Connectionist Hyperprism Neural Network for the Analysis of Ion Mobility Spectra: An Empirical Evaluation

Suzanne E. Bell,*,† W. C. Mead,‡ R. D. Jones,‡ Gary A. Eiceman,† and Robert G. Ewing†

Department of Chemistry and Biochemistry, New Mexico State University, Las Cruces, New Mexico 88003, and Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, New Mexico 87545

Received January 26, 1993

Spectra from ion mobility spectrometry (IMS) consist of low resolution information arising from complex gas-phase ion–molecular reactions. Such spectra lack specific structural information and have been considered inappropriate for analysis by traditional library search and pattern recognition methods. A data set of 292 mobility spectra from 12 organic compounds was treated using a connectionist hyperprism classification (CHC) neural network. Advantages of the CHC network for such spectra included high speed, linear training, and flexibility to adapt training parameters to minimize false negative identifications. The principal disadvantages were a sensitivity to unnecessary spectral information in the training vectors and a high false positive ratio. However, sensitivity exhibited by the CHC neural network assisted in identifying dispensable spectral information and in suggesting the importance of spectral features.

INTRODUCTION

Ion mobility spectrometry (IMS) was introduced over 20 years ago[1] and has recently attracted interest as a portable environmental monitor of hazardous organic compounds.[2] The principal attractions of IMS include small size, low power needs, and light weight. Additional advantages include simple instrumentation, high sensitivity and selectivity to certain toxic compounds, convenience of operation, and comparatively low costs. Examples of successful applications of ion mobility spectrometers include detecting nerve agents,[3] monitoring for contraband,[4] screening toxic rocket propellant on suits of shuttle astronauts,[5] and analyzing ambient air for toxic vapors from industrial emissions.[6]

The central component of an ion mobility spectrometer is the drift tube where vapors are ionized and subsequently characterized. The formation of ions from sample vapors occurs through collisional charge transfers from a reservoir of reactant ions to the neutral vapors (Figure 1). Commonly, these reactions are gas-phase proton or electron transfers as shown in eqs 1 and 2. Ion characterization is accomplished

$$\underset{\text{vapor molecule}}{M} + \underset{\text{reactant ion}}{H^+(H_2O)_n} \rightarrow \underset{\text{product ion}}{M^*H^+} + H_2O \quad (1)$$

$$\underset{\text{vapor molecule}}{M} + \underset{\text{reactant ion}}{O_2^-(H_2O)_n} \rightarrow \underset{\text{product ion}}{M^*O_2^-} + H_2O \quad (2)$$

by injecting ions from the ionization region through an electric shutter[7] to a drift region where the ions move through a voltage gradient toward a detector plate. In the drift region, ions in air at ambient pressure acquire particular velocities measured as drift times through a fixed distance.[8] Ion velocity is governed by the collisional cross-section of the ion.

A mobility spectrum is a plot of detector current versus ion drift time (Figure 2) and contains qualitative and quantitative chemical information about the vapor. Qualitative information is collected from the drift time (in milliseconds) and from the number and shape of peaks.[9] Quantitative information is taken from the relationship between the relative intensities of the reactant ion and product ion peaks.[10] In the two decades

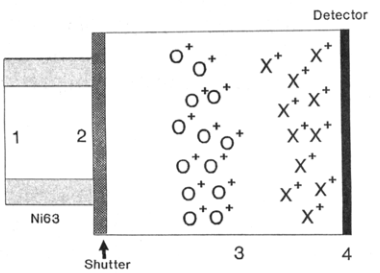

Figure 1. Ion mobility spectrometer. Sample molecules are introduced into the reaction region (1) via a carrier gas and ionized by collisional interactions with β particles from ^{63}Ni. Ion/molecule clusters form and line up along an electronic shutter (2). Pulsing of the shutter allows the ions to enter the drift region under the influence of an electrical field gradient. Separation occurs in the drift region due to differences in the size-to-charge ratios of the ions (3). In this example, the X^+ ions are have smaller size-to-charge ratios and thus a lower (faster) mobility than the larger, slower O^+ ions. Currents induced by arriving ions at the detector plate are converted to voltages (4) and plotted.

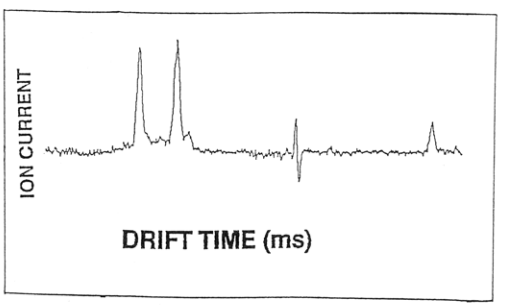

Figure 2. Example mobility spectrum. The x-axis is drift time of the ion/molecule clusters, and the y-axis is intensity of induced current at the detector plate converted to millivolts. The spectrum is simple with relatively few peaks.

following the introduction of ion mobility spectrometry, spectra were regarded as having a low information content. Methods of spectral interpretation developed for other ion separation methods (i.e. mass spectrometry) were considered inappropriate. Consequently, virtually no tools for spectral interpretation were explored, developed, or validated.

IMS operates at atmospheric pressure and thus presents a competitive and complicated environment in which ion/molecule reactions occur. Introduction of more than one type

† New Mexico State University.
‡ Los Alamos National Laboratory.

0095-2338/93/1633-0609$04.00/0

of sample molecule into the instrument can lead to complex interactions and produce indecipherable spectra. Use of a gas chromatograph (GC) as a sample inlet greatly reduces the chance that complex mixtures will enter the instrument and simplifies spectral interpretation. However, low spectral resolution remains a problem.

Previous investigations demonstrated that advanced environmental monitors could be configured around IMS.[11] However, the absence of automated data reduction and libraries of reference spectra limited the usefulness of an otherwise promising technology. Neural networks were selected for use with IMS because mobility spectra for various compounds exhibit common features. Such spectra are not well suited to library search methods. Moreover, rule-based interpretive systems are particularly deficient for mobility spectra due to the absence of any direct connection between drift time and chemical functionality. Traditional pattern recognition techniques such as cluster analysis were also dismissed as less flexible than the network approach.

Neural networks were regarded as promising because of their wide variety of learning modes, topology, and format of data presentation, all of which appeared compatible with the nature and quality of IMS data. The objectives of this work were to examine neural networks with specialized flexibility not offered by back-propagation networks. This flexibility also allowed the assessment of relevance of particular spectral features.

THEORY

Mobility Spectra. Interpretation of mobility spectra was initially thought to be straightforward since the spectra themselves are simple (see Figure 2). Peak shape can be attributed to four parameters including shutter pulse width, ion–ion repulsions, normal diffusion, and reactions in the drift region.[9] Beyond these factors, the relationships between chemical characteristics and spectral features becomes complicated.

Mobility–mass correlations have proven to be imperfect,[12] with many instances of nonlinear relationships. Mobility also depends on the location of the charge in the ion, charge delocalization, and the resulting interactions with the drift gas.[13] In addition, the formation of ion–molecule clusters in the drift region is now known to occur without obvious band-broadening in spectra.[8] Perturbations in base lines of mobility, once disregarded or ignored, are now known to reflect ion stabilities throughout the course of ion trajectory through the drift region.[14] Finally, humidity[15] and temperature[16] affect both the appearance of specific ions and the general profile of mobility spectra.

Unfortunately, the above processes are deduced using low resolution mobility spectra from hybrid instrumentation (IMS coupled to mass spectrometers) not realistic for field monitoring applications. In the absence of high resolution ion mobility spectrometers, mobility spectra will lack spectral detail or fine structure and will reflect a large number of complex molecule processes. In summary, the earlier premises regarding mobility spectra as not informative or too simple are just the opposite: the processes are exceedingly complex.

Neural Networks. Neural networks learn from empirical findings and require no direct knowledge of relevant chemical principles. This makes neural networks compatible with current knowledge and IMS capabilities. The history, theory, and implementation of networks with attention toward back-propagation methods have been thoroughly described[17] as have applications to analytical chemistry.[18] The connectionist hyperprism classification (CHC) network network developed at Los Alamos National Laboratory (LANL) shares some characteristics of Kohonen self-organizing networks and back-propagation networks.[19]

All three network types (back-propagation, Kohonen, and CHC) include multiple identical processing elements, an interconnection architecture, and iterative learning algorithms. The CHC learning algorithm shares the back-propagation of error that gave back-propagation networks their name. CHC-net shares with Kohonen self-organizing networks the iterative adjustment of basis function boundaries. Details differ among the three network types as described in the discussion of the CHC network below. Specific differences between the CHC and back-propagation network can be summarized as follows:

(A) CHC has a single hidden layer and a simple summation of outputs.

(B) CHC uses hyperprisms instead of sigmoidal basis functions.

(C) CHC trains basis functions widths in a self-organizing scheme.

Back-propagation has been favored for applications of networks to analytical chemistry[18] and to spectral interpretation specifically.[20,21] Back-propagation training is supervised learning in which training vectors, along with desired results, are presented. Back-propagation of error is used to adjust weights iteratively so that each training vector submitted to the network will produce the desired response. Thus, the training phase is an iterative approach to minimize the total error in actual output versus desired output for each training vector. Finally, back-propagation networks function even when a chemical process lacks reliable models as with IMS.

Limitations of back-propagation networks[18,19,21] include difficulty of design, long training times, large data set requirements, poor interpolation using small training sets, and weight saturation during training. In the case of spectral interpretation, the need for large training sets and weight saturation are the principal concern.

Weight saturation occurs when the output of a large number of processing elements in the network approaches the limiting value on either the high or low extreme of the transfer function. As the value of the output function approaches an extreme, the derivative of that function, the basis of weight adjustment, approaches zero. Weights are changed little if at all by the back-propagation of error, and the network can enter a paralyzed state. This can occur when the inputs to a network span a very large range (often the case with spectral data as input) and may be reduced through normalization procedures.

A frequent criticism of back-propagation networks applied to analytical problems is the use of small training and test sets.[18,21] Small training sets often lead to overtraining in which the network memorizes the training set and is unable to extrapolate the knowledge beyond the training space. Unfortunately, the amount of information available to train networks is often inescapably limited. The CHC network was devised to address the inevitable problems of a small data set and weight saturation while retaining the inherent advantages of the neural network approach.

CHC Network. The CHC network resembles clustering algorithms[19,22] and uses an iterative approach to define clusters in an n-dimensional input space. The nodes of the CHC network, analogous to the processing elements in a back-propagation network, consist of n-dimensional hyperprism basis functions, as shown in Figure 3. Each basis function has a vector center ($\mathbf{x}_c$), an amplitude, and width dimensions above (b_{ai}) and below (b_{bi}), the number of which depends on the

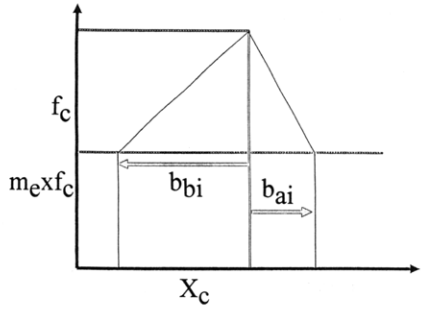

Figure 3. Basis function node of the CHC network. One basis function exists for each class-1 training vector used by the network. Training involves adjusting the amplitude and widths of the hyperprism iteratively to include as much of the input space around it as possible while minimizing enclosed class-0 data points. Refer to the text for a more complete discussion.

dimensionality (n) of the input space. At the center of the basis function, the output is equal to the weight (f_c) and the output of the function is nonzero in its connected space and zero outside of the connected space as defined by the boundaries of the hyperprism. When $x_{cij} - b_{bi} < x_i^p < x_{cij} + b_{aij}$ the ith component of the jth basis function is given by

$$f_{ij}(x) = f_{cij}\{1 - [(1 - m_e)(x_i^p - x_{cij})/b_{wij}]\} \quad (3)$$

where b_{wij} is the appropriate width parameter, either b_{aij} or b_{bij}. Each sample vector (**p**) consists of n-dimensional inputs x_{pi} together with a desired output o_p.

Within the active domain, the output of the basis function is the product of these n components. Outside the active domain, the output is zero. This construct is analogous to the activation function in the back-propagation network. The constant m_e was set to 1.0 in these experiments and the output of the CHC net is $g(x) = \sum f_j(x)$. The constant m_e is the same for all basis functions in the network and was normally set to 1.0.

The network training algorithm contains two parts. One part adjusts the basis function central amplitudes f_c based on the class-1 data while the other part adjusts the widths b_a and b_b based in all training data. The f_c values are adjusted to minimize the root mean square error in the network's calculation of the class-1 data points. The widths are adjusted according to a self-organizing algorithm,[23] adjusting the domain of each basis function to regulate the number of class-0 data points that fall within the active region and biasing the width according to a basis function overlap term. The training algorithm is iterative and is "on-line" in the sense that it need not be static but can grow or shrink to reflect the currently appropriate conditions.

The training algorithm for the jth network central amplitude f_{cj} weighs class-1 data by a manually adjustable parameter, w_1, according to

$$w_p = (w_1 \cdot o_p) + [(1 - w_1)(1 - o_p)] \quad (4)$$

An overlap factor used later in the width training is calculated by computing

$$f_{oj} = g(x_p)/f_{cj}^p \quad (5)$$

for the jth basis function. The iterative correction to the central amplitudes is then given by

$$f_{cj}^{p+1} = f_{cj}^p\{1 + [\alpha w_p((g(x_p) - o_p)]\} \quad (6)$$

where α is a learning rate parameter. For this work, α was typically set to 0.1 and w_1 was set to 1.

The training algorithm that adjusts basis function widths is slightly more elaborate. It includes two sets of terms: one adjusts widths based on data point inclusion, and another revises the widths according to the amount of basis function overlap.

The data point-based adjustment increases the appropriate widths as class-1 points are encountered within the basis function's active domain and decreases the widths as included class-0 points are found. The choice of what widths to adjust is made by calculating the distance of each point included within the active domain to each dimension's nearest basis function edge. This distance is denoted as d_{ij}^p. The data-dependent width adjustment is then given by

$$b_{wij}^{p+1} = (b_{wij}^p)^2 \cdot \{1 + [\beta(b_1 x_i^p + b_0(1 - x_i^p))]\}/b_{av} \quad (7)$$

where β is an overall learning rate (usually 0.5) and b_1 (50.0) and b_0 (−1) are the width adjustment factors associated with class-1 and class-0 data points, respectively.

The overlap-based width adjustment favors a moderate amount of basis function overlap. This part of the training also uses a self-organizing iterative algorithm, analogous to that used for the data point-based width training. On the basis of the overlap factor evaluated above, f_{oj}, the width adjustment is made by splitting the space into two regimes. If $f_{oj} > \chi$, then

$$b_{wij}^{p+1} = (b_{wij}^p)^2 \cdot \{1 + [\delta_+ f_{oj}]\}/b_{av} \quad (8)$$

On the other hand, if $f_{oj} < \chi$, then

$$b_{wij}^{p+1} = (b_{wij}^p)^2 \cdot \{1 + [\delta_- f_{oj}]\}/b_{av} \quad (9)$$

For this work, χ was normally set to 0.1, δ_+ to +0.1 and δ_- to −5.0.

The basis function construct imparts flexibility to the CHC network. Training of widths can be individually controlled so that the function is wider in some dimensions than in others. This allows for weighting of individual components as necessary or appropriate. Weighting, along with the multiple trainable parameters of the basis functions allows for smaller training sets and faster training times when compared to back-propagation networks.[19] The CHC network also is tolerant of extremes in weighting that tend to saturate back-propagation networks, and the problem of local minima is eliminated.[19]

A CHC network was created for each compound in the GC/IMS database. Thus, for each compound, all associated spectra were labeled as class-1 data and all other spectra in the data set, regardless of what compound obtained from, were labeled as class-0 data. Table I lists compounds used and the number of spectra used for each network. One basis function was assigned to each class-1 training vector in the hidden layer.

The basis functions can be thought of as initially residing coincident with each class-1 data point in the input space and training as the process of expanding or contracting the volume around those class-1 points to include as much of the input space as possible while excluding class-0 data points (Figure 4). The result is a region of connected space overlying class-1 data cluster(s). Widths of the basis functions could also be controlled using a basis function overlap term. Any test vector submitted to the trained network that yields a nonzero response in the enveloped class-1 data space is identified as class-1 data for the given network. Any test vector yielding a zero output is classified as a class-0 spectrum.

Each training vector (p) consisted of n inputs (x_{pi}) and the desired output o_p, either 0 or 1. Test vectors were identical

Table I. Compounds Examined

chemical class	compound	no. of spectra[a]	no. of class-0 spectra[b]
background[c]		22	270
ketones	acetone	21	271
	2-butanone[d]	36	256
	2-hexanone	27	265
aromatics	ethylbenzene[d]	20	272
	m-xylene	19	273
	p-xylene	19	273
	toluene	19	273
alcohols	ethanol	27	265
	2-propanol	30	262
miscellaneous	acrolein	30	262
	formaldehyde	22	270
total		292	

[a] Number of spectra of the compound (class-1), test, and training. [b] Number of spectra in the data set not associated with the target compound (292-N). [c] Background spectra collected prior to the elution of any target compounds. [d] Noncontinuous data sets, combined for analysis.

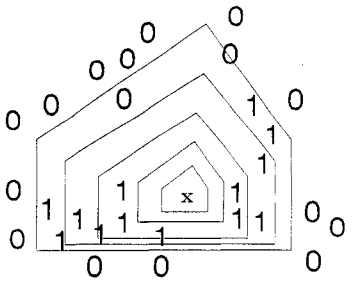

Figure 4. Expansion of a simplified basis function. Dimensions of a basis function are iteratively adjusted to maximize the amount of input space enclosed around the class-1 data point (x) while avoiding enclosure of class-0 points.

except for omission of o_p. The value of n was initially 6, the centroids and amplitudes of the three largest peaks in the spectrum.

EXPERIMENTAL SECTION

Instrumentation. The instrument used to obtain spectra was a prototype high-speed gas chromatograph–ion mobility spectrometer unit called the volatile organic analyzer (VOA; Graseby Ionics, Ltd., Watford, Herts, U.K.). The drift tube was heated to 150 °C and joined through an on-axial[24] design with a 10-m nonpolar GC column. The gas chromatograph was a model 2000 (Microsensor Technology Incorporated, Fremont, CA) modified with an automated air sampling trap with thermal desorption trap, temperature program elements, and a software/hardware interface to a portable MacIntosh computer. The hardware and software interfaces were custom built at Louisiana State University at the Environmental Studies Program and allowed microprocessor control of all GC parameters. Spectra from the ion mobility spectrometer were collected in an IBM-compatible 80386 25-MHz personal computer equipped with a commercially available digital signal averaging boards and the Advanced Signal Processing (ASP) software package (Graseby Ionics). Conditions for analysis are shown in Table II.

In the VOA, data collection is initiated as the shutter is pulsed and a spectrum of n number of data points (user-defined) is collected at a specified sampling frequency. If co-addition of spectra is required, the instrument can collect, add, and average the required number of waveforms. The spectra are stored in binary form and later converted to ASCII for postrun analysis. Real-time display and manipulation is available using the ASP software. For this study, spectra were generated for only the positive ion polarity.

Table II. Experimental Conditions

Gas Chromatograph	
carrier gas	helium
flow rate	set at 25 psi
temperature program	40°–140° ramp
VOA Operation	
drift gas	air
sheath flow	60 mL/min
cell flow	300 mL/min
VOA Acquisition	
waveforms averaged	4
no. of samples	640
frequency	40 kHz
waveforms	1
VOA Gating Pulse	
frequency	40 Hz
delay	0 μs
width	200 μs
source	internal

Procedures. Chemicals used were reagent grade or better. A few microliters of each compound was placed in a 1-L Tedlar bag which was then filled with nitrogen to capacity and gently heated. Vapors were drawn off by the sample pump on the GC and introduced into the trap and sample loop system of the instrument. Final concentrations were in the parts per million range in air.

Networks. The CHC network was constructed and executed on a Sun Sparc 2 (Sun Microsystems, Inc.) workstation. All of the code for the CHC network and preprocessors was written in "C" language at LANL. Training was typically accomplished in 20–40 "epochs" (iterative presentations of the training data set), in a period of 5–10 s, including display overhead. Testing the network's predictions on some 150 data points typically took less than 2 s, also including display overhead. Without the generous output produced, these operations could be significantly faster. In any case, the CHC net's execution time was short enough that it was negligible for present data set sizes and purposes.

RESULTS AND DISCUSSION

Performance of the CHC Network. In the first experiments with the CHC network, combinations of peak centroids, widths, and amplitudes were used as components of the training vectors. Of the 12 compounds analyzed, the networks associated with five (ethanol, 2-propanol, toluene, 2-hexanone, and background) performed well, and the networks associated with the remaining seven did not. Good performance was defined as ca. 75% or greater identification of test vectors and a false positive ratio of less than ca. 1.5. The difficult cases were selected for additional study and were re-examined to isolate the cause of impaired performance. The formaldehyde network was studied first.

The product ions formed by formaldehyde are comparable in size to reactant ions present in the IMS. A mobility spectrum of formaldehyde shows only subtle broadening of the reactant ion peak as opposed to the formation of distinct product ion peaks. A series of formaldehyde spectra are shown in Figure 5.

In the case of formaldehyde, the presence or absence of the compound is revealed in subtle changes in the appearance of the reactant ion peak. Surprisingly, the addition of width information to training vectors (full width at half-maximum) did not improve performance. Further examination of the data showed that although reactant ion peak width did change with the introduction of formaldehyde, the magnitude of that change varied considerably over the formaldehyde training set.

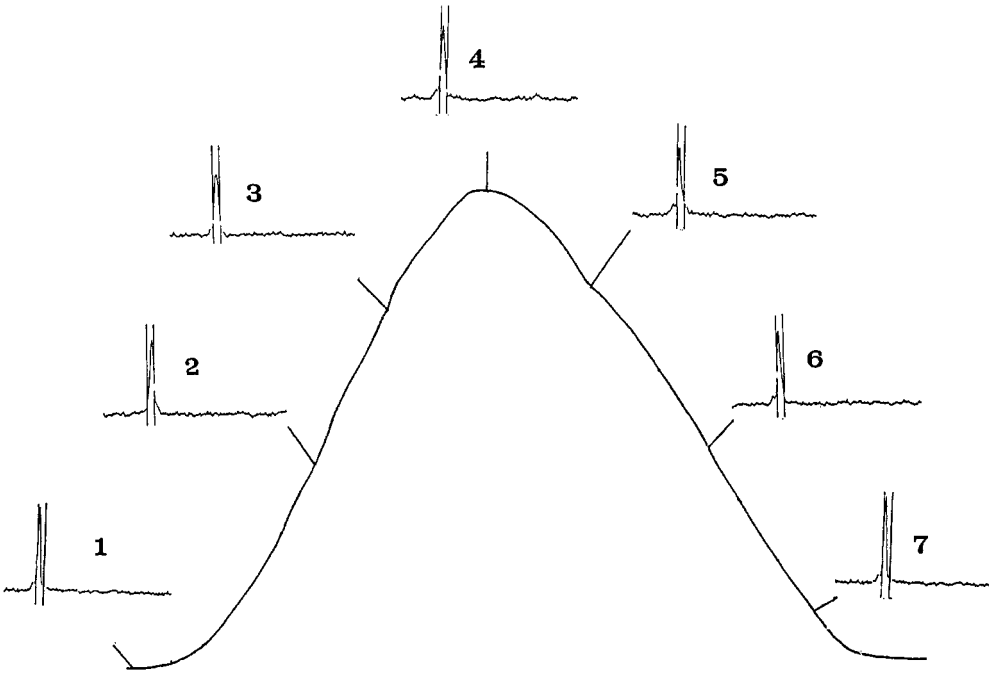

Figure 5. Formaldehyde elution profile. A mobility spectrum shows variability with concentration, analogous to but much more pronounced then the changes in ion intensities seen in GC/MS. Note the similarity between the background spectra and the formaldehyde spectra, and refer to the text for further discussion. The window delineates the region of the spectra where the reactant ion peak may be found as enumerated in Table III.

Table III. Comparison of Reactant Ion and Formaldehyde Product Ion Windows

	low value[a]	high value[b]	apex[c]	width[d]
		Reactant Ion[e]		
av	4132	4430	4281	298
std dev	16.6	80.8	39.9	85.1
% RSD	0.40	1.82	0.93	28.5
min	4100	4238	4238	200
max	4175	4350	4361	550
av − 2SD	4099	4268	4201	128
av + SD	4165	4361	4361	468
		Formaldehyde Product Ion[f]		
av	4185	4505	4345	320
std dev	55.4	61.5	41.3	41.3
% RSD	1.32	1.37	12.89	12.9
min	4125	4400	4263	250
max	4295	4600	4463	400
av − 2SD	4074	4382	4235	238
av + 2SD	4295	4628	4628	403

[a] Left (minimum) cursor position at estimated half-maximum of the peak. [b] Right (maximum) cursor position at estimated half-maximum of the peak. [c] Calculated apex (interpolated). [d] Calculated full width at half-maximum. [e] Taken from both background and target compound spectra. Times in microseconds; $n = 41$. Since the CHC network counts all nontarget spectra as class-0, reactant ion widths from background and target compounds must be considered. [f] Product ion, $n = 21$.

The spectral windows in which the reactant ion and formaldehyde product ion peaks are found are summarized in Table III. The valid region for the reactant ion ranges from roughly 4100 to 4600 μs while the valid region for formaldehyde ranges from 4075 to 4628 μs. An analyst introducing formaldehyde into the system knows the compound is present because the profile of the reactant ion changes as formaldehyde elutes into the instrument. However, the CHC network trains using spectra that are not presented in time sequence and without the benefit of observing evolutionary changes in the appearance of the spectrum. The result is that the overall changes in the appearance of the reactant ion peak profile brought about by formaldehyde is obscured by the large variations of the width of the reactant ion peak.

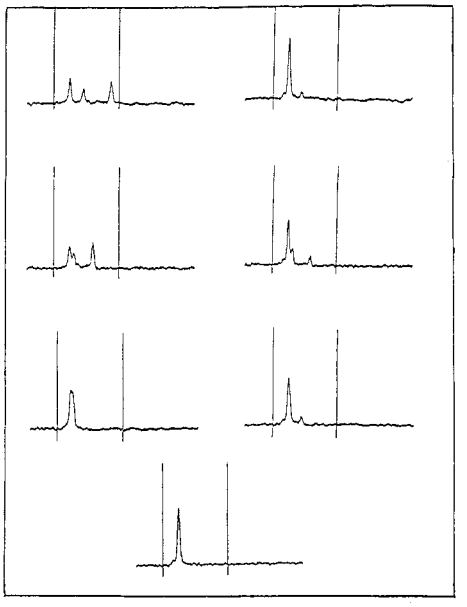

Figure 6. Similarity of spectra. In the right column are spectra at low concentrations, and in the left column are the corresponding high concentration spectra. The three compounds top to bottom are 2-hexanone, 2-butanone, and benzene. The bottom spectrum is a background spectrum of the reactant ion alone. Note similarities down the right column. A possible contaminant peak is visible in the right-hand benzene spectra.

Spectra of 2-butanone and 2-hexanone, while having distinct product ions, were confused by the network for a different but related reason. At low concentrations or on the leading edge of a GC peak, many mobility spectra are nearly identical (Figure 6). As concentration increases, distinctive spectral features develop. Thus, the CHC network appears to have overvalued the low concentration spectra, leading to confusion. Inclusion of more high concentration spectra should alleviate this problem. By forcing the network to generalize training over all possible patterns for a given compound, i.e., by training it on all spectra over the elution profile, similarities among

Table IV. CHC Network Performance[a]

	FN	TPF	FPR	FPF
	6-Inputs			
combined/all	61	0.60	1.87	0.08
combined/easy[b]	13	0.84	1.01	0.08
combined/hard	48	0.35	4.08	0.08
	3-Inputs			
combined/all	32	0.79	2.84	0.16
combined/easy[b]	13	0.84	1.01	0.08
combined/hard	19	0.74	5.07	0.21
ideal value	0	1.0	0	0

[a] FN, false negatives; TPF, true positive fraction, true positives/(true positives + false negatives); FPR, false positive ratio, false positive/true positive; FPF, false positive fraction, false positives/(false positives + true negatives). [b] These compounds were not redone using the three input networks. The results are included for reference.

spectra representing different species are greatly overemphasized.

Peak height was also found to be irrelevant by the network. The reasons for this are straightforward and similar to the peak width issue. Peak intensity is a function of concentration, not chemical structure. The variation of appearance of spectra with concentration has already been described. Until the advent of devices such as servo-controlled inlets for IMS,[25] concentration-dependent variations of spectra are inescapable. The variations in peak heights presented to the CHC network were similar in magnitude (typically 30% or greater) to the variability of width, and thus peak heights appeared to the network as random numbers as did peak widths.

The effects of including peak widths and/or amplitudes confirmed that the CHC network was sensitive to nonessential or misleading random information contained in the preprocessed files.[19,22] Subsequent experimental series were aimed at reducing the amount of extraneous information passed to the network. Amplitude and width values were eliminated from the training vectors, producing three-dimensional input vectors consisting of the centroids of the three largest peaks in the spectrum. Class-1 training vectors were duplicated until they comprised 20% of the entire data set, and the training set was shuffled to randomize the presentation. The results are summarized in Table IV. It is critical to note that spectra were simply duplicated to give the 20% figure. No additional spectra were experimentally recorded, and this point will be addressed in the concluding remarks.

Improvement is evident in the decrease in the absolute number of false negatives and in the increase in the total positive fraction. This outcome is due to expansion of the basis function into the input space, increasing the total volume of connected space, and reducing the number of class-1 data points excluded from that connected space.

The expanded volume of the connected space also accounts for the increase in the number of false positives seen between the three and six input results. Inevitably as the volume of connected space increases, the number of class-0 data points enclosed within it must increase, leading to more false positive identifications. The compromise is tolerable, since a doubling of the total positive fraction (TPF) comes at the cost of a 20% increase in the ratio of false positives (combined/difficult).

The high overall false positive ratio (FPR) (4:1 and 5:1) is directly attributable to the structure and function of the CHC network and to the nature of the IMS data. Expansion of the basis functions inevitably will enclose class-0 data points in the course of enveloping the maximum number of class-1 points. The closer the class-0 points to the class-1 points, the more pronounced the effects will be. Thus, the bunching of IMS data points in the input space is considerable and expansion of the basis functions causes an increase in false positive identifications.

The above cited problem can be addressed in several ways. The first, increasing the number of components in the input vector, has already been discussed. The CHC network views all information except the location of the mobility peak as extraneous. Thus, adding more information obtained from the mobility spectrum is not a viable solution.

The second approach would be to increase the dimensionality of the input through the addition of input vector components independent of the mobility spectrum such as chromatographic retention times. Chromatographic data can be utilized two ways. Retention indices could be used directly as network inputs although this presents many practical problems. Any GC/IMS system using such a network for data reduction would require that only one type and length of column be used under tightly controlled parameters to ensure reproducibility. However, these restrictions negate the flexibility of chromatography as a sample inlet.

The second approach, use of retention times as referee parameters external to the network, was used in early backpropagation experiments. In those cases when false positives did arise, the retention indices were separated by at least 30 units, more than enough for successful identification of the compound in question. This technique preserves the independence of the network from chromatographic information but still restricts the network to use with a specific GC/IMS device. The goal of this project was to create a spectral interpretation program for IMS data only, so chromatographic information was not incorporated into any further experiments.

Human Tests. A final experiment was conducted to compare the performance of the CHC network with humans, both expert and non-experts in IMS. The purpose was two-fold: to see if the CHC performed as well as people and to learn what features humans used to identify spectra. A training set of 2-hexanone spectra was printed out along with a test set of spectra from acetone, formaldehyde, benzene, and 2-hexanone. The test set was given to eight people, three with IMS experience and five without. The subjects were given the test spectra and asked to identify each as 2-hexanone (class-1) or other (class-0). Humans took the test spectra and overlaid them on 2-hexanone spectra for comparison, taking into account all variables such as peak height, location, and width. Thus, humans exercise a type of graphical library search method which took several minutes to complete, whereas the trained network responded instantaneously. The humans scored 81% correct while the CHC network scored 86% correct (17 and 18 correct out of 21, respectively).

CONCLUSIONS

The CHC network is a useful tool for the analysis of ion mobility spectra. The main utility of the CHC construct is the ability to flag spectral features that are considered as irrelevant to the algorithm. In so doing, the network points out the most chemically important characteristics of the spectra. Results of these experiments pointed to the peak location, corresponding to the mobility or drift time of the ion, as the only relevant information for identification of compounds from their spectra. Addition of theoretically important information such as peak width (corresponding to chemical events such as drift tube interactions) and height (related to concentration) only confused the network and thus indicated that these spectral features are not as important as was anticipated.

The CHC network suffered from a tendency to overemphasize similarities between low concentration spectra from different compounds. However, the fault is not solely that of the CHC construct. The low resolution nature of mobility spectra and selection of training sets purposely incorporating these low concentration spectra contribute significantly to this limitation.

One important finding, and one that sets CHC nets apart from back-propagation nets, was the ability to simply duplicate spectra to increase the size of the training set with accompanying gains in performance. No additional actual spectral acquisitions were required using the GC/IMS. Thus, an initial collection of spectra possessing distinctive features coupled to the ability to copy spectra rather than reacquire them should improve the performance of the CHC. Finally, meticulous control of the purity of spectra will be essential due to the sensitivity of the CHC network to extraneous information.

The CHC network was purposely designed to minimize false negative responses. In so doing, the number of false positives was extremely large and much greater than would be expected from a traditional back-propagation network. Although the problem could be addressed using independent information such as chromatographic retention times, the CHC is unlikely to function as a stand-alone spectral identification program in current form. The improvements mentioned above might well change this limitation. However, the CHC has proven itself extremely valuable in other respects and the ability of the CHC to identify relevant features might well be used in design of data sets for other types of pattern recognition and neural networks.

ACKNOWLEDGMENT

Funding for this work was provided by The United States Department of Energy, Los Alamos National Laboratory, and Krug Life Sciences (Subcontract No. 70149). The support of LANL groups X-1 and EM-9, particularly Don Dale, is appreciated.

REFERENCES AND NOTES

(1) Karasek, F. W. The Plasma Chromatograph. *Res. Dev.* **1970**, *21*, 34–37.
(2) Eiceman, G. A. Advances in Ion Mobility Spectrometry: 1980–1990. *Crit. Rev. Anal. Chem.* **1991**, *22* (1,2), 17–36.
(3) Blyth, D. A. A Vapour Monitor For Detection and Contamination Control. *Proceedings of the International Symposium on Protection against Chemical Warfare Agents*, 1983, Stockholm, Sweden; NTIS: Springfield, VA, 1983; pp 65–69.
(4) Karpas, Z. Forensic Science Applications of Ion Mobility Spectrometry. *Forensic Sci. Rev.* **1989**, *1*, 104.
(5) Eiceman, G. A.; Salazar, M.; Limero, T. F.; James, J. T. Detection of Hydrazines in Air Using Ion Mobility Spectrometry. *Anal. Chem.*, in press.
(6) Dam, R. J. Analysis of Toxic Vapors by Plasma Chromatography. In *Plasma Chromatography*; Carr, T. W., Ed.; Plenum: New York, 1984; Chapter 6.
(7) Bradbury, N. E.; Nielsen, R. A. Absolute Values of the Electron Mobility in Hydrogen. *Phys. Rev.* **1936**, *49*, 388–393.
(8) Spangler, G. E.; Collins, C. I. Peak Shape Analysis and Plate Theory for Plasma Chromatography. *Anal. Chem.* **1975**, *47*, 403.
(9) Revercomb, H. E.; Mason, E. A. Theory of Plasma Chromatography/Gaseous Electrophoresis—A Review. *Anal. Chem.* **1975**, *47*, 970–983.
(10) Eiceman, G. A.; Garcia-Gonzales, L.; Yang, Y. F.; Pittman, B.; Burroghs, G. E. Ion Mobility Spectrometry as Flow-Injection Detector and Continuous-Flow Monitor for Aniline in Hexane and Water. *Talanta* **1992**, *39*, 459–467.
(11) Bell, S. C.; Eiceman, G. A. Handheld Gas Chromatographic-Ion Mobility Spectrometry for On-site Analysis of Complex Organic Mixtures in Air or Vapors Over Waste Sites. *Report LA-UR-91-434CONF-910278-2*; Los Alamos National Laboratory: Los Alamos, NM, 1991 (NTIS availability).
(12) Lin, S. N.; Griffin, G. W.; Horning, E. C.; Wentworth, W. E. Dependence of Polyatomic Ion Mobilities in Ionic Size. *J. Chem. Phys.* **1974**, *60 (12)*, 4994.
(13) Karpas, Z.; Berant, Z. Effect of Drift Gas in Mobility of Ions. *J. Phys. Chem.* **1989**, *93*, 3021–3025, 3068–3077.
(14) Eiceman, G. A.; Shoff, D. B.; Harden, C. S.; Snyder, A. P. Fragmentation of Butyl Acetate Isomers in the Drift Region of an Ion Mobility Spectrometer. *Int. J. Mass Spectrom. Ion Process.* **1988**, *85*, 265.
(15) Kojiro, D. R.; Cohen, M. J.; Stimac, R. M.; Wernlund, R. F.; Humphry, D. E.; Norishighe, T. Determination of C1–C4 Alkanes by Ion Mobility Spectrometry. *Anal. Chem.* **1991**, *63*, 2295–2300.
(16) Bell, S. C. Ph.D. Thesis, New Mexico State University, Las Cruces, NM, 1991.
(17) Wasserman, P. D. *Neural Computing: Theory and Practice*; Van Nostrand Reinhold: New York, 1989.
(18) Zupan, J.; Gasteiger, J. Neural Networks: A New Method for Solving Chemical Problems or Just a Passing Phase? *Anal. Chim. Acta.* **1991**, *248*, 1–30.
(19) Jones, R. D.; Lee, Y. C.; Mead, W. C.; et al. Nonlinear Adaptive Networks: A Little Theory, a Few Applications. Los Alamos National Laboratory Report *LA-UR 91-273;* Los Alamos National Laboratory: Los Alamos, NM, 1991 (NTIS availability).
(20) Wythoff, B. J.; Levine, S. P.; Tomelllini, S. A. Spectral Peak Verification and Recognition Using a Multilayered Neural Network. *Anal. Chem.* **1990**, *62*, 2702–2709.
(21) Borggaard, C.; Thodberg, H. H. Optimal Minimal Neural Interpretation of Spectra. *Anal. Chem.* **1992**, *64*, 545–551.
(22) Internal memo, Los Alamos National Laboratory, X-1(6/92)194, 1992, available from the authors.
(23) Kohonen, T. *Self-Organization and Associative Memory*, 2nd ed.; Springer-Verlag: New York, 1988.
(24) St. Louis, R. H.; Siems, W. F.; Hill, H. H. Evaluation of Direct On-Axial Sample Introduction for Ion Mobility Detection after Capillary Gas Chromatography. *J. Chromatogr.* **1989**, *479*, 221–231.
(25) Cram, S.P.; Chesler, S. N. Analytical Fluidic Sampling Systems. *J. Chromatogr.* **1974**, *99*, 267.