

计算机应用于质谱中的差谱技术

乌德春 秦小鸿

(上海科技大学化学系色/质谱实验室)

〔摘要〕 本文报导一种具有一般意义的多峰相对强度计算机计算差谱方法。该方法不仅可用于质谱中的本底扣除和色谱分辨差的相邻两组份的质谱分离, 还可应用于多组份完全重叠混合质谱中已匹配或已知组份的扣除。

前 言

随着GC/MS应用领域的扩展与深入, 经常要求对复杂混合物的各个组份进行快速鉴定^{〔1,2〕}。但由于各种原因, 只能得到一张混合质谱而无法给出一个正确的答案。所谓混合质谱, 可以分为两类: 第一类是由于本底的干扰(如柱流失、离子源系统污染等)而引起的混合质谱。第二类是由于几个组份共存而引起的混合质谱。实际工作中, 此两种情况经常同时出现。当混合质谱的混合干扰达到一定的限度时, 一般的谱库检索方法难以给出可靠的结果。若混合质谱属于第一类, 采用通常的绝对强度本底扣除技术后进行谱库检索, 一般可以得到满意的结果。若混合质谱属于第二类, 简单采用以上方法就有很大的任意性, 得不到满意的结果。Biller-Biemann^{〔3〕}和Dromey^{〔4〕}等曾先后报导过以质量碎片轮廓分析为基础的混合质谱处理方法, 但由于算法对各个组份有一定的色谱分离要求, 不能推广到色谱共流出物或混合物直接进样的混合质谱处理。近年来发展起来的谱库逆检索技术^{〔5-7〕}与差谱技术^{〔8-10〕}的联用则是一个较好的方法。然而, 文献〔8〕报导的是一种独立峰相对强度差谱法, 差谱“质量”完全取决于人为选定的单个峰的“质量”, 因而不可避免地存在着盲目性与主观性。文献〔9〕所叙述的算法, 虽“基本”属于多峰相对强度差谱法, 但终因方法本身是从一个特殊谱库检索算法衍生出来的, 差谱“质量”与混合质谱中各个组份的碎片重叠状况直接相关。因此, 对混合质谱中各个组份碎片重叠较严重的处理, 几乎没有意义。除此之外, 文献〔8, 9〕所论述的方法都与他们各自的检索数据结构密切相关, 因而失去了一般意义。本文所要报导的是一个具有一般意义和主观性较小的多峰相对强度差谱法, 其应用面较为广泛。

差 谱 算 法

本文所叙述的算法, 应用了线性代数原理, 即将一张由n个质荷比及相对强度组成的质谱作为n维线性空间的一个n维向量 $V(v_1, v_2, \dots, v_1 \dots v_n)$ (其中i为质荷比, v_i 表示对应于

1984年4月2日收

i 的相对强度值)。这样,包含组份A和B的混合质谱就可作为一个 n 维向量 $AB(ab_1, ab_2, \dots, ab_i, ab_n)$,包含组份A的质谱就可作为 m 维向量 $A(a_1, a_2, \dots, a_i, \dots, a_m)$,问题就可被描述为 $AB-pA$ (其中 p 为比例因子)。在仪器能够完全重复两次相对强度实验数据的理想情况下,通过求出 $p_i = \frac{ab_i}{a_i}$ ($p_i \neq 0, i = 1, 2, \dots, q$)的最小值 p_{\min} ,令 $p = p_{\min}$,就可得到 p 。但事实是两次实验数据存在着一定的随机偏差,过大或过小的 p 值都将降低差谱的“质量”。因此,合理求出 p 值就是问题的关键。

算法首先假定 AB 中存在着这样一个分量 $A_1'(a_1', a_2', \dots, a_i')$,该分量不存在 B 组份的贡献。基于这一假定,算法第一步是在 A 向量的 m 维空间中求出 $p_i = \frac{ab_i}{a_i}$ ($p_i \neq 0$)的最小值 p_{\min} 。接着从 A 及 AB 中分解出 $A_1(a_1, a_2, \dots, a_i)$ 和 $A_1'(a_1', a_2', \dots, a_i')$, A 与 A_1' 分别由满足 $f_1 p_{\min} \leq \frac{ab_i}{a_i} \leq f_2 p_{\min}$ 条件的 a_i 及 a_i' 所组成(其中 f_1, f_2 是两个参量因子,且 $f_2 \geq -\geq 1$)。然后根据最小二乘法原理^[11a]求出 p_k ,使 $p_k A_1$ 与 A_1' 间距离最小。将 p_k 近似代替 p ,求出 $AB - p_k A$ 的差。如果 $ab_i - p_k a_i < 0$,令 $ab_i - p_k a_i = 0$ 。如果 ab_i 满足条件 $f_1 p_{\min} \leq \frac{ab_i}{a_i} \leq f_2 p_{\min}$,令 $ab_i - p_k a_i = 0$ 。最后进行归一化并剔除相对强度小于千分之一的峰。显然,方法可用于连续依次剔除混合质谱中多个已匹配或已知组份的贡献。

综上所述,差谱“质量”将完全取决于参量因子 f_1 和 f_2 的选择。 f_1 一般可取值为1, f_2 则据仪器的重复性而定。重复性越好, f_2 与 f_1 之差越小。如仪器能完全重复两次相对强度实验数据, f_1 和 f_2 都取为1,按上述理想情况处理。实际处理过程中,为避免盲目性和主观性,并不为 f_1 和 f_2 指定数值,而是由下式^[11b]

$$\cos(\alpha) = \left(\sum_{i=1}^l a'_{i1} a_i \right) / \left(\sum_{i=1}^l a_i^2 \sum_{i=1}^l a_i'^2 \right)$$

计算可靠性定量因子 R ($R = 1000 \cos^3(\alpha)$),综合考虑 R 值(最佳值为1000)及 A_1 的维数 l 后,再来确定 f_1 和 f_2 的取值。一个高的 l 值和 R 值将始终给出一个高“质量”的差谱。

实 验

一种差谱算法“质量”的检验,可通过差谱($AB - A$)与原始 B 谱的相似程度来衡量。基于这种思想,先做了两个实验,并将结果与文献[9]所述方法比较,以便考验程序与算法的“质量”。第三个实验是应用实例。

实验是在美Finnigan-MAT4510型GC/MS/DS仪器上进行的。数据系统由DG公司的NOVA4C小型计算机和Incos (Instrument Computer Operation System)软件组成,该机内存64KB,字长16位。在Incos软件的支持下,此计算机可使用DG公司的FORTRAN IV算法语言。利用Incos的现有程序,直接从仪器的采集数据中输入质谱图,然后用自编FORTRAN程序求差谱,再用Incos程序将差谱数据转换为Incos数据格式。这样,就可使用Incos的谱库检索等程序对差谱作进一步的处理,全过程可用框图表示(图1)。

表1 $I = 16P_K = 0.498$, $R = 999.985$

M/E	AB	A	B	WQB	FMB
50	211	239	158	158	24
51	588	676	432	432	45
52	75	77	63	63	9
53	26	53			
54	7	15			12
55	8	16			14
56	7	15			12
57	18	36			
61	5	9			9
62	25	26	21	20	7
63	50	54	40	40	12
64	17	23	10	10	29
65	101	147	48	47	
66	14	27			
73	7	10	5	4	12
74	49	55	37	37	7
75	18	37			31
76	51	48	47	46	21
77	1000	1000	865	866	379
78	108	116			33
79	246	493	87	87	
80	25	49			
83	3	5			5
84	10	19			17
89	17	35			
90	22	38	6	5	
91	86	70	87	87	84
92	7	14			
93	54	109			10
94	6	11			
95	11	22			
103	2		3	4	3
104	8		14	13	14
105	595	32	1000	1000	1000
106	58	21	76	76	76
107	142	285			

续表 1

M/E	AB	A	B	WQB	FMB
108	203	408			
118	1		1	1	2
119	39		67	67	67
120	208		359	360	360
121	19		32	32	33
122	2	3			3
123	165	331			
与B的相似指数	PURITY			1000	941
	FIT			1000	977
	RFIT			1000	949

(注:表中数值都是经千分之一归一化得到的相对强度值)

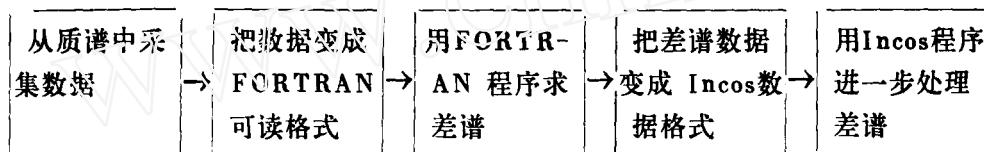


图 1

实验一

本实验中A与B数据(见表1)取自仪器携带的QDEMO色谱示范文件中的两个相邻流出组份(扫描号分别为1071和1015)。AB数据是利用仪器软件功能人为混合A与B数据得到的。WQB数据是利用本文的方法,取 $f_1=1$, $f_2=1.1$ (考虑到数据变化过程产生的误差,估取 $f_2=1.1$)进行AB谱与A谱相减得到的。将A和B分别存入仪器用户谱库后,用文献(9)的方法,将AB谱与A谱相减就得到表中FMB数据。表中最下一栏是用WQB及FMB数据对上建立的用戶谱库检索得到的与B的相似指数⁽⁹⁾(三个参数都是1000时表示两张谱完全相似)。

实验二:

本实验中样品C和D是一对分子式为 $C_{33}H_{24}O_3F_7As$ 的骨架异构体。样品CD是以任意比例混合C和D得到的。实验条件:离子源温度 $140^\circ C$,探头温度 $350^\circ C$,电离电压70电子伏特,灯丝电流0.25毫安。分别对样品C、D及CD直接进样得到图2,3,4。图5,6分别为本算法($f_1=1$, $f_2=2$)及文献(9)算法进行CD与D相减结果。图5,6中相似指数的获得及意义同实验一。

实验三

样品是一种漆的溶剂,要求分析此溶剂中的微量杂质。实验条件:填充柱OV-17,柱温 $100^\circ C$,汽化室温度 $150^\circ C$,进样量0.3微升,载气He,流量20毫升/分,方式GC/EI,离子源温度、电离电压、灯丝电流同实验二。实验得到的重建总离子流图见图7。从图中可知,扫描号为120的组份受到了严重混合干扰。图8是扫描号为120的质谱,利用本算法将扫描号120

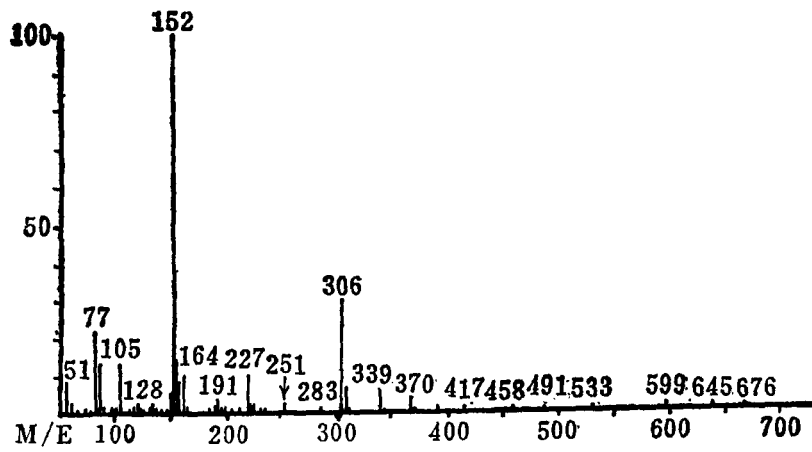


图 2

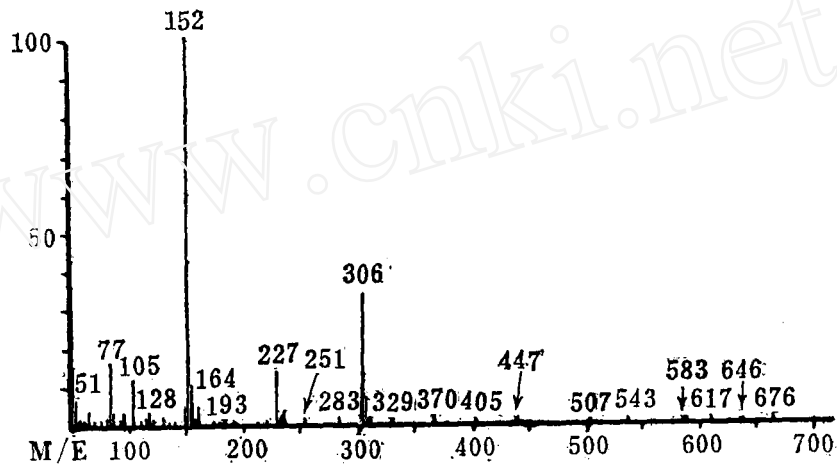


图 3

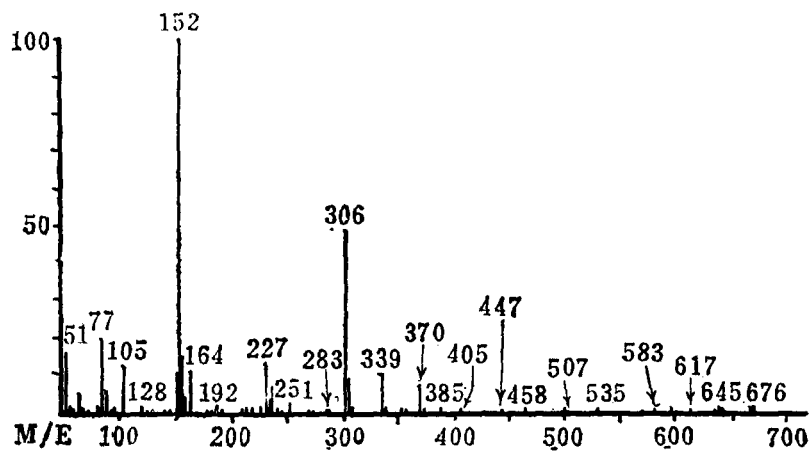


图 4

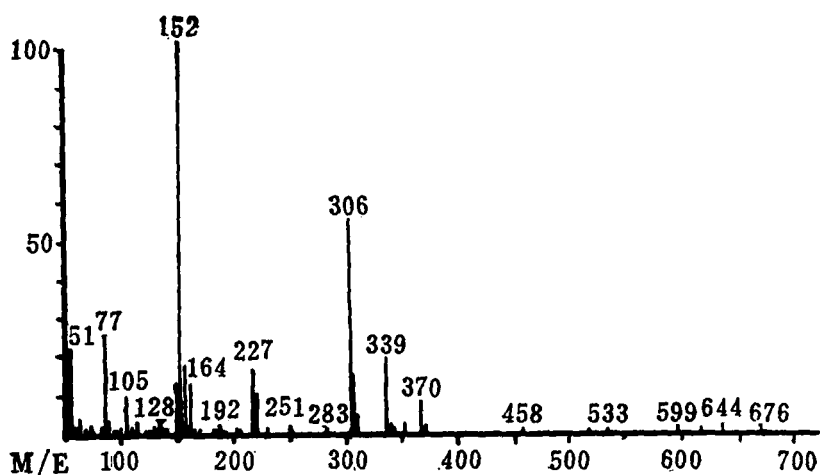


图5 (与C的相似指数

PURITY 865 L=6
 FIT 973 $P_k=0.527$
 RFIT 875 $R=960.951$)

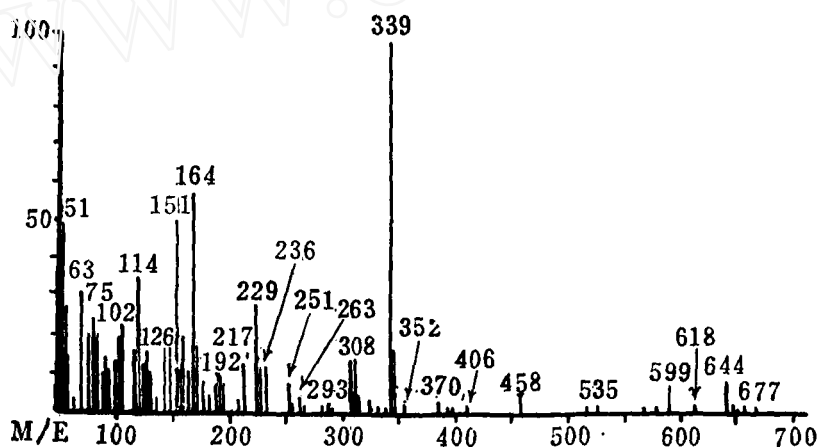


图6 (与C的相似指数

PURITY 147
 FIT 265
 RFIT 407)

和115的两谱相减 ($f_1=1, f_2=2$), 得到质谱图9, 将图9利用检索软件功能对仪器携带的NB谱库检索很快得到甲苯是最佳匹配结果(见图9)。

结果与讨论

从原理上讲, 本算法给出结果的好坏与混合质谱中各组份的相对含量无关。实验一所给出的结果表明, 在假定仪器能完全重复两次实验的相对强度值时, 算法给出的差谱数据(表1中WQB)与原始数据(表1中B)能几乎完全一致。而用文献(9)方法所给出的差谱数据

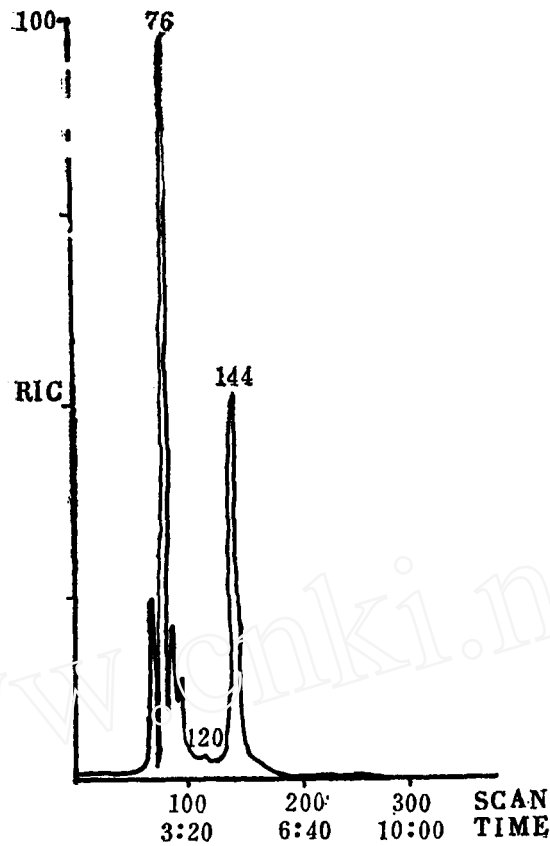


图 7

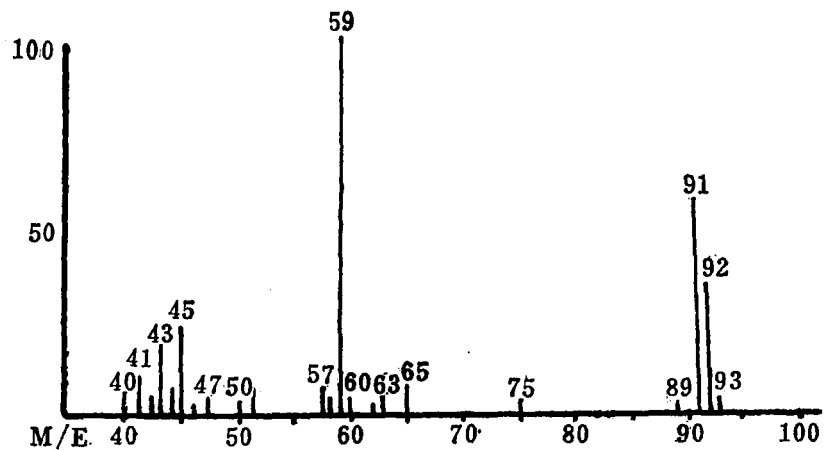


图 8

(表 1 中 FMB) 与原始数据 (表 1 中 B) 之间有较大的差异。由于实际上仪器并不能完全重复两次实验的相对强度值。因此, 本算法给出结果的好坏有时将与混合质谱中各组份的碎片重叠状况有关。一般较轻的重叠将给出一个可靠的结果。应之, 可靠性可能降低。实验二选择一对骨架异构作为分析对象, 正是为了考验算法在一种严重重叠情况下的实际应用能力。结果表明, 利用本算法得到的差谱 (图 5) 与原始谱 (图 2) 仍能较好吻合。相应地利用文

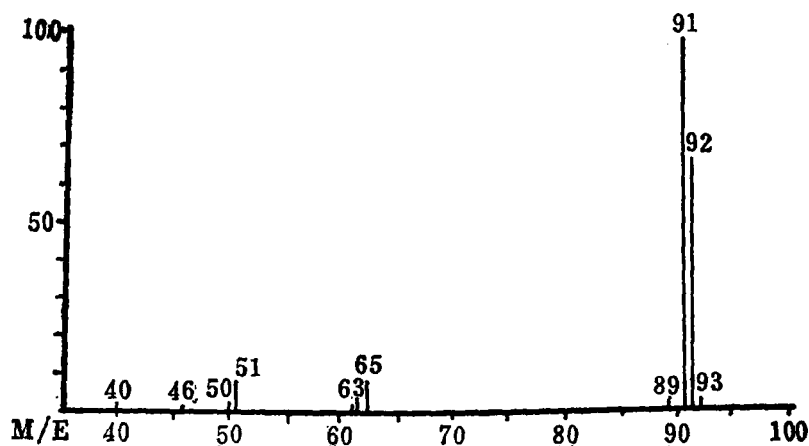


图9) 与甲苯的相似指数

PURITY 959 L=11

FIT 966 $K_k=0.985$

RFIT 982 $K=995.511$

献(9)方法处理得到的差谱(图6),几乎失去了意义。需要特别指出的是,实验二是一个混合物直接进样扣除已知组份(即处理多组份完全重叠情况)的实际例子,显然,这在有机合成等领域中有一定的实用价值。还需指出,即使对于一个完全未知混合物质谱来说,也有可能利用仪器所携带的Incos的Reverse Search程序获得主组份的谱信息,进行逐个扣除逐个解析。实验三是一个分离不好的GC流出组份的定性鉴定例子。当算法与仪器检索系统联用时,很快给出了令人满意的结果。这表明算法具有一般的本底扣除及分离两相邻色谱流出组份的能力。

以上三个实验证明了算法及程序的可靠性和实际应用能力。注意到相对强度较小的数据重复性有更大的偏移,程序开始设有可供分析者调节的“强度窗口”,使算法只能在此窗口内计算 P_k 值,从而保证了 P_k 值的“质量”。此外,程序还可以任意选择混合质谱中 l 个无干扰的峰计算 P_k ,从而提高了程序的实际应变能力。当 $f_1=1$ 时,一个高的 f_2 值和低的 l 值将提示分析者:最初被假定的混合谱中很可能不包含要被减去的那个组份,这是本算法的又一个特点。

以上是多峰相对强度差谱法在有机结构定性方面的应用。希望今后能利用本算法的一些特点,进行用一般方法难于处理的GC共流出物各组份的定量工作。

本工作得到黄关祿、徐信容、汪森超同志的大力支持,中科院上海有机化学所吴元伟同志帮助编写程序,上海科大化学系研究生孔宪起提供样品,在此一并致谢。

参 考 文 献

- (1) Burlingme, A.L., Shackleton, C.H.L., Howe, I, Chizhov, O.S. Anal. Chem. 50, 34bR, (1978)
- (2) Hites, R.A., Biemann, K. Anal. Chem. 42, 855 (1970)
- (3) Biller, J.E., Biemann, K. Anal. Lett. 7, 515 (1974)
- (4) Dromey, R.G., Stefik, M.J., Rindfleisch, J.C., Duffield, A.M. Anal. Chem. 48, 1368 (1978)

- (5) Abramson, F.P. *Anal. Chem.* 47, 45 (1975)
(6) McLafferty, F.W. Hertel, R.H. Villwock, R.D. *Org. Mass Spectrum* 9, 690 (1974)
(7) Pesyna, G.M. Venkataraghavan, R., McLafferty, F.W. *Anal. Chem.* 51, 1945 (1979)
(8) Atwater, B.L., Venkataraghavan, R., McLafferty, F.W. *Anal. Chem.* 51, 1945 (1979)
(9) Finnigan-MAT Incos MSDS manual No. 20000-90193, Vol. II Section 6.
(10) Finnigan Corporation Application Report No. AR8014
(11) 北大数学系《高等代数》, 人民教育出版社(a) P365 (b) P341

Computer Application to Subtraction Spectra in Mass Spectra

Wu Dechun, Qin Xiaohong

(GC/MS Laboratory, Chemistry Department,
Shanghai University of Science and Technology)

Received 2, April 1984

Abstract

This paper describes the general method in which the computer is used for calculating the subtraction spectra according to relative intensity of some peaks. The method is used not only for reducing background from mixture spectra and separating two components from two adjacent peaks of bad resolution of GC, but also for reducing the component which has been matched well or known in a completely overlapped mass spectrum of GC.