Instrumentation

Analysis of Fragrance Mixtures by GC/MS and the Computer

By Mingjien Chien, Senior Research Chemist, Givaudan Corporation, Clifton, New Jersey

In the last decade, gas chromatography-mass spectrometry (GC/MS) has become one of the most powerful analytical instruments used in the fragrance industry.

The advantages GC/MS offers are many. One of the principal advantages is that it utilizes the superior separating power of modern high resolution gas chromatography. Complex mixtures such as fragrances containing no less than one hundred compounds can be resolved with little difficulty. New coating techniques for both polar and non-polar capillary columns add an extra dimension in resolution. These new columns permit one to use higher elution temperatures and obtain reproducible retention patterns. If a fused silica type capillary column is used, an interface between GC and MS can be eliminated. The end of such a capillary column can be brought directly to the MS ion source, so that dead volume is reduced to a minimum.

Another outstanding advantage with a mass spectrometer is that it is a detector both universal and specific. The sensitivity is in the region of low nanograms. A fragmentation pattern is obtained which is specific to each compound in the mixture.

The most overriding superiority of the GC/MS lies in its further coupling to a modern computer data system which can process a huge amount of data generated in a single experiment. Today, computer processed data acquisition, quantitation and consequently the search of a data base for identification have become almost routine practice.

In the analysis of a typical fragrance sample, we normally try to identify each component separated by GC by searching a data base containing the mass spectra of reference compounds. In our laboratory, we have generated a data base containing about 10,000 fragrance or flavor materials. Fortunately, most compounds that we encounter are known and can be identified in this manner. However, if a compound is unknown, a considerable amount of effort may be needed to identify it by interpreting the fragmentation pattern in electron impact and/or molecular weight information derived from chemical ionization mode. In many occasions, other spectroscopic information, mainly nuclear magnetic resonance and infrared data, are necessary for a complete identification.

Another method commonly used in GC/MS is reversed spectral search in which the spectrum of a known compound is compared by the computer with each peak in the GC profile. The same process can also be conducted manually by displaying only selected masses as a function of retention time. This procedure can be used to determine the presence of a specific compound. It is particularly useful in conjunction with sensory studies, since odor is often attributed to compounds present only in trace quantities which in general do not show up in total ion current profile.

In the example in figure 1, the presence of anethole was suggested by a perfumer. When we
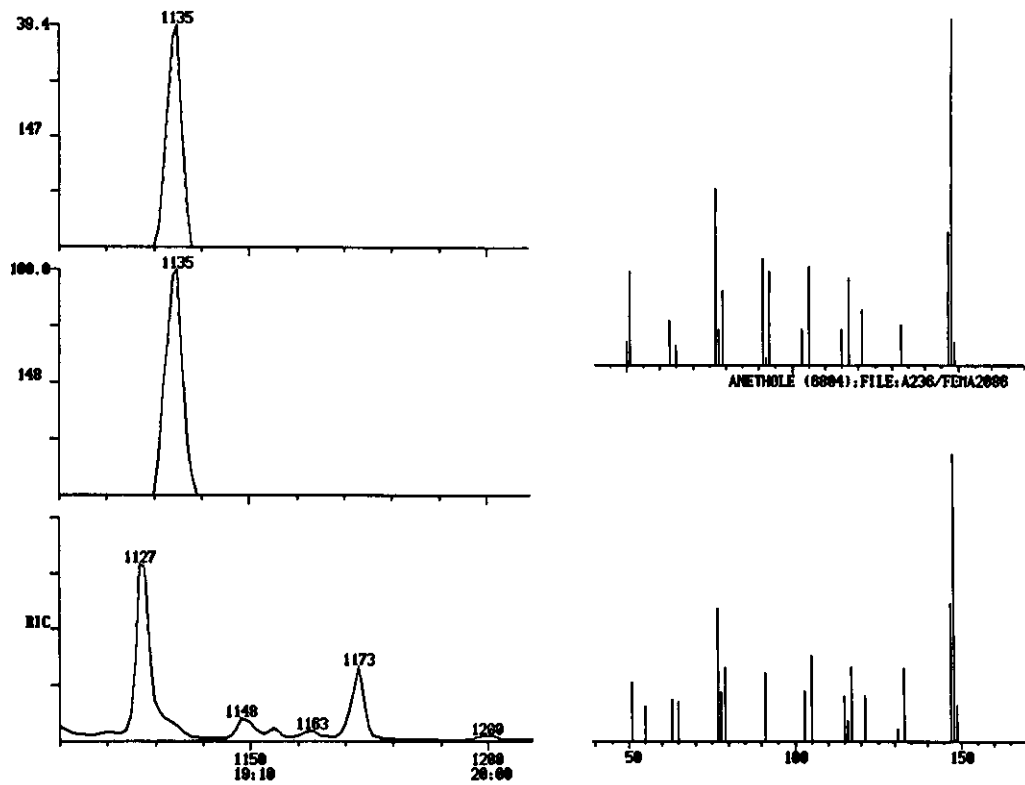
Figure 1. Left: Detection of anethole by plotting m/e 147 and 148. Right: Spectrum at data point (top) 1135 is compared to a reference spectrum of anethole (bottom).
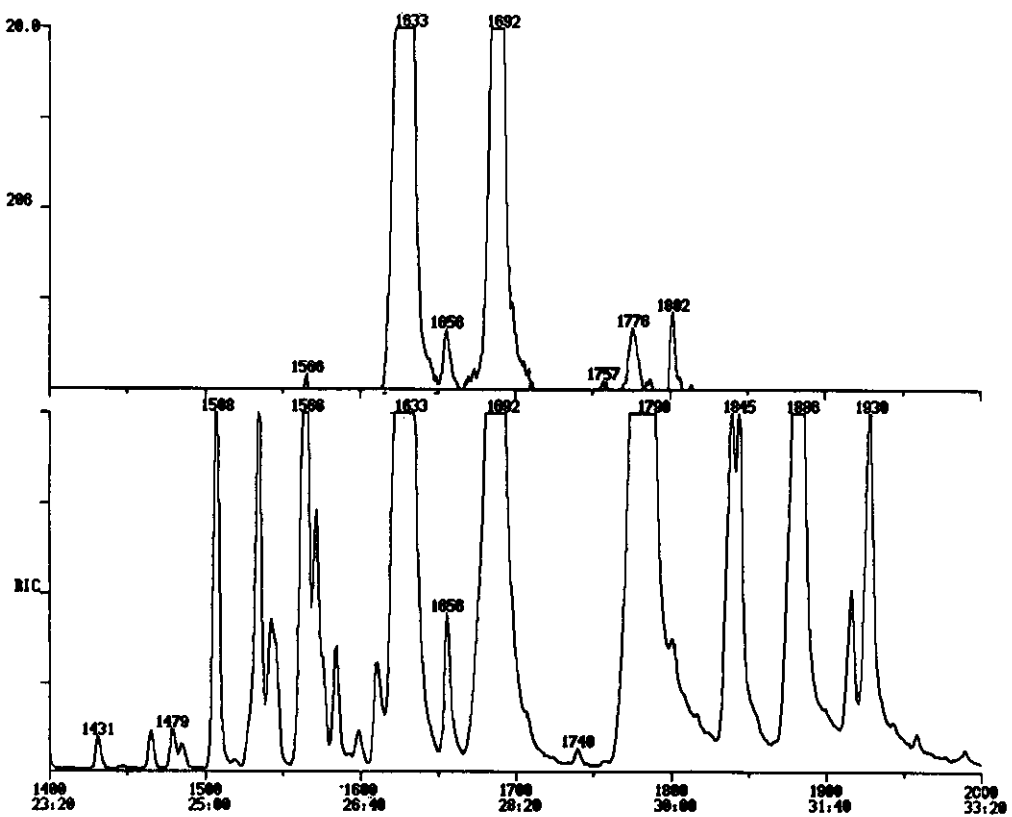

Figure 2. Detection of methyl ionone isomers by plotting m/e 206.

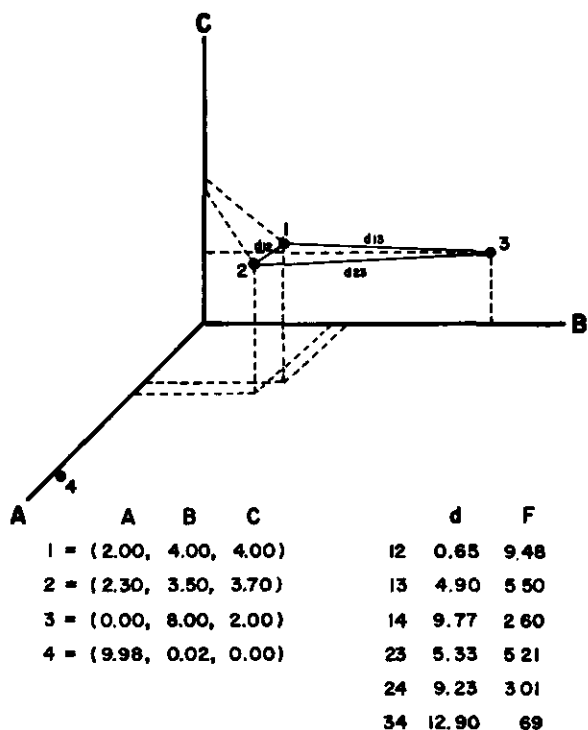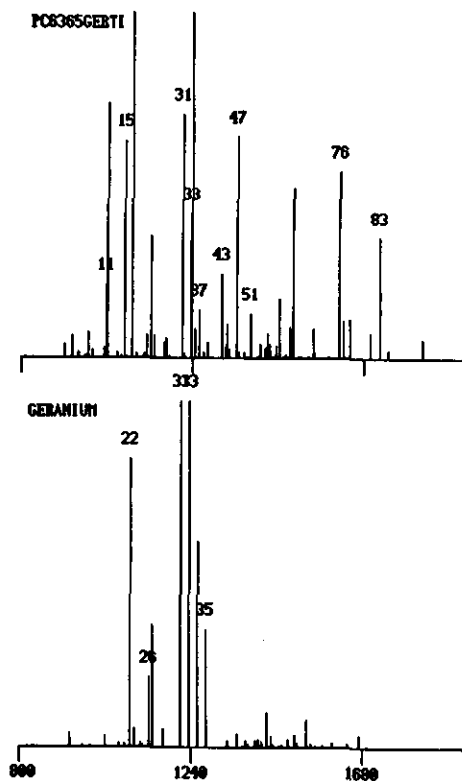|	**A**	**B**	**C**		**d**	**F**
I =	(2.00,	4.00,	4.00)	I2	0.65	9.48
2 =	(2.30,	3.50,	3.70)	I3	4.90	5 50
3 =	(0.00,	8.00,	2.00)	I4	9.77	2 60
4 =	(9.98,	0.02,	0.00)	23	5.33	5 21
				24	9.23	3 01
				34	12.90	69

Figure 3. Similarity among systems 1 through 4 composed of three components A, B and C. Bottom left: composition of the four systems. Bottom right: calculated euclidean distances (d) and fitting parameters (F).

SAMPLE: PC6365GERTI   DATE: 12/27/83
COMPARE TO FILES WITH   3 OR MORE INDEX PEAKS MATCHED

FILE#   FILE NAME   PEAKS   FIT

1: 1   GERANIUM   15   876

COMPARE 15 MOST IMPORTNAT PEAKS IN REFERENCE FILE
AREA(FILE)   AREA(SAMP)   NAME
1   999.99   1191.27   CITRONELLOL
2   621.56   780.68   GERANIOL
3   298.88   232.82   CITRONELLYL FORMATE
4   179.53   110.57   MENTHONE
5   171.21   77.92   GERANYL FORMATE
6   101.59   67.52   MENTHONE
7   40.17   8.71   GERANYL BUTYRATE
8   27.46   42.38   ALPHA-TERPINEOL
9   26.63   22.48   ROSEOXIDE
10   19.81   20.94   GERANYL ACETATE
11   17.44   5.31   GERANYL TIGLATE
12   18.66   45.37   CITRONELLYL ACETATE
13   18.56   3.16   CITRONELLYL BUTYRATE
14   16.49   8.85   GERANYL PORPIONATE
15   7.57   10.93   LINALYL OXIDE

Figure 4. Comparison between a perfume sample and reference file of geranium oil. Also displayed (right) is number of common peaks and their areas found in both sample and reference file.

plotted masses 147 and 148, which are characteristic of anethole, a small peak is revealed even though it is not resolved from another major peak in the mixture. Further confirmation is achieved by comparing the spectrum at this position, after background subtraction, with a reference spectrum of anethole.

In another application, fragments comprising a class of compounds can be displayed. This procedure is very useful in studying isomers which we often encounter in fragrance analysis. Mass spectrometry, in general, is considered disadvantageous in differentiating isomers because they produce similar fragmentation patterns. However, by displaying selected masses we are able to reveal one special class of isomers free of interference from other materials in the mixture. As an example (figure 2), we plot 206 to reveal all methyl ionone isomers in a perfume sample. Again, some of the peaks are so small that they fail to show up in total ion current plot. The identity of each methyl ionone isomer can then be determined simply by their retention order.

Thus, a major disadvantage of MS is turned into a useful analytical procedure.

Much information concerning the identity of each individual compound in the complex mixture is acquired by these procedures. However, in our experience, this type of information alone is often of very little value. To inform a perfumer that his sample contains such common substances as limonene, linalool or diethyl phthalate
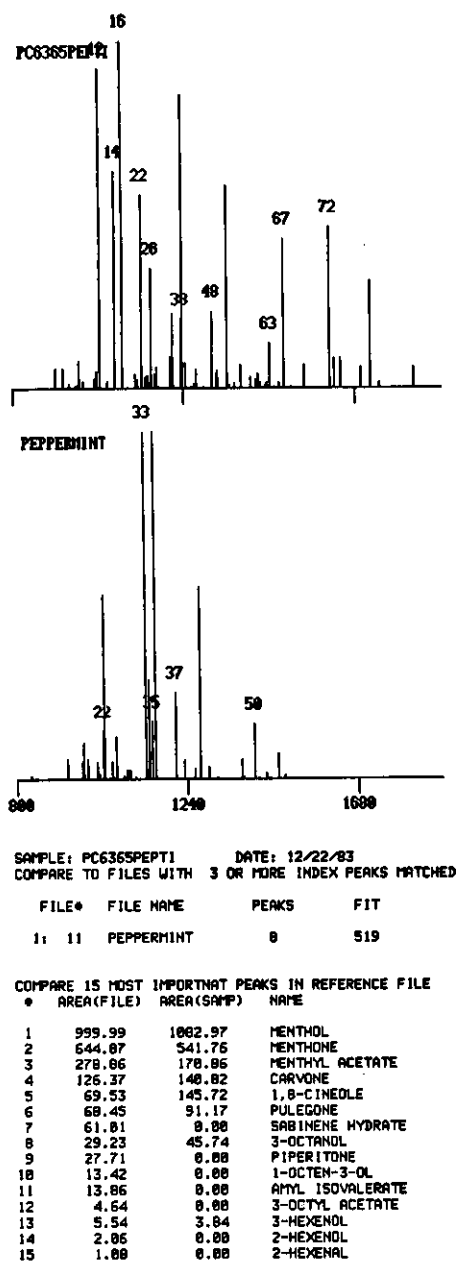
FILE#		FILE NAME	PEAKS	FIT
1:	1	GERANIUM	14	848
2:	18	BEGAMOT	9	719
3:	22	GERANIUM	13	694
4:	36	PATCHOULY	10	620
5:	15	CEDARWOOD	5	567
6:	9	CORIANDER	5	489
7:	6	LEMON	8	335
8:	14	CARDAMON	7	284
9:	27	CARROT SEED	6	267
10:	31	CARDAMON	6	266
11:	8	LIME	6	222
12:	21	CORMINT	4	190
13:	23	BASIL	3	168
14:	3	GRAPEFRUIT	6	157
15:	2	ORANGE	5	150
16:	28	BUCHU LEAVES	3	135
17:	11	PEPPERMINT	1	42

Figure 6. Computer print-out of the search result of an unknown sample.

FILE#		FILE NAME	PEAKS	FIT
1:	11	PEPPERMINT	8	519

```
COMPARE 15 MOST IMPORTANT PEAKS IN REFERENCE FILE
```

#	AREA(FILE)	AREA(SAMP)	NAME
1	999.99	1002.97	MENTHOL
2	644.07	541.76	MENTHONE
3	278.06	170.06	MENTHYL ACETATE
4	126.37	140.02	CARVONE
5	69.53	145.72	1,8-CINEOLE
6	68.45	91.17	PULEGONE
7	61.81	0.00	SABINENE HYDRATE
8	29.23	45.74	3-OCTANOL
9	27.71	0.00	PIPERITONE
10	13.42	0.00	1-OCTEN-3-OL
11	13.86	0.00	AMYL ISOVALERATE
12	4.64	0.00	3-OCTYL ACETATE
13	5.54	3.84	3-HEXENOL
14	2.06	0.00	2-HEXENOL
15	1.08	0.00	2-HEXENAL

Figure 5. Comparison between a perfume sample and reference file of peppermint oil.

means very little. On the other hand, there is indeed much valuable information concealed under the complicated GC pattern and can be brought out only through more sophisticated data manipulation. One of them is to recognize a group of compounds in the mixture as a single identity, e.g., an essential oil in a fragrance. Some of the computer methods that we have developed in our laboratory for this purpose are explained below.

A fitting parameter F_{12} is established between two multi-component systems 1 and 2, based on their composition determined by mass spectral methods described above. This parameter directly reflects the similarity between these two systems. The more similar they are in composition, the higher the F value is. A perfect fitting will have an F value of 1000. F_{12} is calculated based on the K nearest neighbor (KNN) classification rule of computer pattern recognition which was introduced to mass spectral matching by earlier investigators.[1,2] Excellent reviews on this type of pattern recognition and their applications in chemical analysis were given by Kowalski and Bender.[3,4]

Although the details in theory and computation of F value will be presented elsewhere, its principle can be illustrated by considering the example of tri-component systems in figure 3. Systems 1, 2, 3 and 4 are four imaginary systems composed of three compounds A, B and C with various concentrations. Systems 1 and 2 are very similar, 3 is less similar and 4 is most different. They are represented by points in a multidimensional space, each component being a dimension. The euclidean distance between two points clearly reflects the similarity between them (1 and 2 are closest in distance, while 3 and 4 are farthest). F calculated accordingly yields similar results: F_{12} being largest and F_{34} being smallest.

Of course, in reality we are dealing with a much larger number of dimensions. It is hard for human eyes to visualize anything larger than three dimensions, yet a computer can process data in a large number of dimensions without any dif-

ficulty. An example is shown in figure 4 where two multicomponent systems are compared. Each system is expressed in a bar graph of its GC profile where peak intensities are displayed as a function of retention indices.

Figure 4 shows how we detect an essential oil in a perfume sample by this method. The GC pattern of a perfume sample (top trace in figure 4) is compared to the pattern of each essential oil (bottom trace in figure 4) contained in a library followed by a calculation of F. If the pattern of a particular oil is recognized by the computer to be hidden in the more complex pattern of the sample, a high F results. To evaluate this method we conducted the following two types of experiments.

In the first experiment, we spiked a perfume sample with various essential oils and ran the program to determine whether the correct essential oils were recognized or not. Each time the computer came back with correct answers. In the first example (figure 4), the computer successfully detected geranium oil in the sample by recognizing all of its significant components. A high F value of 836 is obtained. However, the results in some other cases are not as clear cut.

In the next example (figure 5) where the sample was spiked with peppermint oil, some of the significant components are not recognized. This is because they are not resolved on GC and subsequently identified. As a consequence, F value is low and the detection of peppermint is somewhat ambiguous. Nevertheless, the computer pattern recognition method still points in the right direction by detecting other major components such as menthone, menthol and menthyl acetate. An eventual confirmation can be achieved by a closer examination.

In the second experiment, to evaluate this method, we made an attempt to determine essential oils in an "unknown" mixture prepared by a perfumer. The individual compounds in the mixture are determined by their retention indices and mass spectral data in a usual manner. Based on this information, the computer compares the pattern of the unknown with each essential oil in the library and calculates the F value. The results are printed out in a descending order of F value (figure 6). They are compared to the sample formulation provided by the perfumer (Table I).

Four of the essential oils used by the perfumers, geranium, cedarwood, patchouli and bergamot are detected without ambiguity. On the other hand, galbanum oil and absolute Burgeon de Cassis, which can be detected easily in a sensory study due to extremely high odor value of

Table I. Formulation of Perfume Sample Prepared by Perfumer

Formula	%
Geranium Oil Bourbon	5.0
Bergamot Oil Italian	15.0
Cedarwood Oil Texas	5.0
Patchouli Oil Indonesian	5.0
Galbanum Oil 1% DEP	2.5
Bourgeon de Cassis Abs. 1% DEP	2.5
Aldehyde C8	2.5
Aldehyde C9	2.5
Linalool Synthetic	5.0
Linalyl Acetate Synthetic	2.5
gamma Methyl Inonone	10.0
Lilial (Givaudan)	5.0
Galaxolide 50 (IFF)	15.0
Hedione (Firmenich)	5.0
Benzyl Acetate	12.5
Traseolide (Naarden) 100%	5.0
	100.0

the nitrogen and sulfur components, have not been picked up by the computer because of their low level of presence.

Thus, we are able to demonstrate how to use a computer pattern recognition method to obtain valuable information which is otherwise buried under a complicated pattern of the extremely complex mixture. This information is often complementary to sensory study.

In conclusion, we would like to point out that the human eye and nose are still the best pattern recognizers. A computer cannot replace a perfumer; however, a computer is much more efficient in processing huge amounts of data and performing repetitive calculations. In the analysis of complex mixtures such as fragrances, we are able to use the computer to select important information which will direct us on how to spend the valuable human effort so often pressed for time.

References

Address correspondence to Mingjien Chien, Senior Research Chemist, Givaudan Corporation, 125 Delawanna Avenue, Clifton, New Jersey 07014 USA.

1. G. T. Rasmussen and T. L. Isenhour, J. Chem. Inf. Comput. Science, 19 (3), 179 (1979)
2. D. P. Martinsen, Applied Spectroscopy, 35 (3), 255 (1981)
3. B. R. Kowalski and C. F. Bender, J. Am. Chem. Soc., 94 (16), 5632 (1972)
4. B. R. Kowalski and C. F. Bender, J. Am. Chem. Soc. 95 (3), 686 (1973)