

SANDIA REPORT

SAND97-0509 • UC-600

Unlimited Release

Printed March 1997

Automated Detection and Reporting of Volatile Organic Compounds (VOCs) in Complex Environments

P. J. Hargis, Jr., B. L. Preppernau, G. C. Osbourn, A. J. Ricco, G. C. Frye

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Approved for public release; distribution is unlimited.



Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
Office of Scientific and Technical Information
PO Box 62
Oak Ridge, TN 37831

Prices available from (615) 576-8401, FTS 626-8401

Available to the public from
National Technical Information Service
US Department of Commerce
5285 Port Royal Rd
Springfield, VA 22161

NTIS price codes
Printed copy: A08
Microfiche copy: A01

Automated Detection and Reporting of Volatile Organic Compounds (VOCs) in Complex Environments

P.J. Hargis, Jr. and B.L. Preppernau
Laser, Optics & Remote Sensing Department

G.C. Osbourn
Vision Science, Pattern Recognition & Multisensor Algorithms Department

A.J. Ricco and G.C. Frye
Microsensor Research & Development Department

Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-1423

Abstract

This paper describes results from efforts to develop VOC sensing systems based on two complementary techniques. The first technique used a gated channeltron detector for resonant laser-induced multiphoton photoionization detection of trace organic vapors in a supersonic molecular beam. The channeltron was gated using a relatively simple circuit to generate a negative gate pulse with a width of 400 ns (FWHM), a 50 ns turn-on (rise) time, a 1.5 μ s turn-off (decay) time, a pulse amplitude of -1000 Volts, and a DC offset adjustable from zero to -1500 Volts. The gated channeltron allows rejection of spurious responses to UV laser light scattered directly into the channeltron and time-delayed ionization signals induced by photoionization of residual gas in the vacuum chamber. Detection limits in the part-per-trillion range have been demonstrated with the gated detector.

The second technique used arrays of surface acoustic wave (SAW) devices coated with various chemically selective materials (e.g., polymers, self assembled monolayers) to provide unique response patterns to various chemical analytes. This work focused on polymers, formed by spin casting from solution or by plasma polymerization, as well as on self assembled monolayers. Response from coated SAWs to various concentrations of water, volatile organics, and organophosphonates (chemical warfare agent simulants) were used to provide calibration

data. A novel visual empirical region of influence (VERI) pattern recognition technique was used to evaluate the ability to use these response patterns to correctly identify chemical species. This investigation shows how the VERI technique can be used to determine the best set of coatings for an array, to predict the performance of the array even if sensor responses change due to aging of the coating materials, and to identify unknown analytes based on previous calibration data.

Acknowledgment

We gratefully acknowledge support of the work in this report by the LDRD program from 1994 to 1996. Phil Hargis and Bryan Preppernau would like to express their appreciation to Ben Aragon for his invaluable assistance in the construction of the high-accuracy vapor mixing manifold and to Fred Franklin for his invaluable assistance in the design and fabrication of the channeltron gate circuitry. Gordon Osbourn would like to express his appreciation for John Bartholomew for helping in developing the VERI algorithms for this application and for analyzing the SAW array data. For support in SAW coatings and vapor testing, Tony Ricco would like to thank Andrea Hoyt, Mark Hill, and Alan Staton and Greg Frye would like to thank Chris Colburn, Don Gilbert, and Richard Kottenstette.

Contents

Introduction	1
Pulsed Molecular Beam Expansion	2
Background	2
Gating Circuit	2
Performance of the Gated Detector	5
Toluene Detection Limit	8
Summary	10
Figures	
1 Schematic diagram of the channeltron gating circuit.	3
2 Digitized oscilloscope traces showing (a) the TTL trigger pulse and (b) the channeltron gated bias pulse.	4
3 Dependence of the photoionization signal on the delay of the channeltron gate pulse relative to the laser-induced photoionization event in the gas pulse. The laser is triggered at $t=0$. The top trace (a) is the signal from 53 ppb of toluene in helium, the bottom trace (b) is the baseline signal with no gas pulse present.	6
4 Wavelength scan of the (1+1) photoionization signal from the toluene S_1 origin (O_0^0) transition. The separation between the vertical lines represents measured laser linewidth. The drop in signal on the right side of the trace shows the signal baseline after the pulsed jet is turned off. The signal was obtained from 53 ppb toluene in helium.	7
5 Data from 1000 ppb toluene in helium used for signal-to-noise (SNR) measurements (a) with and (b) without the channeltron gating circuitry. Data points are not averaged.	8
6 Data from 1000 ppb toluene in helium used to determine the toluene detection limit. Each data point is an average of 100 laser pulses. The baseline is shown vertically expanded by a factor of 100.	9
Pattern Recognition	10
Background	10
Pattern Recognition Technique	11

Experimental Procedures for Developing Training and Test Data Sets	12
Data Preprocessing for Clustering-Based Pattern Recognition	13
Determination of Optimal Arrays	13
Evaluation of Expected Robustness of Array Performance	14
Results	14
Summary	14
Conclusions	15
References	16

Automated Detection and Reporting of Volatile Organic Compounds (VOCs) in Complex Environments

Introduction

The work discussed in this report addresses technological issues associated with the development of sensors for automated VOC (Volatile Organic Compound) monitoring and reporting. Our approach was to develop two types of sensors for VOC monitoring: (i) sensors based on the absorption of ultraviolet (UV) laser radiation and the subsequent detection of the fluorescence or photoionization signal and (ii) sensors based on surface acoustic wave (SAW) detection of VOCs. Intelligent algorithms that can both deconvolve complex chemical mixtures and, in addition, guide the development of an integrated sensor were also investigated in this work. Our goals were to (i) improve the ability of UV and SAW sensors to detect VOCs in complex environments, (ii) optimize the choice of chemically sensitive films for arrays of SAW devices, (iii) apply pattern recognition techniques to achieve chemical selectivity from a small number of nonselective films, and (iv) develop techniques to integrate multiple sensors in a robust, intelligent manner for fully automated VOC monitoring.

The work in this report has shown that by cooling VOCs in a pulsed molecular beam expansion, UV detection sensitivities in the ppb (part per billion) to low ppt (part per trillion) range can be achieved. Achieving these low sensitivities depended critically on the development of a gated channeltron detector to detect the photoionization signal resulting from the absorption of UV laser radiation. The work also led to the development of SAW coatings optimized for the detection of 28 representative compounds. SAW detection limits depend on the volatility of each particular compound, but generally range from 100 ppb for certain organophosphonates to ~100 ppm (part per million) for some highly volatile solvents. Included in the sets of coatings are those for water and CO₂, two common interferents in sensing scenarios. The two sensor approaches investigated in our work generally compliment each other in their sensitivity to specific VOCs.

VERI (Visual Empirical Region of Influence) algorithms were developed and optimized to analyze signals from SAW devices. Key features of the algorithms are the ability to handle nonlinear responses similar to those encountered in SAW devices and the ability to automatically recognize the existence of unknown signals, all with no adjustments or threshold variations. The algorithms run in near real time (seconds) on low-level platforms (486 PC or equivalent) and are suitable for field applications. Application of the algorithms to the analysis of UV data is not expected to require major new algorithm developments.

Pulsed Molecular Beam Expansion

Background

Circuits which can be used to rapidly gate the high-voltage bias for detectors used in photoionization experiments are becoming increasingly useful.^{1,2} In particular, circuits have been developed for use with microchannel plate (MCP) detectors.³ For the work discussed in this report, a gating circuit was developed for a channeltron detector used in experiments to detect resonant laser-induced multiphoton photoionization signals from organic vapors in a supersonic molecular beam expansion. Measurements were carried out to detect very low concentrations (ppb) of organic vapors in helium. At the lowest vapor concentrations, spurious ionization signals interfered with the detection of photoionization signals from the organic vapors. It was found that a gated channeltron detector can be used to reduce the contribution of spurious ionization signals to the channeltron output, thereby reducing VOC detection limits. The spurious signals are believed to originate from scattered laser light (below about 300 nm) and from time-delayed ionization of residual organic vapors in the vacuum chamber.

Gating Circuit

Figure 1 shows the gating circuit which was developed to gate the channeltron bias voltage. The circuit is based upon refinements to a previous circuit developed by Gostein and Sitz for use with MCP's.⁴ In the present circuit a TTL-compatible Siliconix D469 Quad High Current power driver chip is used to gate the high-voltage MOSFET BUK456-1000B. In the laboratory, the TTL-pulse output from a Stanford Research Systems DG535 digital delay generator is used to trigger one channel of the quad power driver.

Figure 2 shows the input TTL trigger pulse and the 1000 Volt gate pulse with a -1000 Volt DC offset. The gate pulse was measured using a Tektronix high voltage probe with the circuit connected to the channeltron detector. Both signals were digitized on a Tektronix 2430A digital oscilloscope and transferred to an AT-PC computer using National Instruments LabView™ 3.0 software. The input TTL trigger pulse has a 200 ns FWHM pulse width and a 1 ns rise and fall time. The output gate pulse shows a 50 ns turn-on time (as measured from the 10 to 90% voltage level) and a 1.5 μ s turn-off time (90 to 10%). After the circuit has completely turned on, the output pulse maintains a constant level for 400 ns. A constant bias level is important since the gain of the channeltron can vary significantly if the bias voltage changes as little as 100 volts. The approach used here to gate the channeltron takes advantage of the steep gain versus bias voltage curve of the channeltron. Thus, it is not necessary to switch the full bias voltage. Instead, one supplies a DC offset pedestal (nominally 1000 to 1500 Volts) from which to gate to the full voltage-on condition.

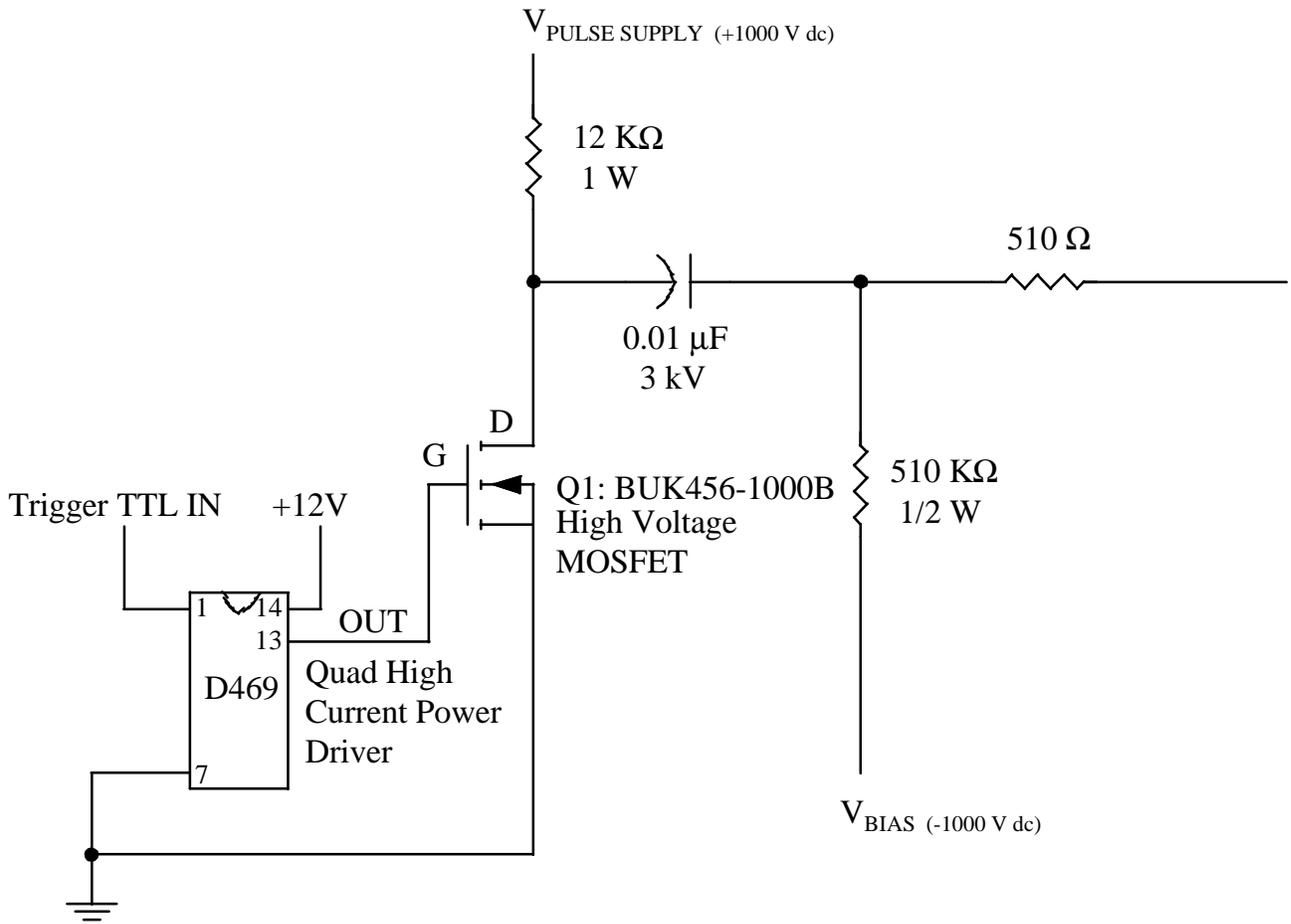


Figure 1. Schematic diagram of the channeltron gating circuit.

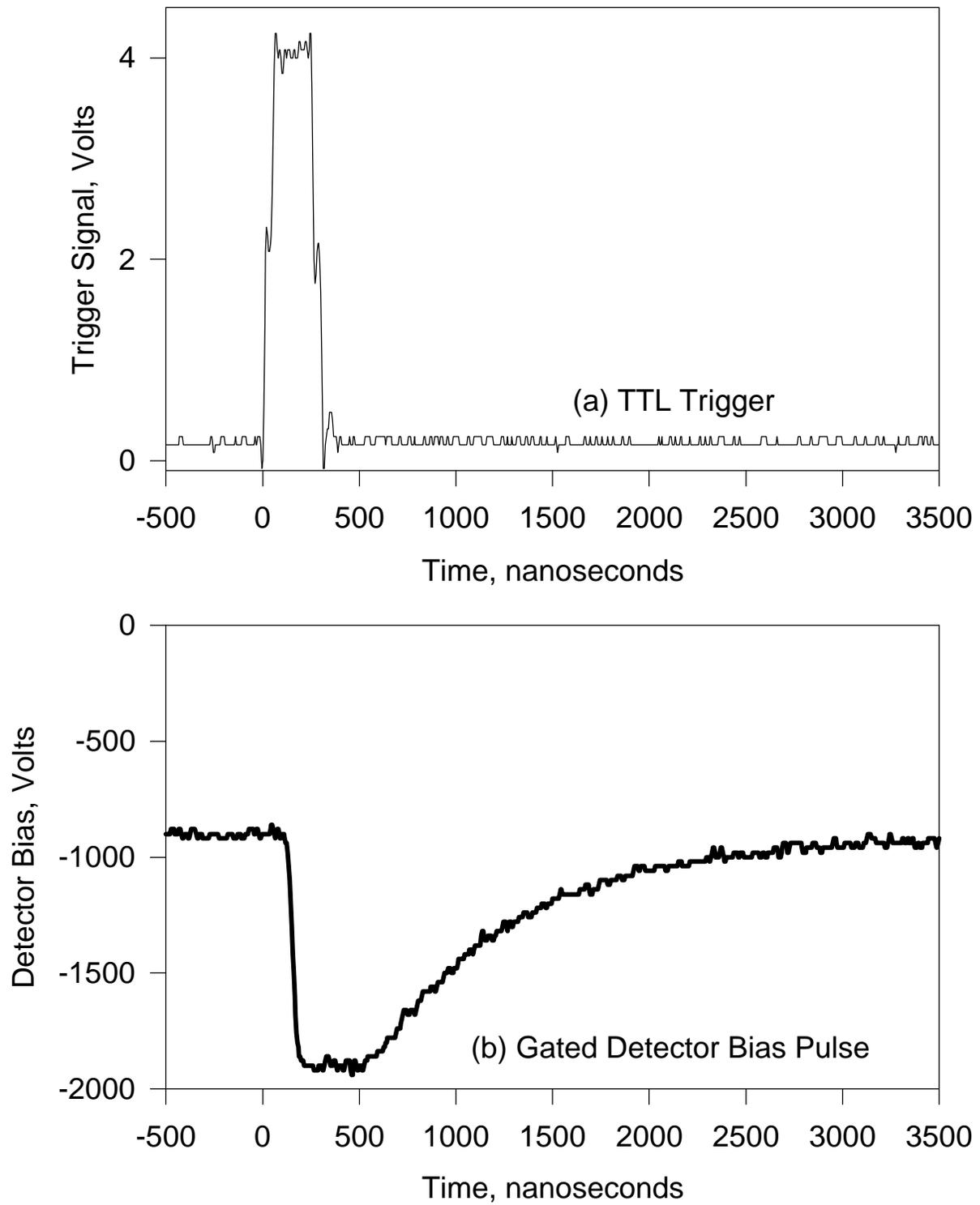


Figure 2. Digitized oscilloscope traces showing (a) the TTL trigger pulse and (b) the channeltron gated bias pulse.

Performance of the Gated Detector

Operation of the gated detector was evaluated in experiments to detect trace concentrations of VOC vapors in helium. In the experiments, the helium and VOC vapor mixture was expanded into a vacuum chamber through a pulsed nozzle. The cooled supersonic expansion was probed by a 10 Hz Quanta-Ray Nd:YAG laser used to pump a pulsed dye laser (Lumonics HD-500) which was frequency-doubled in a beta-barium-borate (BBO) crystal to generate tunable UV laser light. For detection of toluene, ultraviolet laser light at 266.827 nm (vacuum) resonantly photoionizes (1+1) toluene using the S_1 origin (O_0^0) transition.⁵ Toluene at concentrations in the mid-ppb range in 1000 torr of helium (calibrated using a high-accuracy vapor mixing manifold) was supersonically expanded through a General Valve pulsed nozzle with a 0.5 mm diameter orifice. The vacuum chamber was maintained at a pressure of 10^{-8} torr by a turbomolecular pump. The gas pulse was 200 ms long. The average vacuum chamber pressure increased to 1×10^{-5} torr, as measured by an ionization gauge, when the pulsed nozzle was operated at a pulse repetition rate of 10 Hz. The channeltron used in this experiment was a Galileo Model 4831G (1.10×10^8 gain at -2350 VDC).

In Figure 3 two traces are shown. These traces show the effect of changing the delay of the channeltron gate pulse relative to the laser-induced photoionization signal from the gas pulse. On the horizontal time axis in the figure, $t=0$ denotes the time at which the pulsed valve has fully turned on and the laser is fired. The time delay between the triggering of the gas pulse and the laser firing is held constant. The upper trace shows the signal output from the channeltron when the gas pulse is present and the concentration of toluene was 53 ppb. The bottom trace shows the same measurement when there is no gas pulse present. From the traces one can see the contributions from directly scattered laser light as well as from time-delayed ionization signals which are probably due to ions formed from photoionization of residual gas molecules in the vacuum chamber and molecules photodetached from the surfaces of the vacuum chamber. By adjusting the channeltron gate to correspond to the position of the main signal from the gas pulse, signal contributions, which would otherwise be integrated into the total photoionization current response, are rejected.

Figure 4 shows a typical spectral wavelength scan across the toluene resonant origin band using the gated channeltron detector. This data was also obtained at a concentration of 53 ppb toluene in helium. The separation between the vertical lines represents the linewidth of the laser as measured by a Burleigh pulsed wavemeter. The drop in signal on the right-hand side of the trace shows the signal baseline after turning off the pulsed jet. The signal was measured with a Stanford Research Systems Model 250 Boxcar Integrator operated in the last sample output mode. The last sample output signal from the integrator was digitized and input to a computer. For each data point, signals from 10 laser pulses were accumulated and then averaged. The relative response to changes in vapor concentration was determined to be linear over 3 orders of magnitude.

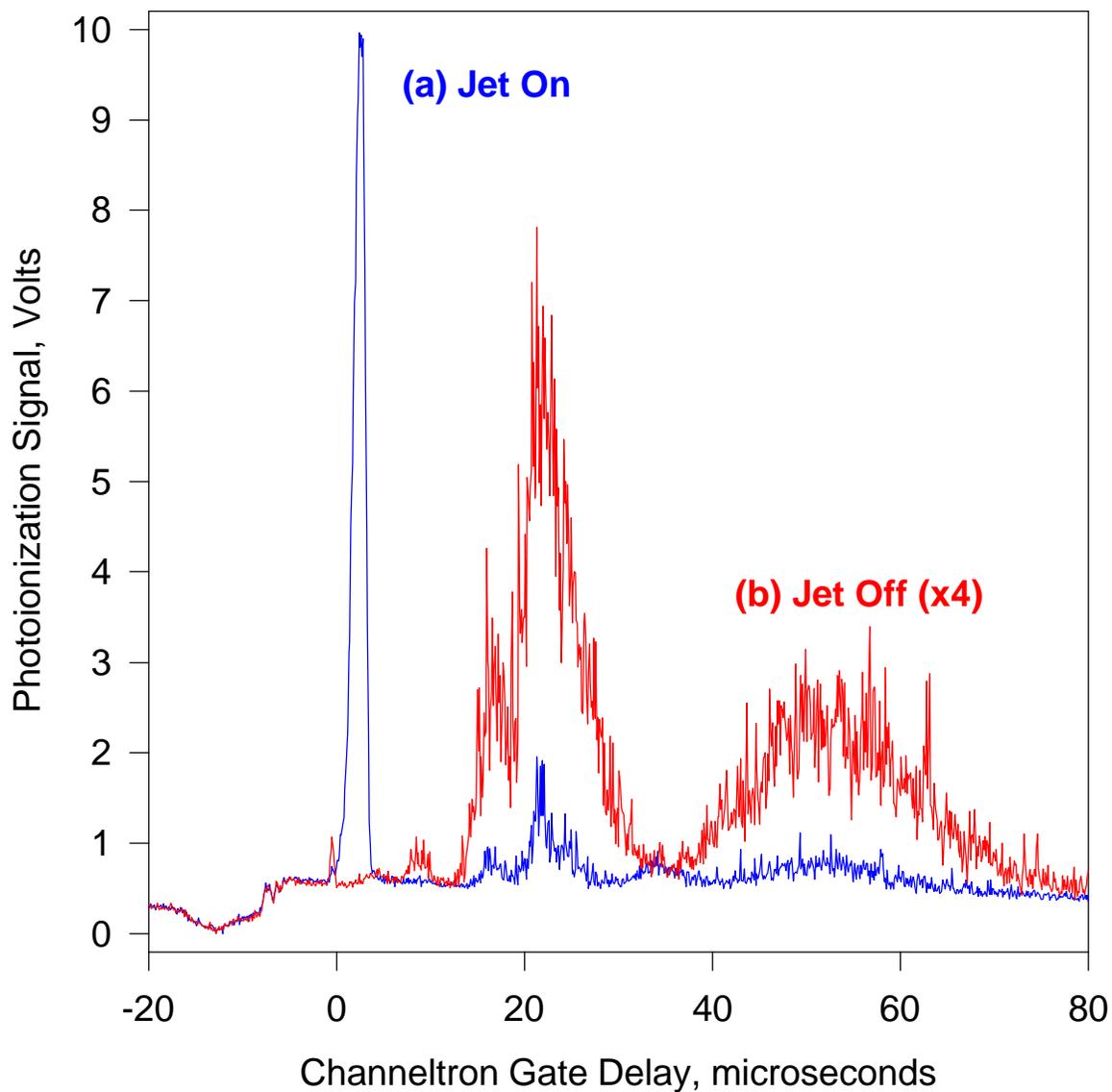


Figure 3. Dependence of the photoionization signal on the delay of the channeltron gate pulse relative to the laser-induced photoionization event in the gas pulse. The laser is triggered at $t=0$. The top trace (a) is the signal from 53 ppb of toluene in helium, the bottom trace (b) is the baseline signal with no gas pulse present.

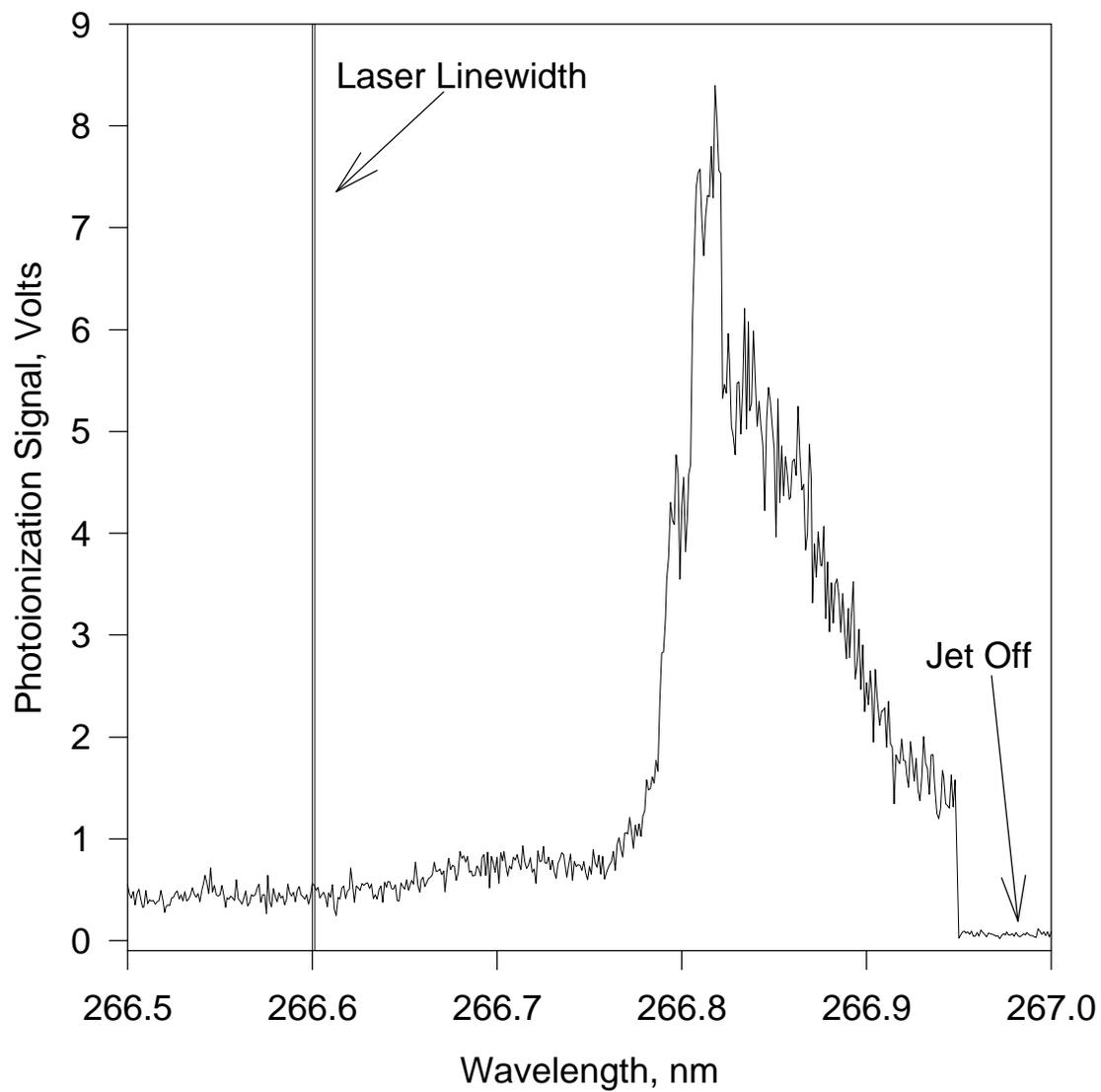


Figure 4. Wavelength scan of the (1+1) photoionization signal from the toluene S_1 origin (O_o°) transition. The separation between the vertical lines represents measured laser linewidth. The drop in signal on the right side of the trace shows the signal baseline after the pulsed jet is turned off. The signal was obtained from 53 ppb toluene in helium.

Toluene Detection Limit

A signal to noise ratio (SNR) determination was made by conducting an “off-on-off” experiment. For this measurement, the laser wavelength was set to the maximum resonant photoionization transition and the signal was measured both with and without the gas pulse present. The SNR is defined as the average value of the signal from the gas pulse divided by the standard deviation of the noise in the baseline with the gas pulse off. A comparison showing the benefits of gating the channeltron is shown in Figure 5. The bottom trace is the signal without gating (constant -1500 VDC bias) and the top trace shows the effect of gating the detector (gate pulse is 400 ns wide and -500 Volts amplitude with a 1000 Volt DC offset bias). The concentration in the jet is 1000 ppb. The data presented here is single shot data with no averaging. Signal to noise ratios calculated using this data are shown associated with each trace. The SNR improves by an order of magnitude using the gated channeltron.

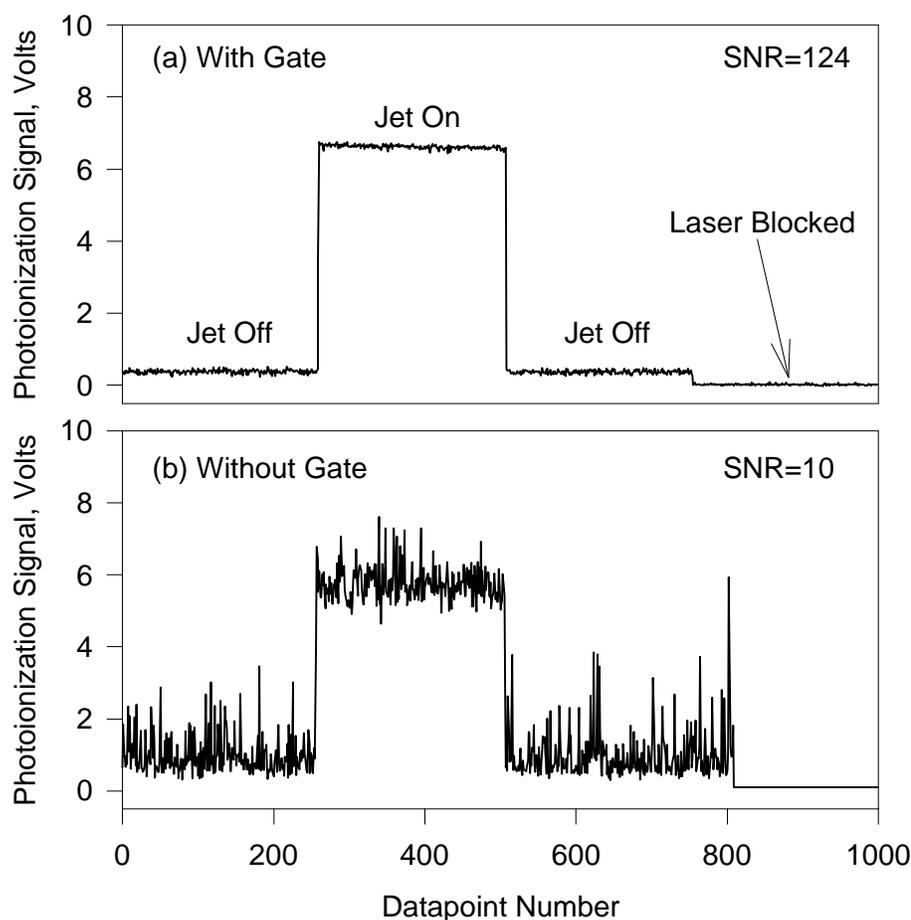


Figure 5. Data from 1000 ppb toluene in helium used for signal-to-noise (SNR) measurements (a) with and (b) without the channeltron gating circuitry. Data points are not averaged.

Figure 6 shows data obtained for determining the toluene detection limit. The toluene concentration is again 53 ppb in helium and each data point is an average of 100 shots. The slight overshoot in signal when the gas pulse is turned on is thought to be due to release of toluene accumulated on the interior of the pulsed valve assembly when the valve is off. In this case a SNR of 1500 is achieved which implies a detection limit (SNR=1) of 0.035 ppb. This can be compared to the detection limit of 0.05 ppb measured by T.A. Cool's research group in their resonance-enhanced multiphoton time-of-flight mass spectrometry apparatus.⁶ Similar low detection limits have been obtained for other organic compounds with the gated channeltron detector.

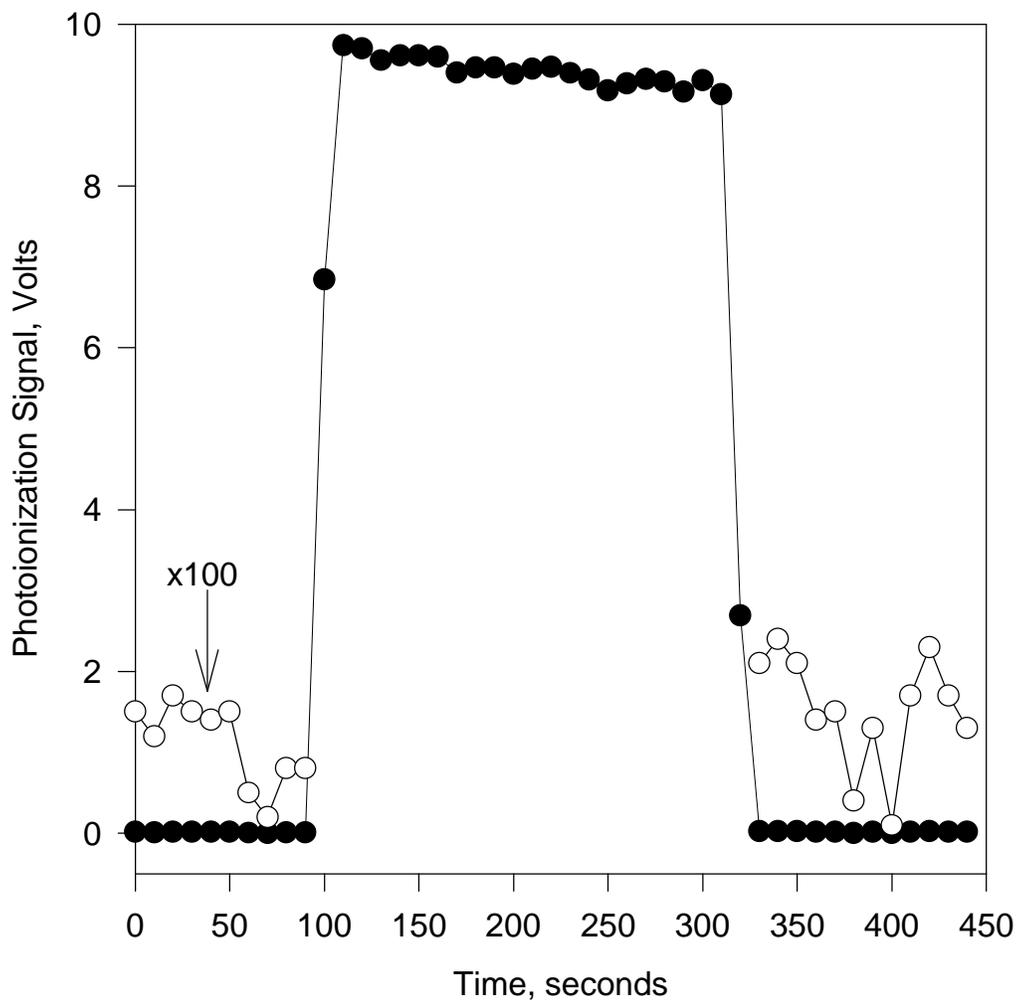


Figure 6. Data from 1000 ppb toluene in helium used to determine the toluene detection limit. Each data point is an average of 100 laser pulses. The baseline is shown vertically expanded by a factor of 100.

Summary

The above experimental results demonstrate the effectiveness of using a gated channeltron for resonant multiphoton ionization detection of trace organic molecules in a pulsed molecular beam apparatus. By using a gated channeltron, improvements in signal to noise ratios of at least one order of magnitude are easily obtainable. Furthermore, as demonstrated, the gated channeltron detector can be used for very low concentration detection of volatile organic compounds. The detection limits are as good as those obtained from other detection techniques such as resonance-enhanced multiphoton time-of-flight mass spectrometry. In addition, the use of a channeltron with a relatively simple gating circuitry enables a much simpler and more compact detection apparatus for the detection of VOCs.

Pattern Recognition

Background

There is much current interest in the use of pattern recognition (PR) techniques to identify and quantify chemical analytes based on the multivariate responses of chemical microsensor arrays^{7,8}. PR techniques used previously measured array responses of the class categories of interest, called training set data, to infer the class category associated with new array measurements (test data). The class categories can be the chemical identities of the analytes and can also represent the analyte concentrations. Each sensor in the array provides one dimension of the multidimensional data vectors. The application of PR techniques to microsensor-based chemical sensing is motivated by two advantages: (i) chemical identification is possible using only partially chemically selective microsensors rather than highly selective (and more difficult to develop) microsensors; (ii) PR can in principle identify a large number of chemical species using a small, fixed set of sensors.

The desired arrays of sensors are those that yield distinctive response patterns for different chemicals, i.e., so that the clusters of pattern vectors associated with different chemical classes are spatially separated and distinguishable in the pattern vector space. However, not all arrays yield useful separations of the chemical classes, so much of the work required to develop a useful chemical pattern recognition scheme is in finding a useful combination of chemical sensors.

SAW devices that have been coated with different partially selective films are of current interest for array applications.⁸⁻¹⁰ The variety of coatings that can be fabricated suggests the possibility of detecting a wide range of analytes with SAW devices. However, coated SAW devices often respond nonlinearly and occasionally nonmonotonically as a function of analyte concentration.¹¹ These devices also do not necessarily yield additive responses to chemical mixtures. Such responses present severe difficulties for conventional chemometric techniques, which work well for linear response signals.

We developed a new PR technique which can treat nonlinear and nonmonotonic sensor responses. The technique is computationally efficient, provides warning for new data points that are atypical or that cannot be reliably identified, and requires no user-adjustable parameters. The usefulness of the technique was demonstrated by using it to optimize arrays of coated SAW

devices for identifying individual VOCs, organophosphonates and water. The usefulness of PR techniques was further demonstrated by the ability to detect, using one of the best arrays found from the optimization study, water, six VOCs (acetone, DIMP, *i*-octane, *n*-propanol, TCE, toluene) and their 21 binary combinations over wide concentration ranges. VOC detection and monitoring is of interest for a number of industrial and DOE applications¹⁰ and organophosphonate detection is important for early warning of chemical warfare agents⁹.

Pattern Recognition Technique

Our PR technique relies on a new empirical approach to *k*-dimensional cluster analysis that incorporates measured human visual perceptions of difficult 2-dimensional clusters.¹² We use this new clustering method because it: (i) mimics human cluster perception; (ii) requires no prior knowledge about the final cluster result (e.g., the number of clusters); (iii) outperforms commercially available cluster methods on benchmarking tests.¹² Here we describe the use of the clustering results to carry out PR or to evaluate the expected PR performance using sets of coated SAW sensor responses. Details of the clustering method are described in Ref. 12.

PR is accomplished by computing clusterings of a new array (multivariate) measurement with the set of training measurements. These training measurements characterize the response of the array to the chemicals of interest, and the different sets of VOC categories are referred to as classes. The chemical identities of the training points, which are found to cluster to the new measurement, are used to assign a chemical identity as follows. Three possible clustering cases occur for the chemical class assignment of each new test point: (i) The new point is clustered only with training data of a single class and is assigned that same class -- this is the desired result; (ii) The new point is clustered with none of the training data and is not assigned a class -- these outlier points can result from detection of a new chemical class that is not present in the training data set, although sensor measurement errors can also produce outliers; (iii) The new point is clustered with training points from multiple classes -- these points occur where the training set classes overlap and cannot be reliably distinguished, so that the class assignment is ambiguous. We note that cases (ii) and (iii) provide useful information that is often unavailable or unreliable from existing PR techniques.

It is usually necessary to estimate the expected performance of the sensor array using only the measured training data set. We use the "leave-one-out" approach¹³, where each data point is individually removed from the complete training set and examined as a new test point by the PR technique. This approach maximizes the use of the available training data and avoids the favorable bias of the results that occurs by including the point that is being classified. This approach is not always convenient in some PR methods that require additional computations for each version of the training set with a point removed (e.g., retraining a neural net). The known class identity of each training data point is compared with the result inferred from the clustering approach. There are again three possible cases which can arise: (i) The point is clustered only with other training data of the correct class -- such points are correctly identified; (ii) The point is clustered with none of the other training data -- these outlier points are too far away from other points in the same class for clustering to occur. Outlier points in the leave-one-out analysis are usually due to inadequate sampling of the class in the parts of pattern vector space near the outliers, and suggest the need for more training points. However, these can also result from sensor measurement errors; (iii) The point is clustered with other training points from one or more

classes that do not match the correct class. These points identify overlaps between classes in pattern vector space where the sensor array is unable to reliably distinguish chemical identities. Statistics on these three cases for the entire training set provide three separate figures of merit for the sensor array under consideration, i.e. the percentage correctly identified, the percentage that indicates inadequate training set sampling, and the percentage that indicates class overlap. These figures of merit are used here to optimize the selection of the array sensors.

Experimental Procedures for Developing Training and Test Data Sets

We analyze experimental results from three sets of coated-SAW arrays and analytes. In all studies we make use of only the frequency shifts which result from the uptake of the analyte by the SAW device coatings. Other measurable responses, e.g. attenuation¹⁰, are also available for PR, but these were not considered in this work.

Chemical Set 1: The first set of chemicals includes water, an alcohol (isopropanol), a ketone (acetone), unchlorinated hydrocarbons (d-limonene, n-hexane, and toluene) and chlorinated hydrocarbons (CCl₄, chloroform, and trichloroethylene or TCE). These analytes were sensed using SAW devices coated with ethyl cellulose (EC), polyisobutylene (PIB) and nafion polymer films. The data were acquired by alternating between a chemical-free purge stream and a stream at a constant chemical concentration. Chemical concentration was controlled by the relative flow rates of a chemical-free mix-down stream and a saturated stream (obtained by passage through a bubbler containing the chemical). Four concentrations per chemical were obtained. This produced 12 data points for both the chlorinated and nonchlorinated classes.

Chemical Set 2: The second set of chemicals includes water, an alcohol (isopropanol), a ketone (acetone and methyl ethyl ketone or MEK), unchlorinated hydrocarbons (n-hexane, and toluene) and chlorinated hydrocarbons (CCl₄, chloroform, TCE, and perchloroethylene or PCE). These analytes were sensed using SAW devices coated with 13 polymer films: EC, PIB, poly(vinyl acetate) (PVA), poly(vinyl propionate) (PVP), a fluorinated polysiloxane (OV-215), a cyanide modified polysiloxane (OV-275), polyethylene glycol adipate (PEGA), poly(diphenoxy phosphazene) (PDPP), poly(epichlorohydrine) (PECH), poly(ethylenimine) (PEI), a hydrophobic fluoropolyol (FOX), poly(vinyl tetradecanal) (PVT), and a hydrogen bond acid modified polysiloxane (HBAS). The data were acquired as with Chemical Set 1. Ten concentrations in a linear progression from 5 to 50% of the saturation vapor pressure were obtained for each chemical.

Chemical Set 3: The third set of chemicals includes water, organophosphonates (diisopropyl methyl phosphonate or DIMP, dimethyl methyl phosphonate or DMMP), a ketone (acetone), aromatic hydrocarbons (benzene, toluene), chlorinated hydrocarbons (CCl₄, TCE), aliphatic hydrocarbons (cyclohexane, isooctane), and alcohols (methanol, pinacolyl alcohol, n-propanol). These analytes were sensed using SAW devices coated with nine self-assembled monolayer (SAM) films and plasma-grafted polymer films (PGFs). These films and the data acquisition method are described in Ref. 14. We retained 21 points per chemical, with concentrations ranging from 9% to 49% of the saturation vapor pressure.

Chemical Set 4: The last set consists of water, six VOCs (acetone, DIMP, *iso*-octane, *n*-propanol, TCE, toluene) and their 21 binary combinations over wide concentration ranges. This

set was done to evaluate the best set of SAW coatings obtained from the optimization study using the second set of individual VOCs. In particular, this study demonstrates that SAW arrays chosen for their ability to distinguish individual VOCs provide useful performance on binary mixtures as well. We also used this best 6-film array to investigate the ability of this approach to treat binary mixtures as a function of SAW sensitivity aging. The films were $\text{SH}(\text{CH}_2)_{15}\text{CH}_3$, $\text{SH}(\text{CH}_2)_{10}\text{COOH}$, $\text{SH}(\text{CH}_2)_{10}\text{COO}^-/\text{Cu}^{2+}$, plasma-grafted acrylic acid: 5 min. base + 30 min. graft, polyisobutylene - plasma-grafted acrylic acid: 5 min. base + 15 min. graft, and Eugenol-30. Binary VOC mixture data were acquired, starting from a high concentration/ low concentration mixture, by increasing the concentration of the low component while decreasing the concentration of the high component, ending when the opposite low/high mixture limit is reached.

Data Preprocessing for Clustering-Based Pattern Recognition

It is common for coated SAW devices to yield ranges of responses that are quite different in magnitude. In order for each device to have a comparable effect on the class separations in pattern vector space, the magnitudes of the device responses must be comparable. We equalize the responses of each sensor by a multiplicative scale factor, so that the largest absolute training data response of each sensor is unity.

The coated SAW sensor responses and the lengths of the multivariate pattern vectors generally increase as the chemical concentrations are increased. This tends to spread the class vectors away from the origin and each other as the chemical concentrations are increased. We have normalized the pattern vectors to unit length after the device responses have been equalized for the first two data sets.

Determination of Optimal Arrays

A key element in the successful development of a pattern recognition system is the selection of a useful set of measurements that can separate the classes in pattern vector space. This corresponds to selecting a set of SAW coatings and a set of chemical concentrations to include in the training set. One might expect that arrays with larger numbers of distinct sensor signals, k , would always be preferable for PR applications. This is not true for several reasons. From a practical viewpoint, arrays with the fewest devices are likely to require simpler fabrication and data acquisition procedures. This motivates the use of the smallest arrays that provide the desired performance. From a PR performance viewpoint, discriminating ability reaches a maximum and then declines as k is increased beyond approximately $N/2$, where N is the number of training data points per class. The reasons for this well-known behavior of PR systems are discussed in Ref. 13. Here we emphasize that the number of SAW films that can be usefully included in an array is limited by the number of chemical concentrations per class that is included in the training data set. This well-known condition for reliable PR analysis of arbitrary k -dimensional data is often not satisfied in published sensor PR work. Lastly, we have determined that using more SAW devices than the minimal necessary to distinguish the chemical of interest can lead to more rapid degradation of chemical recognition performance as the SAW devices age. Including unnecessary devices in an array can decrease the effective multivariate signal to noise of the system.

For large sets of distinct devices, we must search for the smallest subsets that produce good separation of the chemical classes of interest in a leave-one-out analysis. We do this explicitly for the nine SAMs and PGFs by examining all combinatorial sets of arrays. The computations are done on a SPARC 10 workstation. The primary figure of merit we seek to minimize is the percentage of incorrect leave-one-out points in the data set. We also seek to minimize the number of outliers among the array combinations with the smallest incorrect percentages. The number of data points per chemical class is sufficient to usefully include all of the SAM-coated and PGF-coated SAW devices in the array, but we are interested in identifying the smallest subset of devices with the best PR performance.

Evaluation of the Expected Robustness of Array Performance

The SAW sensors have been observed to alter their sensitivities as a function of age, and these changes can affect the ability of the PR system to distinguish the VOCs. The PR method can provide an estimate of this degradation if the quantitative effects of the sensor sensitivity changes are added to a training data set. We have performed this analysis for individual and binary mixture VOC data sets. Aging effects were simulated by artificially multiplying the responses of each film by a constant less than one. These artificial sensitivity reductions were varied randomly between no shift (constant=1) and a limiting value. The PR results directly show how much aging can be tolerated before the individual and binary mixture VOCs recognition becomes unreliable.

Results

Chemical Set 1: The leave-one-out results for the ethyl cellulose, polyisobutylene and nafion coatings yield no incorrect points for the class combinations chosen. These results show that a variety of chemical classes can be distinguished using these three SAW coatings. The three individual chemicals within the chlorinated and nonchlorinated VOC classes are not reliably separated, so that additional measurements are needed to distinguish these chemicals. However, more concentration values are required for the individual chemical classes to achieve this improvement.

Chemical Set 2: With these films, it was found that very high correct classifications (>95%) could be obtained for a large number of potential arrays. In an investigation into the effects of (modeled) aging, high performance was maintained with 20% added error, indicating these arrays are fairly robust. With the addition of 50% aging, significant degradation in performance was observed with the best array giving only 73% correct classification. With this level of aging, it was clear that the best arrays were of intermediate size, with only 4-6 sensors. The best four sensor array (EC, PIB, PEGA, and PVA) gave 72.3% correct. For five sensor arrays, adding OV-275 to this set gave 72.7% correct while adding PEI gave 72.9%. For six sensor arrays, adding both OV-275 and PEI gave the best result of 73.1%. It is clear, however, that the added complexity of going from four to six sensors would not be justified by these small improvements if these large aging effects were expected.

Chemical Set 3: The results of the array optimization showed that all three-film arrays generate some incorrect chemical identifications. The overall PR performance improves as films are added and, for four or more films in the arrays, arrays exhibiting 100% correct responses were found for all array sizes. The best film combinations were found to be: three films -- HS(CH₂)₁₀COO⁻/Cu²⁺, Eugenol-30, PIB-PGAA 5+15, four films -- HS(CH)₁₅CH₃, PGAA-PGAA 5+15, Eugenol-30, PIB-PGAA 1+15, five films -- HS(CH₂)₁₅CH₃, HS(CH)₁₅CH₃, PGAA-PGAA 5+15, PGAA-PGAA 5+30, PIB-PGAA 1+15, and six films -- HS(CH₂)₁₀COOH, HS(CH₂)₁₀COO⁻/Cu²⁺, HS(CH)₁₅CH₃, Eugenol-30, PIB-PGAA 1+15, PIB-PGAA 5+15.

Chemical Set 4: The results from this best 6-film array show that good chemical differentiation of both individual and binary mixtures is achieved by applying PR to this set of six coated SAW films. Similar to the results for Chemical Set 2, the effects of (modeled) aging are

small for distinguishing of individual VOCs for SAW sensitivity losses of up to 30%; However, for distinguishing binary VOCs, identification accuracy becomes increasingly unreliable for SAW sensitivity losses beyond about 10%.

Summary

We have developed a new pattern recognition (PR) technique for chemical identification using arrays of microsensors. The technique relies on a new empirical approach to k-dimensional cluster analysis which incorporates measured human visual perceptions of difficult 2-dimensional clusters. The method can handle nonlinear SAW array data, detects both unexpected (outlier) and unreliable array responses, and has no user-adjustable parameters. We use this technique to guide the development of arrays of thin-film-coated SAW devices that produce optimal PR performance for distinguishing a variety of volatile organic compounds, organophosphonates and water. The technique also can be used to indicate the expected losses in chemical recognition performance as the sensors exhibit age-induced changes in sensitivities.

Conclusion

We have shown that a pulsed molecular beam expansion combined with the use of a gated channeltron detector provides a very sensitive means to detect VOCs. Detection limits in the ppt range were obtained for toluene. Detection limits for other VOCs such as acetone, xylenes, and trichloroethylene were in the ppt to ppb range. Based on the work in this report, it should be possible to develop field-portable VOC sensors with sensitivity in the high ppt to low ppb range. The high spectral resolution associated with pulsed beam expansion measurements enhances the identification of VOCs in complex chemical environments. Further improvements in the identification of VOCs are possible with the use of pattern recognition algorithms such as the VERI algorithm developed to analyze signals from arrays of SAW devices.

We have also shown that good chemical differentiation can be achieved in arrays of SAW devices by using pattern recognition algorithms to deconvolve signals generated by the individual coated SAW films, if the choice of films is optimized. Furthermore, we have shown that performance in differentiating chemical species is maintained for modest drops in film sensitivities that occur due to SAW film aging. The choice of array size is important here as we have found that the performance of larger than optimal arrays can degrade more rapidly with SAW film aging. The results further emphasized the crucial role of PR analyses in designing practical array systems for field operation. Our work in this report conclusively demonstrates that pattern recognition algorithms can be used to improve the performance of SAW devices in identifying VOCs and that they play a crucial role in the design of SAW sensors that operate reliably under field conditions.

References

1. R.J. Baker and B.P. Johnson, *Electron. Lett.* **29**, 56 (1993).
2. R.J. Baker and B.P. Johnson, *Rev. Sci. Instrum.* **64**, 1655 (1993).
3. M. Gostein and G.O. Sitz, *Rev. Sci. Instrum.* **65**, 3036 (1994).
4. M. Gostein and G.O. Sitz (private communication).
5. P.J. Breen, J.A. Warren, and E.R. Berstein, *J. Chem. Phys.* **87**, 1917 (1987).
6. T.N Tanada, J. Velazquez, N. Hemmi, and T.A. Cool, *Ber. Bunsenges. Phys. Chem.* **97**, 1516 (1993).
7. X. Wang, J. Fang, P. Carey, S. Yee, *Sensors and Actuators* **B13**, 455 (1993).
8. S. L. Rose-Pehrsson, et. al., *Anal. Chem* **60**, 2801 (1988).
9. A. J. Ricco, C. Xu, R. M. Crooks, and R. E. Allred in *Interfacial Design and Chemical Sensing*, ACS Symposium Series No. 561 (Amer. Chem. Society, Washington, D. C., 1994), Ch. 23, p. 264.
10. G. C. Frye, S. J. Martin R. W. Chernosek and K. B. Pfeifer, *Int. J. of Environ. Conscious Manufacturing* **1**, 37 (1992).
11. G. C. Frye and S. J. Martin, "Velocity and Attenuation Effects in Acoustic Wave Chemical Sensors", *Proc. 1993 Ultrasonics Symposium (IEEE, Piscataway, NJ, 1993)* p.379.
12. G. C. Osbourn and R. F. Martinez, *Pattern Recognition* **28**(11), 1793 (1995).
13. D. J. Hand, *Discrimination and Classification* (John Wiley and Sons, New York, 1981).
14. A. J. Ricco, G. C. Osbourn, J. W. Bartholomew, R. M. Crooks, C. Xu, and R. E. Allred, *Technical Digest 1994 Solid-State Sensor and Actuator Workshop*, (Transducer Research Foundation, Cleveland, 1994), p. 180.

DISTRIBUTION:

5	MS1423	P.J. Hargis, Jr., 1128
1	MS1423	G.N. Hays, 1128
1	MS1423	B.P. Aragon, 1128
1	MS1423	F.R. Franklin, 1128
5	MS1423	G.C. Osbourn, 1155
5	MS1425	G.C. Frye, 1315
5	MS1425	A.J. Ricco, 1315
1	MS0971	A.R. Lang, 5703
1	MS0972	A.J. Medina, 5722
1	MS0755	W. Einfeld, 6612
1	MS0755	D.S. Blair, 6612
1	MS1188	I.R. Shokair, 9512
2	MS0188	Laboratory-Directed Research/Development, 4523
1	MS9018	Central Technical Files, 8940-2
5	MS0899	Technical Library, 4414
2	MS0619	Review & Approval Desk, 12609 For DOE/OSTI