at the rising edge; the direction determined by w_{1j} converges to the direction of the pure spectrum of the first-eluted compounds in the

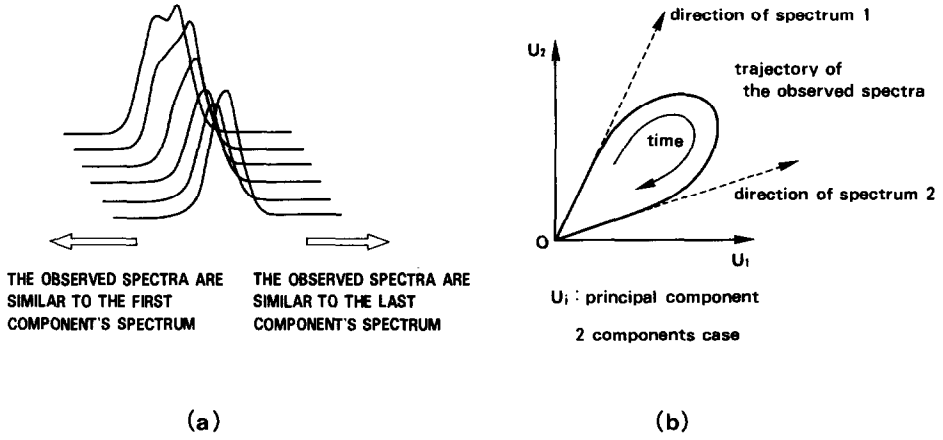

Fig. 1. Estimation of a spectrum by extrapolation. (a) Three-dimensional chromatogram of two-component unresolved peak. (b) Trajectory of observed vectors in the space determined by the two principal components.

initial part of the trajectory. Thus, by extrapolating the w_{1j} in the reverse direction of time, the spectrum of the first-eluting compounds can be estimated as a convergence direction of the observed spectra in the plane determined by $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. The convergence direction can be calculated numerically by approximating the trajectory of the tip of the observed spectrum vectors by a polynomial of parameters such as the length of the trajectory from the origin. Similar argument can be used for more than three-component cases. Hence one can at least estimate the spectrum of the component which elutes first at the rising edge of an unresolved peak and that of the component which elutes at its trailing edge.

Estimation of relative intensity of an estimated three-dimensional chromatogram

In the above methods, elution profiles $\boldsymbol{C}_k$ of all the components in an unresolved peak and spectrum $\boldsymbol{S}_1$ of the component that elutes at the rising edge or trailing edge of the unresolved peak can be estimated. Hence we can determine the shape of the three-dimensional chromatogram P_1 of the components eluting at the rising (or trailing) edge of the unresolved peak:

$$P_1 = \boldsymbol{S}_1 \boldsymbol{C}_1^{\mathrm{T}} \tag{15}$$

To resolve an unresolve peak into each component, the relative intensity of this three-dimensional chromatogram P_1 must be estimated. In this estimation, rank annihilation is adopted[16–18].

Let α be the best estimate of relative intensity of P_1. As the data matrix after subtracting the three-dimensional chromatogram of the first-eluting component contains only $L-1$ components, the degrees of freedom of the matrix $(D-\alpha P_1)$ decreases from the original number of components L to $L-1$. This condition can be expressed as

$$\text{rank}\,\{(D - \alpha P_1)(D - \alpha P_1)^{\mathrm{T}}\} = L-1 \tag{16}$$

The analytical solution of eqn. 16 was given by Lorber[19]. Let U and V be matrices constructed by $\boldsymbol{u}_k$ and $\boldsymbol{v}_k$ in eqn. 7 as follows:

$$\begin{aligned} U &= [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_L] \\ V &= [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_L] \end{aligned} \tag{17}$$

Define $\boldsymbol{a}$ and $\boldsymbol{b}$ as

$$\begin{aligned} \boldsymbol{a} &= U^{\mathrm{T}} \boldsymbol{S}_1 \\ \boldsymbol{b} &= V^{\mathrm{T}} \boldsymbol{C}_1 \end{aligned} \tag{18}$$

Note that $\boldsymbol{a}$ and $\boldsymbol{b}$ are L-dimensional vectors. Let a_i and b_i be the ith element of $\boldsymbol{a}$ and $\boldsymbol{b}$, then the estimated α is calculated as follows:

$$1/\alpha = \sum_{i=1}^{L} a_i b_i / \xi_i \tag{19}$$

Note that ξ_i is the square root of the kth eigenvalue expressed as enq. 7.

$D - \alpha P_1$ contains only $L-1$ components. By iterating these procedures, we can resolve an unresolved peak into three-dimensional chromatograms of existing components.

Computer program

Programs performing the above-described algorithm were developed on an NEC PC-9801 personal computer. The programs were written in C language.

A brief description of the flow of operation is as follows:

(1) Determine the area of an unresolved peak using the cursor displayed on the CRT of the computer together with the contour plot of the three-dimensional chromatogram.

(2) Calculate eigenvectors and eigenvalues of the second moment matrix D^TD.

(3) Determine the number of components contained in the unresolved peak by analysing the eigenvalues calculated in step 2.

(4) Estimate the elution profiles of each component by minimizing eqn. 12 under the constraints.

(5) Determine the region of the rising edge using the displayed cursors together with a three-dimensional chromatogram or contour plot of an unresolved peak displayed on the CRT of the computer. After setting the region, the computer automatically extrapolates the observed spectra and calculates the estimated spectrum of the first- (or last-) eluting component.

(6) Calculate the relative intensity α of the three-dimensional chromatogram P_1 in the given unresolved peak data using eqns. 17–19.

(7) Subtract the estimated three-dimensional chromatogram of the component eluting at the rising edge or trailing edge of the unresolved peak.

(8) If the remaining unresolved peak contains more than two components repeat steps 2–7.

In the estimation of elution profiles, non-linear programming with constraints must be performed. An augumented Lagrangian algorithm was adopted for that purpose[20].

Procedures

Unresolved peak data were prepared both by numerical calculation of artificial chromatograms and by the actual measurement of unresolved peaks by HPLC with multi-wavelength detection. The developed algorithm was tested in the resolution of artificial and actual unresolved peak data. In order to investigate the performance of the developed method compared with the method based on principal component analysis followed by multiple regression analysis, the same data were analysed using multiple regression analysis where the spectrum of each component was estimated by curve fitting of the estimated elution profiles by multiple regression analysis.

The artificial three-dimensional chromatograms were generated assuming that it contained two or three components, that the three components had the spectra shown in Fig. 2 and that each chromatogram was Gaussian with the same variance (peak width). Several three-dimensional chromatogram data were calculated at different resolutions, R_s, and peak-height ratios. Also, random noise of the order of 1% of the standard deviation of the chromatographic data were added to the data. The size of the

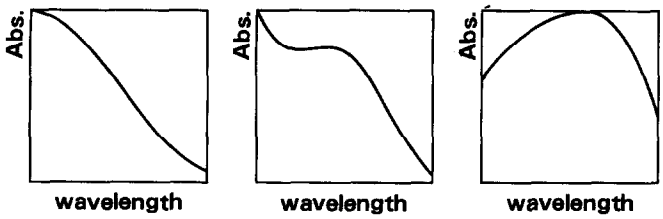

Fig. 2. Spectra of three components used in artificial unresolved peak calculation.

data matrix was 30 × 30, which meant that 30 chromatograms were measured at 30 different wavelengths and the intensities (absorbances) of the chromatograms were measured at 30 different points of time.

In the experiment on unresolved peaks of actual samples, the three-components sample mixtures shown in Table I were separated by reversed-phase HPLC under the conditions shown in Tabe II. The areas to be analysed by the algorithms were determined by manual operation using cursors displayed on the CRT display of the computer system together with the contour plot of the three-dimensional chromatogram obtained. The data matrix size was 30 × 40, which meant that the chromatograms were measured at 30 different wavelengths and at 40 different points in time. Caffeine eluted first, N-methylaniline second and *o-tert.*-butylphenol third in this separation system. The measured resolutions were as follows: (i) with 84.0%

TABLE I

SAMPLES USED IN THE EXPERIMENTS

Samples were dissolved in acetonitrile.

Sample	*Caffeine concentration (wt.-%)*	*N-Methylaniline concentration (vol.-%)*	*o-tert.-Butylphenol concentration (vol.-%)*
A	0.043	0.0033	0.033
B	0.051	0.0040	0.020
C	0.064	0.0025	0.025
D	0.073	0.0029	0.014
E	0.128	0.0000	0.000
F	0.000	0.0100	0.000
G	0.000	0.0000	0.100

TABLE II

HPLC SEPARATION CONDITIONS

Packing	Hitachi 3056 (ODS)
Column	50 mm × 4 mm I.D.
Eluent	acetonitrile–water: (1) 84.0:16.0 (v/v); (2) 80.6:19.4 (v/v)
Flow-rate	1.0 ml/min
Injection volume	10 μl
Detector	MCPD-350 diode-array detector (Otsuka Electronics, Osaka, Japan) with laboratory-made data acquisition system[21]

acetonitrile as the eluent, R_s (caffeine/N-methylaniline) = 0.33 and R_s (N-methylaniline/*o-tert.*-butylphenol) = 0.22; and (ii) with 80.6% acetonitrile as the eluent, R_s (caffeine/N-methylaniline) = 0.46 and R_s (N-methylaniline/*o-tert.*-butylphenol) = 0.33.

The results of the unresolved peak separation were evaluated using the spectrum shape similarity, defined as the correlation coefficients between the actual spectrum and the estimated spectrum for qualitative aspects and the peak area of the estimated elution profile compared with the actual peak area for quantitative aspects.

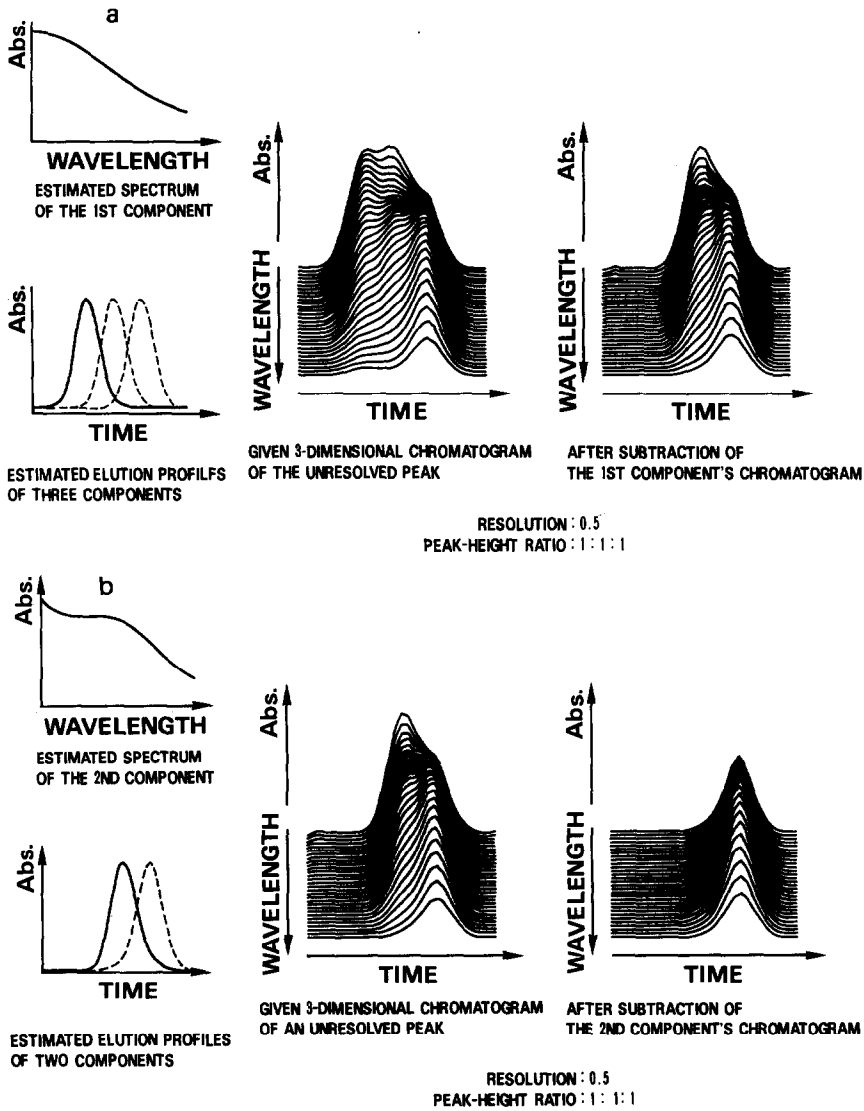

Fig. 3. Example of peak resolution of artificial three-component unresolved peak. (a) Subtraction of the first component; (b) subtraction of the second component.

RESULTS

Fig. 3 shows an example of peak resolution by the proposed method. Fig. 3a shows subtraction of the first components from the original three-dimensional chromatogram and Fig. 3b subtraction of the second components from the residual two-component three-dimensional chromatogram.

Table III shows the estimated peak area (peak volume summation of the peak area at every wavelength) of each component of a two-component unresolved peak with various resolutions. For an unresolved peak with resolution $R_s < 0.4$, the estimation errors were larger than 10% and for $R_s > 0.5$ they were less than 10%. Table IV shows the results obtained by multiple regression analysis. In the resolution of those artificial unresolved peaks with $R_s < 0.4$, the developed method was not as efficient as the method using multiple regression analysis. However, in the analysis of those artificial unresolved peaks with $R_s > 0.5$, the results obtained by the two methods did not show clear differences.

TABLE III

RESULTS OF PEAK RESOLUTION OF ARTIFICIAL TWO-COMPONENT UNRESOLVED PEAK BY THE DEVELOPED METHOD

The values in parentheses represent the actual peak volumes.

Peak-height ratio	*Resolution, R_s*	*Component 1*		*Component 2*	
		Total peak area	*Error (%)*	*Total peak area*	*Error (%)*
		(4701)		(5180)	
1:1	0.1	6889	46.5	2990	−42.3
	0.2	5340	13.6	4539	−12.4
	0.3	4014	−14.6	5866	13.2
	0.4	5311	13.0	4570	−11.8
	0.5	5059	7.6	4823	−6.9
	0.6	4935	5.4	4948	−4.5
	0.7	4991	6.2	4892	−5.6
		(4701)		(2573)	
2:1	0.1	5738	22.1	2694	−39.7
	0.2	5020	6.8	2697	−11.8
	0.3	3987	−15.2	3155	+28.4
	0.4	5171	10.0	2395	−17.6
	0.5	4962	5.6	2664	−9.4
	0.6	4892	4.1	2757	−6.7
	0.7	4842	3.0	2788	−4.9
		(7522)		(2058)	
4:1	0.1	8381	11.4	2213	−41.2
	0.2	5769	−23.3	3345	−85.8
	0.3	6647	−11.6	2363	+43.1
	0.4	7083	−5.8	2313	+22.0
	0.5	7434	−1.2	1734	5.0
	0.6	7623	1.3	2000	−4.2
	0.7	7668	1.9	1717	−6.4

TABLE IV

RESULTS OF PEAK RESOLUTION OF ARTIFICIAL TWO-COMPONENT UNRESOLVED PEAK BY THE METHOD OF PRINCIPAL COMPONENT ANALYSIS FOLLOWED BY MULTIPLE REGRESSION ANALYSIS

The values in parentheses represent the actual peak volumes.

Peak-height ratio	*Resolution, R_s*	*Component 1*		*Component 2*	
		Total peak area	*Error (%)*	*Total peak area*	*Error (%)*
		(4701)		(5180)	
1:1	0.1	4560	−3.0	5319	+2.7
	0.2	4604	−2.1	5276	+1.9
	0.3	4320	−8.1	5560	+7.3
	0.4	4980	5.7	4902	−5.4
	0.5	4608	−1.9	5274	+1.7
	0.6	4492	−4.5	5391	+4.1
	0.7	4383	−6.8	5500	+6.2
		(4701)		(2573)	
2:1	0.1	4596	−2.2	2694	+4.7
	0.2	4593	−2.3	2697	+4.8
	0.3	4135	−12.0	3155	+22.6
	0.4	4896	4.1	2395	−6.9
	0.5	4628	−1.6	2664	+3.5
	0.6	4535	−3.5	2757	+7.2
	0.7	4504	−4.2	2788	+8.4
		(7522)		(2058)	
4:1	0.1	7379	−2.0	2213	+7.5
	0.2	6247	−17.0	3345	+62.5
	0.3	7229	−3.9	2363	+14.8
	0.4	7280	−3.2	2313	+12.4
	0.5	7859	4.5	1734	−15.7
	0.6	7594	1.0	2000	−2.8
	0.7	7878	4.7	1717	−16.6

Table V shows the correlation coefficients between the estimated spectrum and actual spectrum of the first and the second components in artificial three-component unresolved chromatograms. Table VI shows the results of the estimated peak volume. As shown in the two-component case, the peak areas of the component having the highest intensity in those unresolved peaks with $R_s > 0.4$ were estimated with errors of about 10%. Table VII shows the results by multiple regression analysis on the same data and it is clear that the developed system did not show any superiority.

However, the results for the unresolved peaks of an actual three-component mixture showed differences between the proposed method and the method with multiple regression analysis. Fig. 4 shows the actual elution profiles of the three components using 84.0% and 80.6% acetonitrile as the eluent. The peaks of caffeine and *o-tert.*-butylphenol were relatively wide and all the peaks showed tailing. Fig. 5 shows an example of the resolution of the unresolved peak of an actual sample mixture. The three-dimensional chromatogram of caffeine was subtracted from the originally observed three-dimensional chromatogram.

TABLE V

CORRELATION COEFFICIENTS BETWEEN ACTUAL SPECTRUM AND ESTIMATED SPECTRUM OF THE FIRST AND SECOND COMPONENTS FOR PEAK RESOLUTION OF ARTIFICIAL TWO-COMPONENT UNRESOLVED PEAK BY THE DEVELOPED METHOD

Resolution	*Peak-height ratio*	*Component 1*	*Component 2*
0.4	1:1:1	0.9996	0.9995
0.5	1:1:1	1.0000	0.9989
0.6	1:1:1	1.0000	0.9996
0.7	1:1:1	1.0000	1.0000
0.4	2:2:1	0.9999	0.9893
0.4	4:4:1	0.9999	0.9853
0.4	2:1:1	1.0000	0.9696
0.4	4:2:1	1.0000	0.9706
0.4	4:1:1	1.0000	0.9518
0.5	2:2:1	1.0000	0.9990
0.5	4:4:1	1.0000	1.0000
0.5	2:1:1	1.0000	0.9999
0.5	4:2:1	1.0000	0.9997
0.5	4:1:1	1.0000	0.9971

Figs. 6 and 7 show the results of concentration estimation by the proposed methods and Figs. 8 and 9 those obtained by multiple regression analysis. The estimated concentration was calculated from the estimated peak area of each component and the peak area of a standard one-component sample with known concentrations. Figs. 6 and Fig. 8 show the results of peak resolution where R_s (caffeine/N-methylaniline) = 0.33 and R_s (N-methylaniline/*o-tert.*-butylphenol) = 0.22. Figs. 7 and Fig. 9 show the results of peak resolution R_s (caffeine/N-methylaniline) = 0.46 and R_s (N-methylaniline/*o-tert.*-butylphenol) = 0.33.

For the unresolved peak resolution of actual three-component mixtures, the correlation coefficients between the estimated and actual concentrations were 0.916 and 0.976 by the proposed method and 0.803 and 0.952 by the method of principal component analysis followed by multiple regression analysis. These results showed that the proposed method was superior for the estimation of components in actual unresolved peaks.

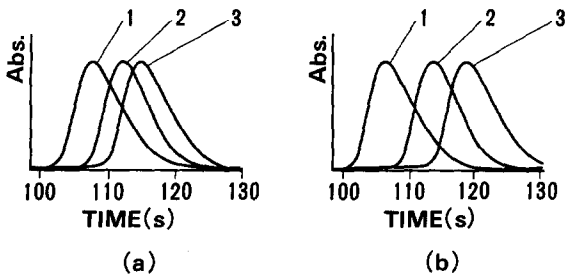

Fig. 4. Actual elution profiles of (1) caffeine, (2) N-methylaniline and (3) *o-tert.*-butylphenol. Eluent: acetonitrile–water, (a) 84.0:16.0 (v/v) and (b) 80.6:19.4 (v/v).

TABLE VI

RESULTS OF PEAK RESOLUTION OF ARTIFICIAL THREE-COMPONENT UNRESOLVED PEAK BY THE DEVELOPED METHOD

Resolution, R_s	*Peak-height ratio*	*Component 1*			*Component 2*			*Component 3*		
		Actual peak volume	*Estimated peak volume*	*Error (%)*	*Actual peak volume*	*Estimated peak volume*	*Error (%)*	*Actual peak volume*	*Estimated peak volume*	*Error (%)*
0.4	1:1:1	4701	4619	−1.7	5180	6354	22.7	6890	5878	−14.7
0.5	1:1:1	4701	4760	1.3	5180	4916	− 5.1	6890	7099	3.0
0.6	1:1:1	4701	4916	4.6	5180	5240	1.2	6890	6619	− 3.9
0.7	1:1:1	4701	4510	−4.1	5180	5536	6.9	6890	6728	− 2.3
0.4	2:2:1	5641	5729	1.6	6216	5279	−15.1	4134	4985	20.6
0.4	4:4:1	5641	5613	−0.5	6216	5638	− 9.3	2067	2674	29.4
0.4	2:1:1	5641	5666	0.4	3108	2145	−31.0	4134	5074	22.7
0.4	4:2:1	5641	5675	0.6	3108	2301	−26.0	2067	2842	37.5
0.4	4:1:1	5641	5694	0.9	1554	1167	−24.9	2067	2403	16.3
0.5	2:2:1	5641	5500	−2.5	6216	6003	− 3.4	4134	3795	− 8.2
0.5	4:4:1	5641	5675	0.6	6216	6470	4.1	2067	1775	−14.1
0.5	2:1:1	5641	5596	−0.8	3108	3384	8.9	4134	3904	− 5.6
0.5	4:2:1	5641	5539	−1.8	3108	3532	13.6	2067	1744	−15.6
0.5	4:1:1	5641	5094	−9.7	1554	2788	79.4	2067	1382	−33.1

TABLE VII

RESULTS OF PEAK RESOLUTION OF ARTIFICIAL THREE-COMPONENT UNRESOLVED PEAK BY THE METHOD OF PRINCIPAL COMPONENT ANALYSIS FOLLOWED BY MULTIPLE REGRESSION ANALYSIS

Resolution, R_s	*Peak-height ratio*	*Component 1*			*Component 2*			*Component 3*		
		Actual peak volume	*Estimated peak volume*	*Error (%)*	*Actual peak volume*	*Estimated peak volume*	*Error (%)*	*Actual peak volume*	*Estimated peak volume*	*Error (%)*
0.4	1:1:1	4701	4761	4.2	5180	4710	− 9.1	6890	7302	6.0
0.5	1:1:1	4701	4726	0.5	5180	5224	0.8	6890	6824	− 0.9
0.6	1:1:1	4701	4667	− 0.7	5180	4981	− 3.8	6890	7127	3.5
0.7	1:1:1	4701	4368	− 7.0	5180	4954	− 4.4	6890	7453	8.2
0.4	2:2:1	5641	5881	4.3	6216	6402	3.0	4134	3709	−10.3
0.4	4:4:1	5641	5608	− 0.6	6216	5427	−12.7	2067	2890	39.8
0.4	2:1:1	5641	5573	− 1.2	3108	3378	8.7	4134	3934	− 4.8
0.4	4:2:1	5641	5542	− 1.8	3108	3240	4.2	2067	2035	− 1.5
0.4	4:1:1	5641	5603	− 0.7	1554	1827	17.6	2067	1833	−11.3
0.5	2:2:1	5641	4853	−14.0	6216	7404	19.1	4134	3737	− 9.6
0.5	4:4:1	5641	5476	− 2.9	6216	6096	− 1.9	2067	2353	13.8
0.5	2:1:1	5641	5638	− 0.1	3108	3283	5.6	4134	3965	− 4.1
0.5	4:2:1	5641	5749	1.9	3108	3044	− 2.1	2067	2025	− 2.0
0.5	4:1:1	5641	5028	−10.9	1554	2100	35.1	2067	2135	3.3

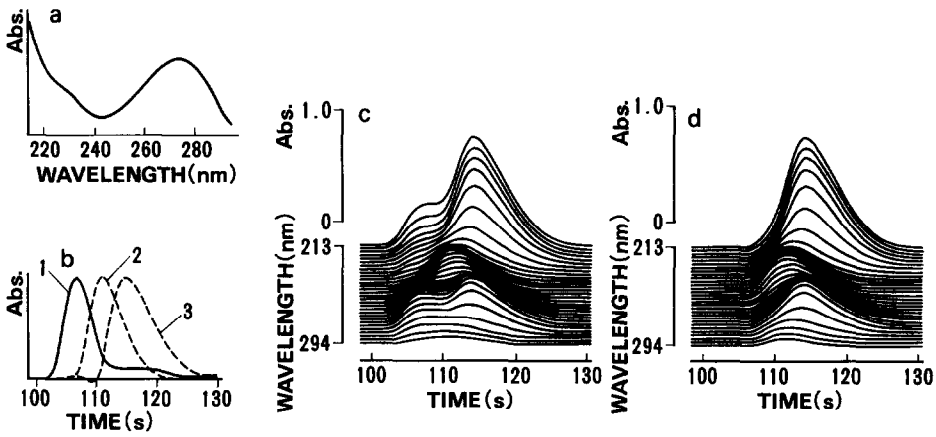

Fig. 5. Example of peak resolution of actual three-component unresolved peak to illustrate elution profile estimation and observed spectra extrapolation followed by rank annihilation. (a) Estimated spectrum of caffeine. (b) Estimated elution profiles of (1) caffeine, (2) N-methylaniline and (3) *o-tert.*-butylphenol. (c) Three-dimensional chromatogram of caffeine, N-methylaniline and *o-tert.*-butylphenol. (d) Three-dimensional chromatogram after subtraction of caffeine. Sample: 0.043% (w/w) caffeine, 0.033% (v/v) N-methylaniline and 0.033% (v/v) *o-tert.*-butylphenol (10 μl). Eluent: acetonitrile–water (84.0:16.0, v/v). Resolution: R_s (caffeine/N-methylaniline) = 0.33; R_s (N-methylaniline/*o-tert.*-butylphenol) = 0.22.

DISCUSSION

Effects of peak resolution on the estimated results

The results of the resolution of unresolved peaks on both artificial unresolved chromatograms and actual chromatograms showed that when $R_s >$ than 0.4, the dominant component in the unresolved peak can be determined with a maximum error of about 10% by the developed method. Computer simulations were performed for the two-component case where unresolved peaks with $R_s < 0.3$ were resolved. The estimation errors were larger for an unresolved peak with $R_s < 0.4$.

The results for the actual unresolved peaks also showed that the peak areas of three components could be estimated with a correlation coefficient of 0.976 when $R_s = 0.46$ and 0.33. On the other hand, when $R_s = 0.22$ and 0.33, the correlation coefficient between the estimated peak areas and the actual peak areas decreased to 0.916. These results also showed that the developed method could separate unresolved peaks with $R_s \geqslant 0.4$. These results coincided with those for the artificial unresolved peaks. Hence it is considered that $R_s = 0.4$ is a certain criterion that determines the limits of the ability of the developed peak resolution method. The main factor determining this limitation is considered to be errors in spectrum estimation. The results of the experiments on peak resolution by multiple regression analysis showed that the elution profiles were well estimated, so that the estimation errors under the condition of small R_s were relatively small, as shown in Table IV. It is considered that the errors in estimation using the developed method mainly came from errors in spectrum estimation in the experiments on the artificial unresolved peaks. It is difficult to extrapolate the observed spectra only by curve fitting to a polynomial in the case of small resolutions because the observed spectra suffer from distortion by several

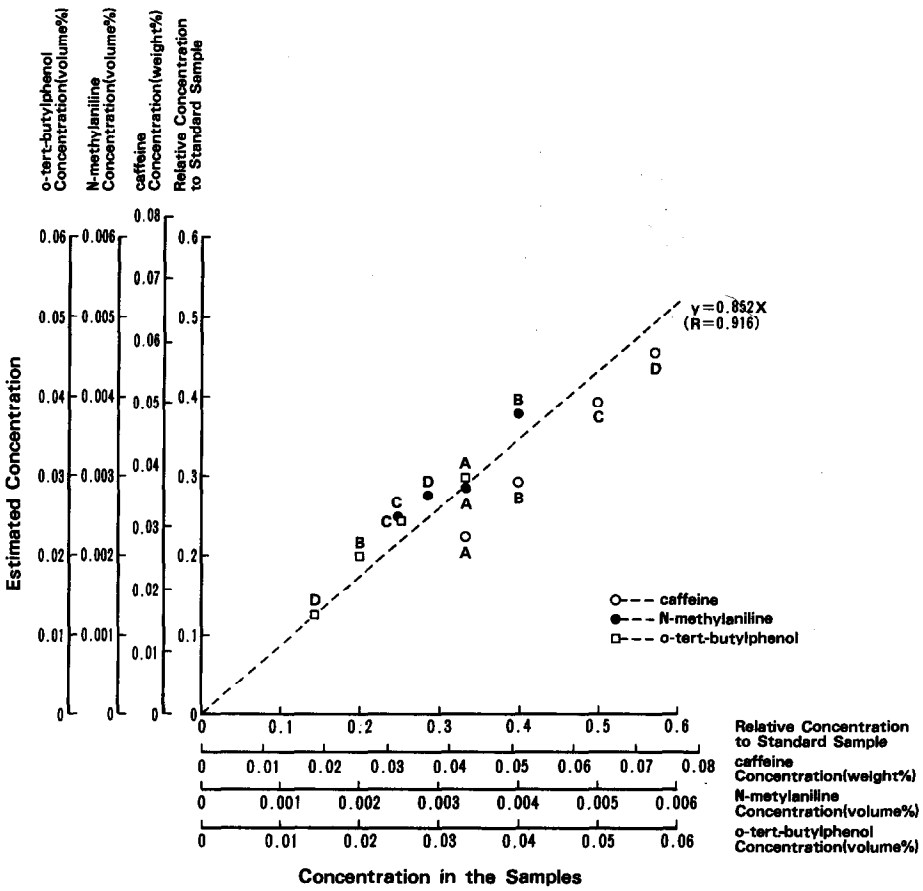

Fig. 6. Results of quantative analysis by the developed method. Eluent: acetonitrile–water (84.0:16.0, v/v). Resolution: R_s (caffeine/N-methylaniline) = 0.33; R_s (N-methylaniline/*o-tert.*-butylphenol) = 0.22. Quantitation: (○) caffeine, peak area at 262 nm; (●) N-methylaniline, peak area at 246 nm; (□) *o-tert.*-butylphenol, peak area at 227 nm. A, B, C and D correspond to the sample shown in Table I.

co-eluting components. Hence it is considered that the present method was not suitable for the resolution of strongly overlapped peaks with small resolutions.

Effects of peak-height ratio on the estimated results

The results of computer simulation on the artificial unresolved peaks showed that if the peak height of one component is much smaller than those of co-eluting components, the errors in its area estimation were relatively large in comparison with those for the dominant component in the unresolved peak. However, those components which had the largest intensities in a unresolved peak could be well determined with small errors when there were large differences in peak height between the co-eluting compounds. Hence it is considered that the developed method can be applied to the quantitative analysis of the main component's peak area in an unresolved peak with estimation errors $< 10\%$ provided that $R_s \geqslant 0.4$.

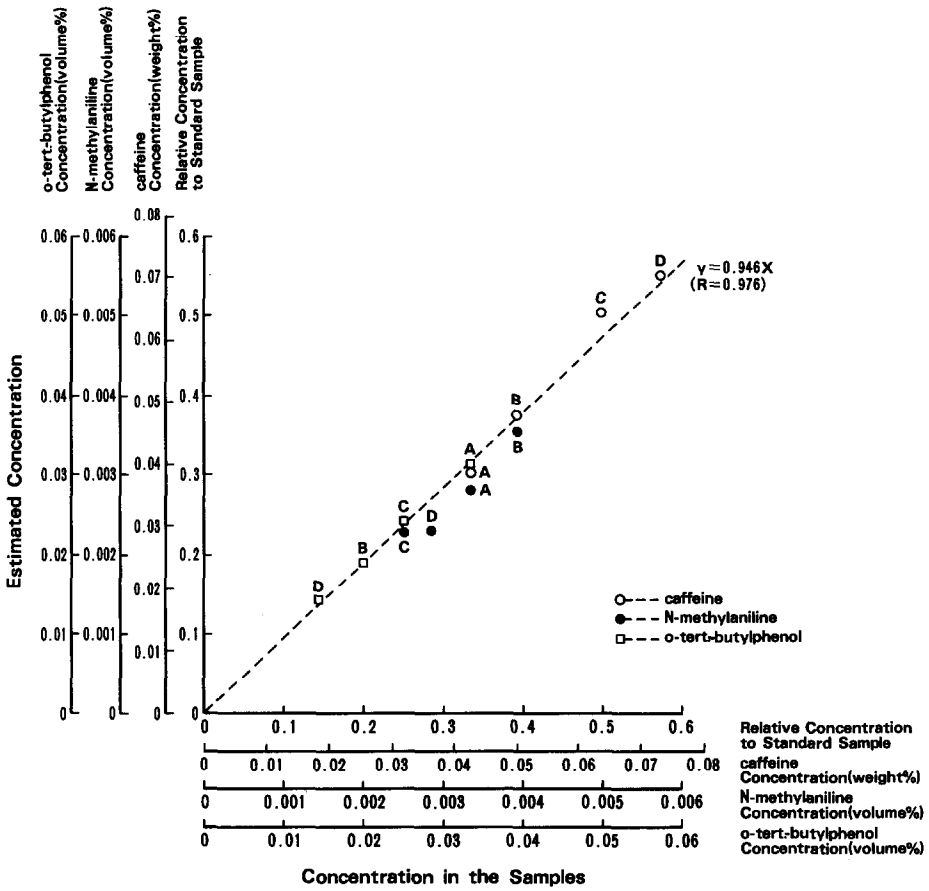

Fig. 7. Results of quantative analysis by the developed method. Eluent: acetonitrile–water (80.6:19.4, v/v). Resolution: R_s (caffeine/N-methylaniline) = 0.46; R_s (N-methylaniline/*o-tert.*-butyl phenol) = 0.33. Quantitation as in Fig. 6. A, B, C and D correspond to the samples shown in Table I.

Comparison of the developed peak resolution method and the conventional method based on principal component analysis and multiple regression analysis

In this work, the two methods for unresolved peak resolution were compared by resolving the same unresolved peak data obtained through computer simulations and HPLC separations of actual three-component mixtures.

One method was the developed method where elution profile estimation based on principal component analysis, spectrum estimation by extrapolating the observed spectra and rank annihilation to estimate the amount of a component were performed (method 1). The other was the conventional method where elution profile estimation based on principal component analysis followed by multiple regression analysis was performed (method 2).

The results for the artificial unresolved peaks showed that the performance of method 1 was almost the same in the two-component case and slighty poorer in the three-component case compared with that of method 2. In the computer simulation of

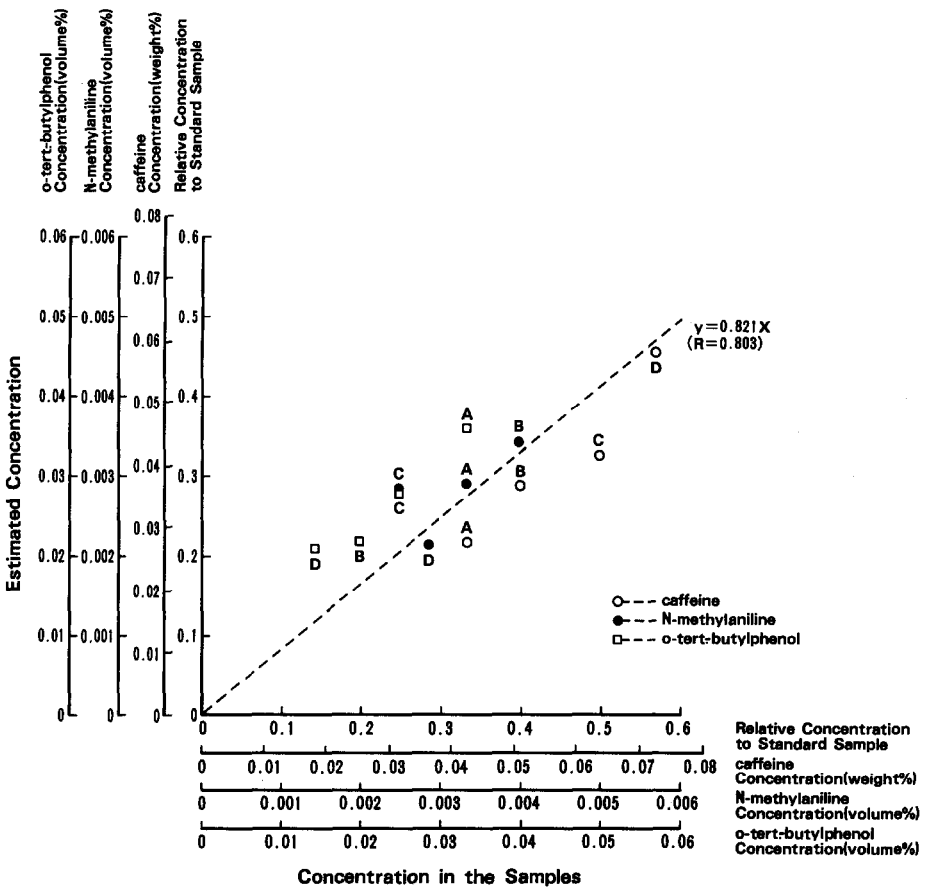

Fig. 8. Results of quantative analysis by the method of principal component analysis followed by multiple regression analysis. Eluent: acetonitrile–water (84.0:16.0, v/v). Resolution: R_s (caffeine/N-methylaniline) = 0.33; R_s (N-methylaniline/*o-tert.*-butylphenol) = 0.22. Quantitation as in Fig. 6. A, B, C and D correspond to the samples shown in Table I.

artificial unresolved peak resolutions, the elution profiles of three components were assumed to have the same simple Gaussian profiles with the same peak width. They did not show any tailing. The elution profiles were assumed to be ideal. Hence the condition of minimizing the entropy of time derivatives of elution profiles expressed as eqn. 12 was well satisfied. As a result, elution profile estimation was performed with good accuracy. The correlation coefficients between the estimated elution profiles and the true ones were greater than 0.99 except for three cases, (where they were 0.98). It is considered that if the elution profile of each component in the unresolved peak has an ideal shape, and the estimation can be made with good accuracy, we can obtain good qualitative estimations even by method 2. This was the case for the three-component artificial unresolved peaks. In method 1, the estimation errors come from the elution profile estimation and the spectrum estimation. Because the elution profile estimation was fairly good, errors in spectrum estimation were dominant in the total quantitative

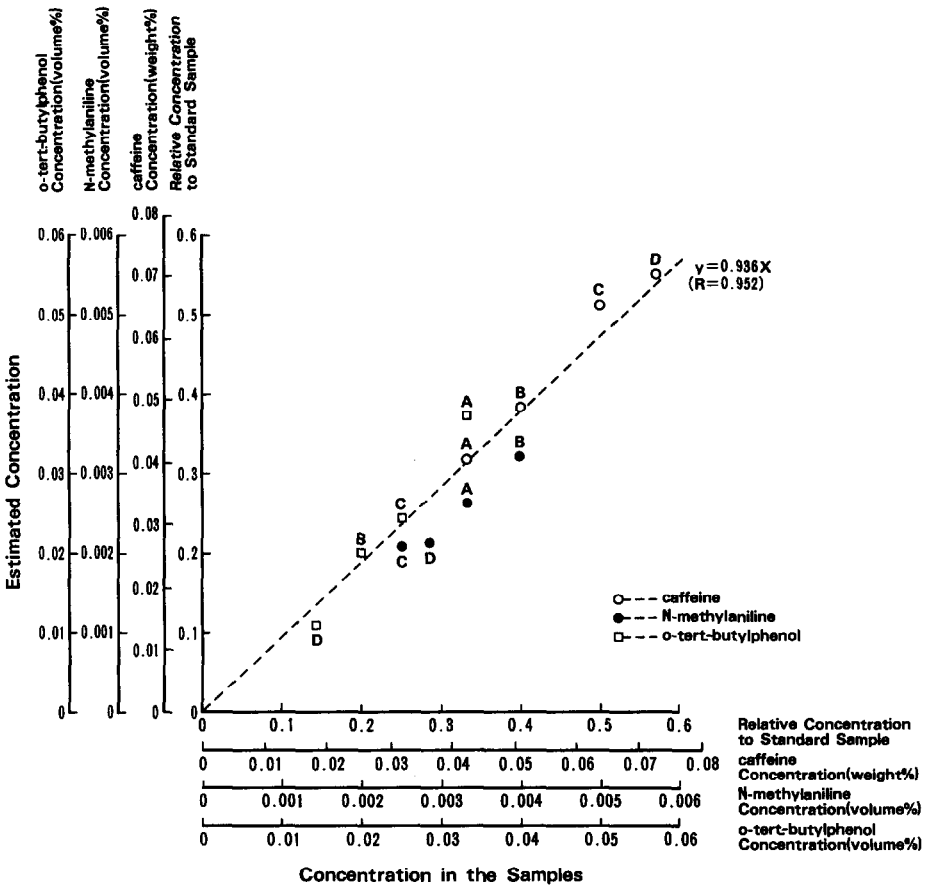

Fig. 9. Results of quantative analysis by the method of principal component analysis followed by multiple regression analysis. Eluent: acetonitrile–water (80.6:19.4, v/v). Resolution: R_s (caffeine/N-methylaniline) = 0.46; R_s (N-methylaniline/*o-tert.*-butylphenol) = 0.33. Quantitation as in Fig. 6. A, B, C and D correspond to the samples shown in Table I.

estimation errors in the three-component case. As a result, we could not estimate the peak areas of each component by method 1 with smaller errors than by method 2.

On the other hand, in HPLC separations of actual samples, the elution profiles of caffeine, N-methylaniline and *o-tert.*-butylphenol showed extensive tailing and the shapes of the elution profiles differed from each other, as shown in Fig. 4. They differed from the ideal elution profiles such as Gaussian. Minimization of eqn. 12 leads to localization (sharpening) of the elution profiles. In these situations, the method overestimates the elution profiles. The band width of the estimated elution profile of caffeine in Fig. 5 was narrower than that of the actual elution profile in Fig. 4. To express the tailing of real data, unrealistic peak tailings appeared in the estimation result. This result showed the problems that arise in elution profile estimation based on principal component analysis. In that method, a function expressing the ideal characteristics of elution profiles was introduced to estimate the elution profiles of

components. However, in the actual HPLC separation, the elution profiles sometimes differ from the ideal elution profiles. In this situation, the estimation method is not consistent with the actual elution profiles. Hence errors in the elution profile estimation directly affect the spectrum estimation in method 2, as it uses estimated elution profiles in the subsequent multiple regression analysis for spectrum estimation.

On the other hand, in method 1, the elution profiles and spectra are estimated independently. Even if the result of elution profile estimation is erroneous, this error will not directly affect the results of spectrum estimation. If the spectrum of the component in the unresolved peak is well estimated, the errors in quantitative estimation mainly come from the elution profile estimation. For actual sample mixtures, the elution profile estimation was not so good because of the existence of tailing, but the spectrum estimation was considered to have only small errors. Consequently, it is considered that the errors in the total estimation were smaller than those by method 2. Even when $R_s = 0.46$ and 0.33 and the elution profiles were estimated with smaller errors, the results showed an improvement in the accuracy of estimation by the developed method, as shown in Figs. 7 and 9.

When elution profile distortion such as peak tailing is observed, where the elution profile estimation by principal components analysis might be erroneous, the proposed method is considered to be superior to conventional methods.

An additional advantage of the developed method over the conventional method is that we can easily take the spectrum information into consideration when the absorption spectra of components in an unresolved peak are available before peak resolution. Although the elution profile stability is not so good in HPLC, as the elution profiles differ slightly in each separation and are affected by deterioration of the analytical column, it is often capable of obtaining reliable spectrum information about components in unresolved peaks. If we can obtain a standard spectrum of a component in an unresolved peak beforehand, we can utilize it immediately by the proposed method.

Feasible application area of the developed method compared with the experimental method for optimization of HPLC separations

Generally, optimization of HPLC separations by experimental approaches such as selection of columns and eluents and adjustment of flow-rate and gradient sequences are performed in order to obtain good chromatographic separations. We must first try these experimental methods before using the developed peak resolution method. There is no need to resort to the chemometric approach when we obtain good data by actual experiments. Chemometric approaches such as the developed method are only useful when these experimental approaches are impossible or very difficult to use. It is considered that this condition is often encountered in the analysis of biological fluids. In the HPLC separation of biological fluid samples such as urine and serum, there are many co-eluting compounds together with the compound of interest. The amounts of the co-eluting compounds fluctuates widely from sample to sample, or different kinds of co-eluting compounds may be present in different samples. The degrees of overlap may differ from sample to sample. As a result, even if one can find appropriate separation conditions for the analysis of a certain sample, these conditions may not be optimum for another sample. Under these conditions, the developed peak resolution method is useful for compensating for the low separation capability of the actual HPLC system.

CONCLUSION

A method for the resolution of unresolved peaks obtained by HPLC with multi-wavelength detection was developed. The method estimates the elution profiles and absorption spectrum of a component eluting at the rising edge or trailing edge of the unresolved peak and estimates the relative intensity of the derived three-dimensional chromatogram of one component by rank annihilation. The method can estimate peak areas with errors $\leqslant 10\%$ when $R_s \geqslant 0.4$. In comparison with the method based on principal component analysis followed by multiple regression analysis, the estimation accuracy was considered to be superior, especially when the elution profiles of the components are distorted from Gaussian, such as when tailing occurs, where the elution profile estimation by principal components analysis seems erroneous.

REFERENCES

1 D. W. Osten and B. Kowalski, *Anal. Chem.*, 56 (1984) 991.
2 B. Vandeginste, R. Essers, T. Bosman, J. Reijnen and G. Kateman, *Anal. Chem.*, 57 (1985) 971.
3 B. G. M. Vandeginste, F. Leyten, M. Gerritsen, J. W. Noor and G. Kateman, *J. Chemometr.*, 1 (1987) 57.
4 W. Lindberg, J. Oehman and S. Wold, *Anal. Chem.*, 58 (1986) 1415.
5 D. H. Burns, J. B. Callis and G. D. Christian, *Anal. Chem.*, 58 (1986) 1415.
6 P. J. Gemperline, *Anal. Chem.*, 58 (1986) 2656.
7 J. K. Strasters, H. A. H. Billet, L. de Galan, B. G. M. Vandeginste and G. Kateman, *Anal. Chem.*, 60 (1988) 2745.
8 W. H. Lawton and E. A. Sylvestre, *Technometrics*, 13 (1971) 617.
9 E. A. Sylvestre, W. H. Lawton and M. S. Maggio, *Technometrics*, 16 (1974) 353.
10 W. P. Carery, K. R. Beebe, E. Sanchez, P. Geladi and B. R. Kowalski, *Sensors Actuators*, 9 (1986) 223.
11 E. R. Malinowski, *Anal. Chem.*, 49 (1977) 606.
12 E. R. Malinowski, *Anal. Chem.*, 49 (1977) 612.
13 K. Sasaki, S. Kawata and S. Minami, *Appl. Opt.*, 22 (1983) 3599.
14 K. Sasaki, S. Kawata and S. Minami, *Appl. Opt.*, 23 (1984) 1955.
15 S. Kawata, H. Komeda, K. Sasaki and S. Minami, *Appl. Spectrosc.*, 39 (1985) 610.
16 C. N. Ho, G. D. Christian and E. R. Davidson, *Anal. Chem.*, 50 (1978) 1108.
17 C. N. Ho, G. D. Christian and E. R. Davidson, *Anal. Chem.*, 52 (1980) 1071.
18 C. N. Ho, G. D. Christian and E. R. Davidson, *Anal. Chem.*, 53 (1981) 92.
19 A. Lorber, *Anal. Chim. Acta*, 164 (1984) 293.
20 E. Aiyosi and K. Shimizu, *Nonlinear Programming* (in Japanese), Nikka-Giren, Tokyo, 1978, Ch. 10, p. 237.
21 I. Sakuma, T. Dohi, N. Takai, H. Funakubo, Y. Okabe, T. Katoda, Y. Sekiguchi and T. Isomura, *Annual Report of Engineering Research Institute Faculty of Engineering, University of Tokyo*, University of Tokyo, Tokyo, Vol. 42, 1983, p. 53.