SANDIA REPORT

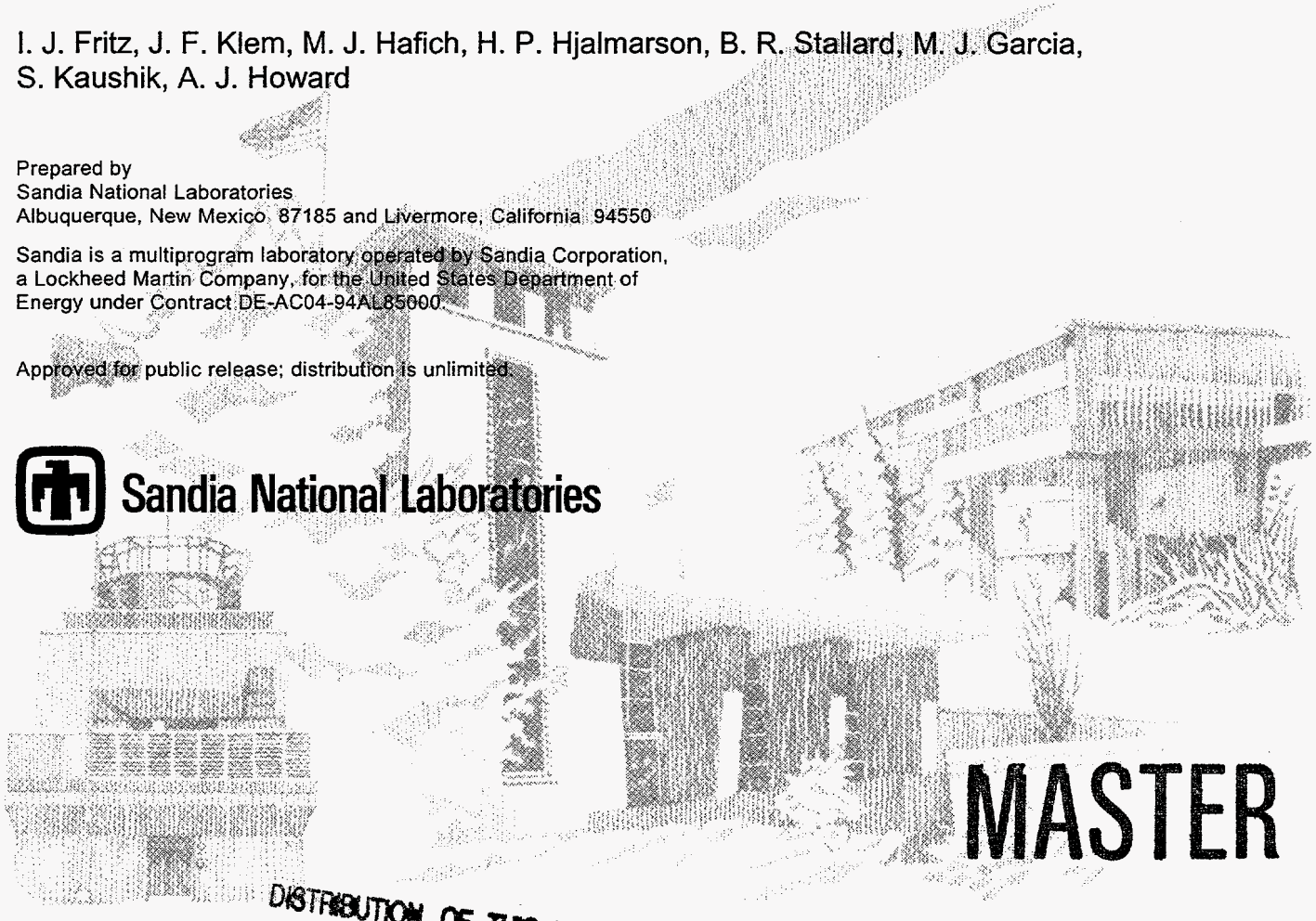
SAND97–0369 • UC–410
Unlimited Release
Printed April 1997

Compact Environmental Spectroscopy Using Advanced Semiconductor Light-Emitting Diodes and Lasers

I. J. Fritz, J. F. Klem, M. J. Hafich, H. P. Hjalmarson, B. R. Stallard, M. J. Garcia, S. Kaushik, A. J. Howard

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.



Sandia National Laboratories

SF2900Q(8-81)

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

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

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

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

NTIS price codes
Printed copy: A04
Microfiche copy: A01

SAND 97-0369
Unlimited Release
Printed April 1997

Distribution
Category UC-410

Compact Environmental Spectroscopy Using Advanced Semiconductor Light-Emitting Diodes and Lasers

I. J. Fritz
Photonics Research Department
J. F. Klem and M. J. Hafich
Compound Semiconductor Materials and Processes Department
H. P. Hjalmarson
Advanced Compound Semiconductor Technologies Department
B. R. Stallard, M. J. Garcia and S. Kaushik
Surface and Sensor-Controlled Processes Department

Sandia National Laboratories
P. O. Box 5800
Albuquerque, NM 87185-0603

A. J. Howard
Intel Corporation
Rio Rancho, NM

Abstract

This report summarizes research completed under a Laboratory Directed Research and Development program funded for part of FY94, FY95 and FY96. The main goals were 1) to develop novel, compound-semiconductor based optical sources to enable field-based detection of environmentally important chemical species using miniaturized, low-power, rugged, moderate cost spectroscopic equipment, and 2) to demonstrate the utility of near-infrared spectroscopy to quantitatively measure contaminants. Potential applications would include monitoring process and effluent streams for volatile organic compound detection and sensing head-space gasses in storage vessels for waste management. Sensing is based on absorption in the 1.3-1.9 μm band from overtones of the C-H, N-H and O-H stretch resonances. We describe work in developing novel broadband light-emitting diodes emitting over the entire 1.4-1.9 μm wavelength range, first using InGaAs quantum wells, and second using a novel technique for growing digital-alloy materials in the InAlGaAs material system. Next we demonstrate the utility of near-infrared spectroscopy for quantitatively determining contamination of soil by motor oil. Finally we discuss the separability of different classes of organic compounds using near-infrared spectroscopic techniques.

Contents

Introduction

This report summarizes research completed under a Laboratory Directed Research and Development program funded for part of FY94, FY95 and FY96. It includes material from five journal publications as listed in the Table of Contents.

There were two main goals to this program: 1) to develop novel compound-semiconductor based optical sources to enable field-based detection of environmentally important chemical species using miniaturized, low-power, rugged, moderate cost spectroscopic equipment, and 2) to demonstrate the utility of near-infrared spectroscopy to quantitatively measure contaminants. Potential application areas include a wide variety of important environmental problems such as monitoring process and effluent streams in manufacturing for volatile organic compounds and sensing head-space gasses in storage vessels for waste management. In addition, miniature spectroscopic sensors would find utility in areas such as monitoring the health of nuclear weapons, at-line process monitoring in the chemical and pharmaceutical industries, real-time non-invasive monitoring of medical patients and quality control in the food, textile and other industries.

For this program, we investigated sensing based on optical absorption in the 1.4-1.9 μm band from overtones of the C-H, N-H and O-H stretch resonances. In this report we describe work in developing novel broadband light-emitting diodes emitting over the entire 1.4-1.9 μm wavelength range. This work utilized and built on Sandia's extensive expertise in strained-layer compound-semiconductor technology. Our first devices utilized three InGaAs quantum wells with different compositions and demonstrated, for the first time, the feasibility of obtaining broadband emission in the desired 1.3-1.9 μm range. To improve the uniformity of the light output with respect to wavelength, we next developed a novel technique for growing digital-alloy materials in the InAlGaAs material system and then incorporated these materials into broadband emitter structures. This work on broadband light-emitting diodes is described in the following three sections.

In the last two sections of this report we address the question of the suitability of near-infrared spectroscopy for detection of chemical species important for environmental monitoring. As a specific example, we demonstrate the utility of near-infrared spectroscopy for quantitatively determining contamination of soil by motor oil, and propose a design for a miniature spectrometer for this application based on our broadband light-emitting diodes. Finally we use advanced statistical techniques to investigate the separability of three different classes of organic compounds using infrared spectroscopy in different wavelength regimes and show that there is sufficient class separability in the near-infrared to ensure delectability of a given class in the presence of interference from substances in the other classes.

Broadband light-emitting diode for 1.4-2.0 μm using variable-composition InGaAs quantum wells

I. J. Fritz, J. F. Klem, M. J. Hafich, A. J. Howard and H. P. Hjalmarson

Sandia National Laboratories, P. O. Box 5800, Albuquerque, NM 87185-0603

ABSTRACT

We describe a novel broadband light-emitting diode employing InGaAs quantum wells with different bandgaps. The device structures were grown by molecular-beam epitaxy on InP substrates and consist of three $In_xGa_{1-x}As$ quantum wells (x = 0.4, 0.53 and 0.66) imbedded in the p-doped region of an InAlAs p-n junction diode. Electrons injected into the p-region are captured into the wells and recombine with holes to produce radiation at the three bandgap wavelengths. Broadband emission over a large wavelength range (1.4 - 2.0 μm) is easily achieved with this approach. These novel solid-state sources have a variety of potential applications, for example in miniature spectrometer and sensor systems.

Broadband light-emitting diodes (BBLEDs) are novel multiple-quantum-well photonic sources with a range of interesting potential applications in systems such as miniature spectrometers[1], fiber-optic gyroscopes[2] and sensors for the automotive and avionics industries.[3] Our motivation for developing BBLEDs in the near-infrared stems from an evolving interest in customized miniature spectrometers for environmental sensing[1] (e.g. process and effluent monitoring) and biomedical applications[4] (e.g. blood sugar and alcohol monitoring). For our initial work we have targeted the 1.3 μm - 2.0 μm wavelength region where absorption from N-H and C-H stretching modes provides important signature information for detecting organic compounds.

To obtain optical emission over the desired spectral region, we have chosen to work with InGaAs/InAlAs multiple-quantum-well structures grown on InP substrates. It is possible to grow several quantum wells with different bandgaps in a single device structure, with easily accommodated strain levels. Overlapping emission from the individual quantum wells then provides the broadband source. Previous work using multiple-bandgap quantum wells[3] was done with lattice-matched AlGaAs materials, which restricts available wavelengths to ~700-900 nm. Our work represents the first use of multiple strained quantum wells with different bandgaps to achieve broadband emission at wavelengths beyond 1.3 μm. Other work on BBLEDs operating near 1.5 μm has emphasized the use of superluminescence from single-composition quantum wells.[2] This approach has yielded spectral widths up to 200 nm, about a factor of three smaller than achieved in the present work.

Our devices were grown by molecular-beam epitaxy (MBE) and are illustrated schematically in Fig. 1. Three $In_xGa_{1-x}As$ quantum wells, with x=0.4, x=0.53 and x=0.66, emit at their individual bandgap wavelengths of λ_1, λ_2 and λ_3. The three wells are each 10 nm thick and are under planar strains of -0.9% (x=0.4), 0 (x=0.53) and +0.9% (x=0.66), well within the limits for coherent epitaxy. Only single In, Al and Ga effusion cells are available in our growth chamber, limiting the range of compositions that can be grown in a single structure. Since growth interruptions to grow different InAlGaAs alloys for the barriers and wells is not feasible (due to interfacial contamination), we chose to use lattice-matched $In_{0.52}Al_{0.48}As$ for the barrier and cladding layers, without separate confinement regions, for our initial studies. The three InGaAs compositions were produced by linearly varying the temperature of the Ga cell during the growth of the triple-quantum-well region of the structure. The choice of using a tensilely strained well for the short wavelength emission might be expected to lead to a reduced emission from this well; however, its strength relative to the other wells also depends on the transport and capture properties of the structure, so that the overall emission spectrum may be tailored by adjusting other parameters, such as doping and barrier width. The quantum wells are separated by 6-nm-thick barriers of lattice-matched $In_{0.52}Al_{0.48}As$.

As indicated in Fig. 1, the quantum wells are located in the p-doped region of an InAlAs p-n junction diode. This configuration was chosen to minimize difficulties in transporting holes through the barriers, which we believe to be the mechanism that most strongly limits injection of electron-hole pairs uniformly through the quantum-well region. The dopants are Si ($n=2\times10^{18}$ cm^{-3}) and Be ($p=(1-2)\times10^{18} cm^{-3}$). A separation of d_j=24 nm between the junction and the first well is included to avoid depletion of holes from the wells. The growth sequence was from left to right in Fig. 1. The uppermost InAlAs layer is 0.4 μm thick and is capped with a 2-

nm-thick $In_{0.53}Ga_{0.47}As$ contacting layer doped at $p=2x10^{18}$ cm^{-3}. Standard photolithography and wet chemical etching were used to define square mesa-isolated devices between 0.4 and 1.5 mm on a side. The n-type substrates were contacted with Ge/Au/Ni/Au and the tops of the mesas were contacted with rings of Ti/Pt/Au. The device wafer was cleaved along the mesa edges to allow observation of emission from the side of the devices as well as from the top.

The expected emission wavelengths of the three wells were calculated by two methods. First, three (separate) superlattice band structure calculations[5] were done using the three individual well-barrier combinations. Second, a tunneling-resonance calculation of the actual triple-well structure was made. In both calculations band-bending effects due to the doping were ignored, as were the compositional variations within each well. The calculations were done in the two-band envelope-function formalism and yielded identical results. The heavy-hole (HH) bandgaps for the x=0.53 and x=0.66 wells are predicted to be at λ_2=1550 nm and λ_3=1765 nm, respectively. As the x=0.4 well is in tension, its bandgap involves the light hole (LH) and is predicted to occur at λ_1=1445 nm. The HH transition for x=0.4 is expected at 1380 nm.

Spectrally resolved electroluminescence data were obtained as follows. The device was mounted on a rotation stage to allow collection from different emission angles. The emission was collected at an aperture of f/3.5, chopped, dispersed by a 0.25 m monochromator and detected by a liquid-nitrogen cooled InAs detector feeding a lock-in amplifier. Similar spectra were taken from a calibrated tungsten-halogen lamp to correct for the wavelength-dependent response of the measurement system. Absolute emission intensity at the wavelength of peak LED output was measured by comparing the LED and standard lamp outputs in a small-aperture geometry where the absolute intensity of the standard is known.

Typical emission spectra ("shape normalized" to correct for the relative system response) are shown in Fig. 2. The data are from a 0.023 cm^2-area device under dc excitation of 500 mA. Similar spectra were obtained on a smaller device ($4.9x10^{-3}$ cm^2 in area) at the same current level. Emission normal to the top surface of the device with no polarization analysis is given by the solid curve. The positions of four observed optical transitions are marked by vertical lines. The longest emission peak, ascribed to the HH transition of the x=0.66 well (well #3), is observed at 1820 nm. Emission from the HH transition of the x=0.53 well (well #2) is seen at 1600 nm. Both LH and HH transitions are observed for the x=0.4 well (well #1). In top emission these are seen as broad structure between 1400 and 1600 nm. The dashed and dotted curves of Fig. 2. show edge-emission spectra for light polarized parallel to the epilayers (TE) and normal to them (TM). To a good approximation, HH transitions are forbidden in the TM spectrum, allowing the assignment of the peak at 1520 nm to the LH transition of the x=0.4 well. The peak seen at shorter wavelength (1475 nm) in the TE spectrum is assigned to the HH transition of the x=0.4 well based on its position relative to the other transitions. The observed transitions are all at somewhat longer wavelengths than predicted by our calculations, but their relative positions agree well. The discrepancy between calculation and observation stems from small run-to-run variations in the growth conditions; transitions in other devices (not shown here) are in better agreement with theory.

The insert to Fig. 2 shows the light-current characteristic for top emission at the peak wavelength. The data are linear below 10 mA of current and show a gradual saturation at higher levels. The shapes of the spectra in Fig. 2 change very little with drive current, the main effect being some small line broadenings and shifts at the highest currents, probable due to heating.

To determine the total output of our device it is necessary to know in detail the angular and polarization dependence of the emission spectrum. To make a preliminary estimate of the integrated output, we have measured the angular dependence of the output at several wavelengths. Typical data are shown in Fig. 3. Figure 3a shows the angular dependence of emission from the top of the device, for TE and TM polarization, obtained at 1800 nm. The insert figure shows the expected angular dependences for isotropic radiation through a single interface.[6] The discrepancy in the shapes of the theoretical and experimental curves is due to multiple internal reflections and scattering from the unpolished rear surface. If we approximate the angular dependence as simply a cosine function, independent of wavelength, then the total top emission is given by the product of (a) the area under the solid curve in Fig. 2, (b) the angular integral of the cosine dependence and (c) the absolute peak intensity, measured as described above. The result is an integrated output of ~250 μW. Emission from the side of the device is more complicated because of optical interference (quasi-waveguiding) effects and because of the polarization-dependent selection rules. We have therefore not attempted to estimate the integrated output. However, we find the side emission to be comparable in strength to the top emission. The angular dependence of the side emission at 1489 nm and 1800 nm are shown in Fig. 3b. Due to the quasi-waveguiding effects, the spectra peak at angles of ~40-50 degrees away from the side, depending on wavelength.

The data of Fig. 2 show that significant optical output is obtained between ~1400-2000 nm. This 600 nm range is useful for a wide variety of spectroscopic applications.[7] There is, however, still opportunity for considerable improvement in device performance. The present data were obtained under dc operation. Pulsed operation would allow operation at higher power levels. Improvements in light collection (i.e., from a bonded lens) and in spatial uniformity of emission (by reducing current crowding) are also expected to significantly improve performance. As illustrated by our results, the bandgaps of the three quantum wells are not evenly spaced, even though their compositions are. This is a consequence of the built-in strain and leads to an undesirable asymmetry in the spectrum. Thus an improvement in device performance should be realized by utilizing a quantum-well structure with equally spaced HH transitions. Further enhancement in performance should be possible by including a "separate confinement" region to enhance carrier capture into the wells. The latter two improvements are not feasible with the present growth method. We are currently exploring the use of digital-alloy materials to provide more flexibility in tailoring bandgaps and strains for optimum BBLED devices.

ACKNOWLEDGMENTS

We are pleased to acknowledge expert technical assistance from D. B. Webb. This work was supported by the U. S. Department of Energy under Contract DE-AC04-94AL85000.

REFERENCES

1. X. Wang, D. E. Vaughan, V. Pelekhaty and J. Crisp, "A novel miniature spectrometer using an integrated acousto-optic tunable filter," *Rev. Sci. Instrum.*, vol. 65, pp. 3653-3656, 1994.

2. S. Kondo, H. Yasaka, Y. Noguchi, K. Magari, M. Sugo and O. Mikami, "Very wide spectrum multiquantum well superluminescent diode at 1.5 μm," *Electron. Lett.*, vol. 28, pp. 132-133, 1992.

3. A. J. Moseley, D. J. Robbins, C. Meaton, R. M. Ash, R. Nicklin, P. Bromley, R. R. Bradley, A. C. Carter, C. S. Hong and L. Figueroa, "Broadband GaAs/$Al_XGa_{1-X}As$ multi-quantum well LED," in *Quantum Optoelectronics, 1991*. Washington: Optical Society of America, 1991, Technical Digest Series vol. 7, pp.193-196.

4. D. M. Haaland, M. R. Robinson, G. W. Koepp, E. V. Thomas and R. P. Eaton, "Reagentless near-infrared determination of glucose in whole blood using multivariate calibration," vol. 46, pp. 1575-1578, 1992.

5. I. J. Fritz, J. F. Klem, T. M. Brennan, J. R. Wendt and T. E. Zipperian, "Superposition of light-hole and heavy-hole bandgaps in InGaAs/InAlAs strained quantum-well structures," *Supperlatt. and Microstr.*, vol. 10, pp. 99-105, 1991.

6. F. Stern, "Transmission of isotropic radiation across an interface between two dielectrics," *Appl. Optics*, vol. 3, pp. 111-113, 1964.

7. Note that vertical emission is not polarization dependent. The polarization dependence of the side emission does not pose any problem for most spectroscopic sensing applications (because gasses and liquids are isotropic) but might need to be considered in some system designs.

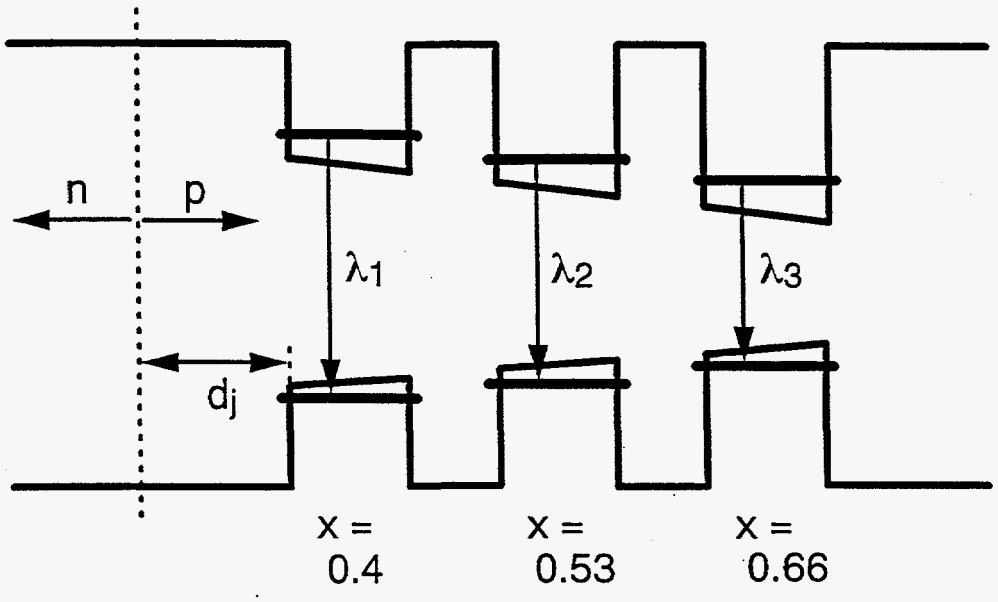


Fig. 1 Schematic of conduction and valence bands in a triple-quantum-well, broadband, light-emitting diode. InGaAs quantum wells with three different compositions emit light at three different wavelengths. The wells are in the p-doped region of an InAlAs junction diode grown on an InP substrate. Band bending due to doping is not included in this schematic.

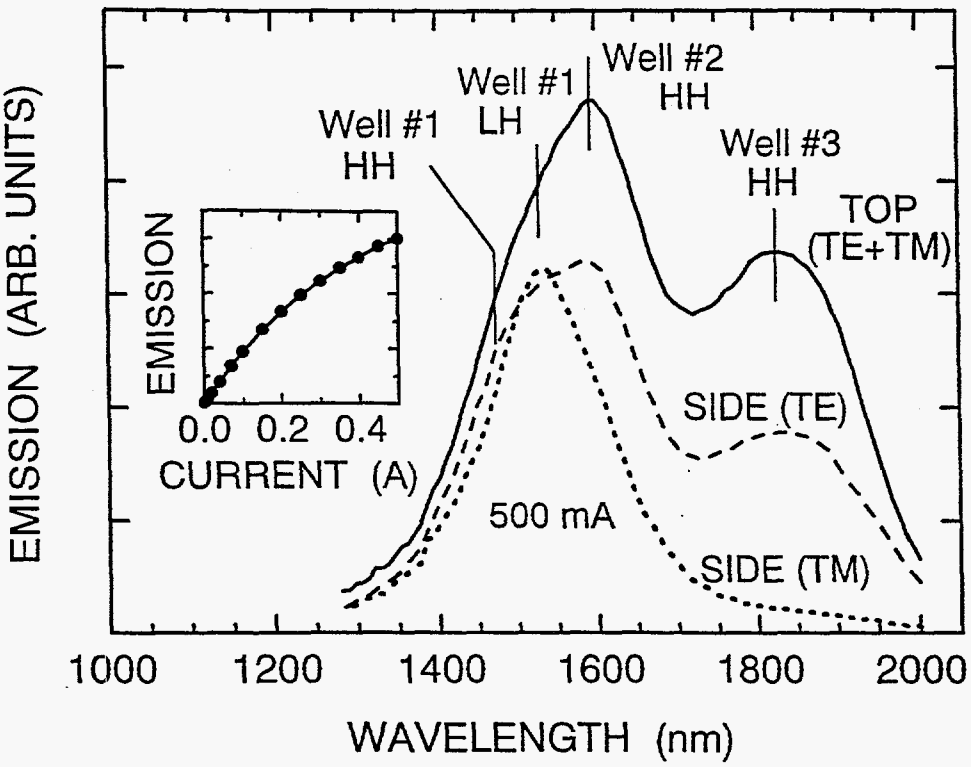


Fig. 2 Emission spectrum from a typical device with 500 mA of forward current. Solid curve: emission normal to top surface without analyzing polarizer. Dashed curve: Edge emission with TE polarization. Dotted curve: Edge emission with TM polarization. Short vertical lines indicate positions of various bandgap transitions. Insert: relative light emission vs. drive current at the peak wavelength.

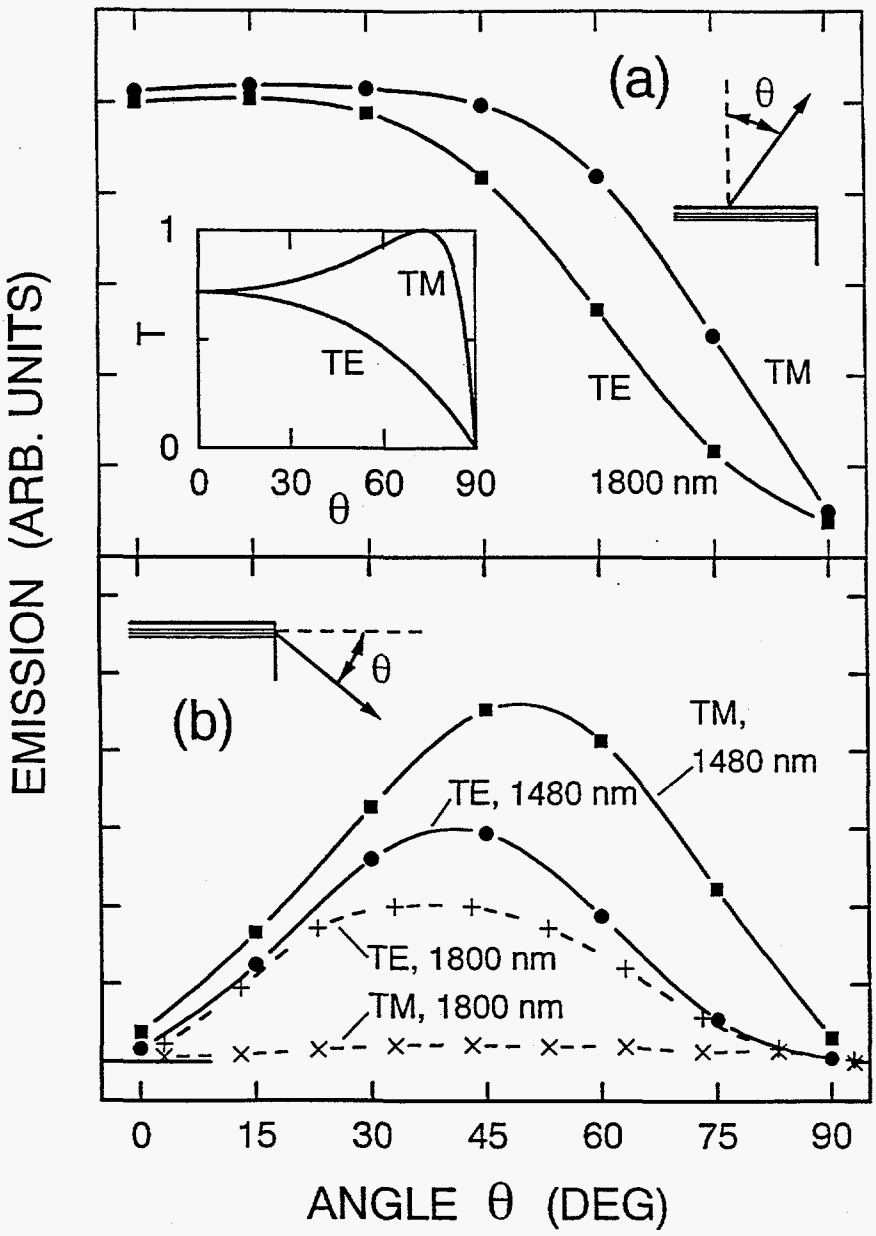


Fig. 3 Angular dependence of emission intensity. (a) top emission at 1800 nm. (b) edge emission at 1480 nm and 1800 nm. Insert: predicted top emission variation assuming isotropic radiation into a single interface. Discrepancy between experiment and prediction for top emission is due to multiple internal reflections and scattering. Peak emission away from θ=0 in (b) is due to quasi-waveguiding effect.

Novel pseudo-alloy approach to epitaxial growth of complex InGaAlAs multilayer structures

I. J. Fritz, J. F. Klem, M. Hafich and A. J. Howard

Sandia National Laboratories, Albuquerque, New Mexico, 87185-0603

ABSTRACT

We describe growth by molecular-beam epitaxy of InGaAlAs multilayers using a versatile and agile technique that allows a wide range of bandgaps and strains for the individual layers. In our approach, the layers are most generally pseudo-alloys produced by growing short-period superlattices containing ternary or quaternary layers, in various combinations, without changing the temperatures of the single In, Ga, Al and As effusion cells. To illustrate the method, we have designed, grown and characterized a separate-confinement, strained-layer light-emitting diode operating at 1.5 μm.

Epitaxial semiconductor growth capabilities continue to be challenged by the need for increasingly complex device structures. Devices requiring multiple materials with different bandgaps, refractive indices, strains, etc. form a long list, including such devices as graded-index separate-confinement heterostructure lasers, graded-bandgap heterojunction transistors, vertical-cavity lasers and modulators, quantum-cascade lasers and broadband light-emitting diodes. All of these devices have been demonstrated, but increased flexibility and agility in material growth is needed for continued improvement in their performance. In this Letter we consider a novel approach to the growth of materials with a wide range of bandgaps and strains, using the InGaAlAs materials system as an example.

To understand our approach, consider the use of a molecular-beam epitaxy (MBE) machine having single sources of In, Ga, Al and As available for growing InGaAlAs alloys. With the temperatures of the individual effusion cells fixed, this machine can grow three binary compounds (GaAs, InAs and AlAs), three ternary compounds (GaAlAs, AlInAs and InGaAs) and one quaternary compound (InGaAlAs). These seven compounds correspond to seven "states" of the machine, each state being specified by a different combination of shutter openings. By rapidly switching between pairs of machine states (i.e., growing short-period superlattices), one can grow pseudo-alloys having compositions intermediate to the compositions of the two individual states. More complicated pseudo-alloys are possible by switching among several machine states.

In the past there has been considerable work on using short period superlattices to approximate ternary alloys (e.g., GaAs/$Al_{0.3}Ga_{0.7}As$ for $Al_xGa_{1-x}As$[1], InAs/GaAs for $In_xGa_{1-x}As$[2], GaSb/InSb for $Ga_xIn_{1-x}Sb$[3], etc.). Our approach is an extension of that work to include various combinations of ternary and quaternary alloys for the individual layers, an area in which little work has been done.[4] Of course it is possible to grow any composition by switching among the three binary states, but in most instances this is not convenient because of the large lattice mismatch involved.

To demonstrate our machine-state switching approach to pseudo alloys, we have designed, grown and tested a separate-confinement, strained-layer light-emitting diode (LED) operating at a wavelength of 1.5 μm. This device was grown as a part of a program to develop broadband emitters over the 1.3-1.8 μm region using multiple quantum wells with different bandgaps. Results on multiple-quantum-well emitters will be reported separately.[5]

In Fig. 1 we illustrate schematically the machine-state setup for growing our LED. This figure shows bandgap vs. lattice-constant for the InGaAlAs material system. The machine states are labeled in vector notation, with (1,0,0), (0,1,0) and (0,0,1) representing InAs, GaAs and AlAs, respectively, and the appropriate sums of these vectors representing the ternary and quaternary states. For our demonstration device, we set the relative fluxes of the sources so that the (1,1,0) state represents the ternary alloy $In_{0.66}Ga_{0.34}As$ and (1,0,1) represents $In_{0.52}Al_{0.48}As$ (lattice matched to InP).

Since the relative fluxes of the In, Ga and Al sources are fixed once the compositions of the (1,1,0) and (1,0,1) states are set, the compositions corresponding to the (0,1,1) and (1,1,1) states are also fixed. Because the InGaAlAs materials are grown at a relatively low substrate temperature with a large As overpressure, it is a good approximation to assume that the growth rates are additive and determined by the group III flux rates. Thus the quaternary (1,1,1) state

corresponds to $In_{0.41}Ga_{0.21}Al_{0.38}As$. Note that this alloy has a smaller lattice constant than InP.

By machine-state switching among the three states (1,1,0), (1,0,1) and (1,1,1), one may now grow pseudo-alloys with all compositions bounded by the heavy line in Fig. 1. This range of compositions encompasses a wide range of bandgaps and a moderate range of lattice constants. Since the available lattice constants straddle that of InP, one may design strained-layer quantum-well structures using compressively strained wells and strain-compensating barriers with a variety of well depths and bandgaps.[6] In addition, a range of compositions matched to InP enables tailoring of the properties of thick, graded bandgap or graded index regions. Previous work on quantum-well structures in the InAlGaAs/InP material system has mainly involved single-composition, ternary InGaAs wells, with a much narrower range of available bandgaps.

For the active quantum-well region of our demonstration device, we use materials along the (1,1,0) - (1,1,1) line in Fig. 1. To grow these quaternary pseudo-alloys, the In and Ga shutters are left open, and the Al shutter is switched cyclically with a duty cycle q. We may denote the resultant effective machine state as **Q**=q(1,1,1)+(1-q)(1,1,0)=(1,1,q). In Fig. 2 we show the properties of the quaternary as a function of q assuming homogeneous mixing. Figure 1a shows the individual In, Ga and Al mole fractions. Figure 1b gives the bandgap wavelength, calculated using a standard interpolation procedure[7] ignoring strain and confinement effects. Part (c) of the figure shows the lattice mismatch between the alloy and InP.

Our device structure, shown in Fig. 3, consists of a 10-nm-thick q=0.18 quantum well, surrounded by separate confinement layers with q=0.72, each 9.3 nm thick. This strain-balanced structure is sandwiched between $In_{0.52}Al_{0.48}As$ cladding layers ~0.5 μm thick. The quantum-well structure is located in the p-doped region of a p-n diode, with the diode junction located 30 nm below the bottom separate-confinement layer. Within the quantum-well layer the individual layer thicknesses are 0.29 nm for the (1,1,1) composition and 1.35 nm for (1,1,0). For the separate-confinement layers these values are 1.17 nm and 0.47 nm, respectively. Wet chemical etching was used to isolate square devices between 0.2 mm and 1.8 mm on a side. For the results presented below, the device wafer was cleaved near the mesa edges to allow observation of edge emission. The devices had current-voltage characteristics similar to companion devices with bulk InGaAs quantum wells and InAlAs cladding.

In Fig. 3 we show electroluminescent emission spectra from the pseudo-alloy device described above along with data from a companion reference sample. The reference diode consists of a 10-nm-thick $In_{0.53}Ga_{0.47}As$ quantum well imbedded in the p-region of an $In_{0.52}Al_{0.48}As$ junction diode, as shown in the figure. The spectra were obtained by collecting the side emission at an f-number of ~3.5, dispersing the radiation with a 0.25 m grating monochromator and detecting it with a cooled InAs detector. The spectra are shape-normalized to spectra from a calibrated standard lamp. For each device we show the TE and TM polarized emission separately. To first order, emission from the heavy-hole bandgap transition is forbidden in TM polarization.

The main feature of the data is that the new pseudo-alloy device is about three times brighter than the reference device. We believe this is mainly due to the action of the separate-confinement layers, which make the carrier-capture width of the pseudo-alloy device about three times larger than that of the reference structure. The observed brightness therefore demonstrates that there is no significant device degradation from the use of many thin layers in the separate-

confinement and active regions. An additional feature of the data is the difference in the relative strengths of the TE and TM peaks for the two devices. This difference is due to the separate-confinement structure having a compressively strained quantum well. The strain increases the splitting of the light and heavy hole states, thereby decreasing the emission from the light hole.

The observed emission peaks of our device are at 1510 nm (heavy hole) and 1435 nm (light hole). To compare these with theoretical expectation we have performed tunneling resonance calculations[8] of the various energy levels in our structure. This calculation was done using two-band envelope-function boundary conditions,[9] with material parameters given previously.[10] The calculated bandgap wavelengths are 1525 nm and 1450 nm for the heavy- and light-hole transitions, respectively. The close agreement between calculation and measurement is within the bounds set by the uncertainty in the InGaAs (1,1,0) composition and suggests that the tunneling-resonance approach should be useful in designing more complicated multiple-quantum-well structures.

In conclusion, we have demonstrated a novel approach to growing a range of materials with different bandgaps and strains in a single device structure by using machine-state switching to produce pseudo alloys. This approach, which is an extension of previous work using short-period superlattices or "digital alloys", avoids the necessity of using multiple sources or changing source fluxes. A separate-confinement quantum-well LED grown with pseudo-quaternaries composed of ternaries and quaternaries is significantly brighter than that of a more traditional single quantum-well device using bulk lattice-matched alloys. Our general method should find applications in extending the range of complex structures that can be grown for optoelectronic applications.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge expert technical assistance from J. L. Dishman, J. Avery, and J. M. Sergeant. This work was supported by the United States Department of Energy under Contract DE-AC04-94AL85000.

REFERENCES

1. R. C. Miller, A. C. Gossard, D. A. Kleinman and O. Munteanu, Phys. Rev. **B29**, 3740 (1984).

2. M. C. Tamargo, R. Hull, L. H. Green, J. R. Hayes and A. Y. Cho, Appl. Phys. Lett. **46**, 569 (1985).

3. L. R. Dawson, Bull. Am. Phys. Soc. **39**, 406 (1994).

4. I. J. Fritz, L. R. Dawson, J. A. Olsen and A. J. Howard, Appl. Phys. Lett. **67**, 2320 (1995).

5. I. J. Fritz, J. F. Klem, M. Hafich and A. J. Howard, Electron. Lett. **31**, 829 (1995).

6. G. C. Osbourn, P. L. Gourley, I. J. Fritz, R. M. Biefeld, L. R. Dawson and T. E. Zipperian, in *Semiconductors and Semimetals,* edited by R. K. Willardson and A. C. Beer (Academic, San Diego, 1987), vol. 24, chap. 8.

7. T. H. Glisson, J. R. Hauser, M. A. Littlejohn and C. K. Williams, J. Electron. Mater. 7, 1 (1978).

8. A. Harwit, J. S. Harris and A. Kapitulnik, J. Appl. Phys. **60**, 3211 (1986).

9. G. Bastard, Phys. Rev. B **25**, 7584 (1982).

10. I. J. Fritz, J. F. Klem, T. M. Brennan, J. R. Wendt and T. E. Zipperian, Superlatt. and Microstr. **10**, 99 (1991).

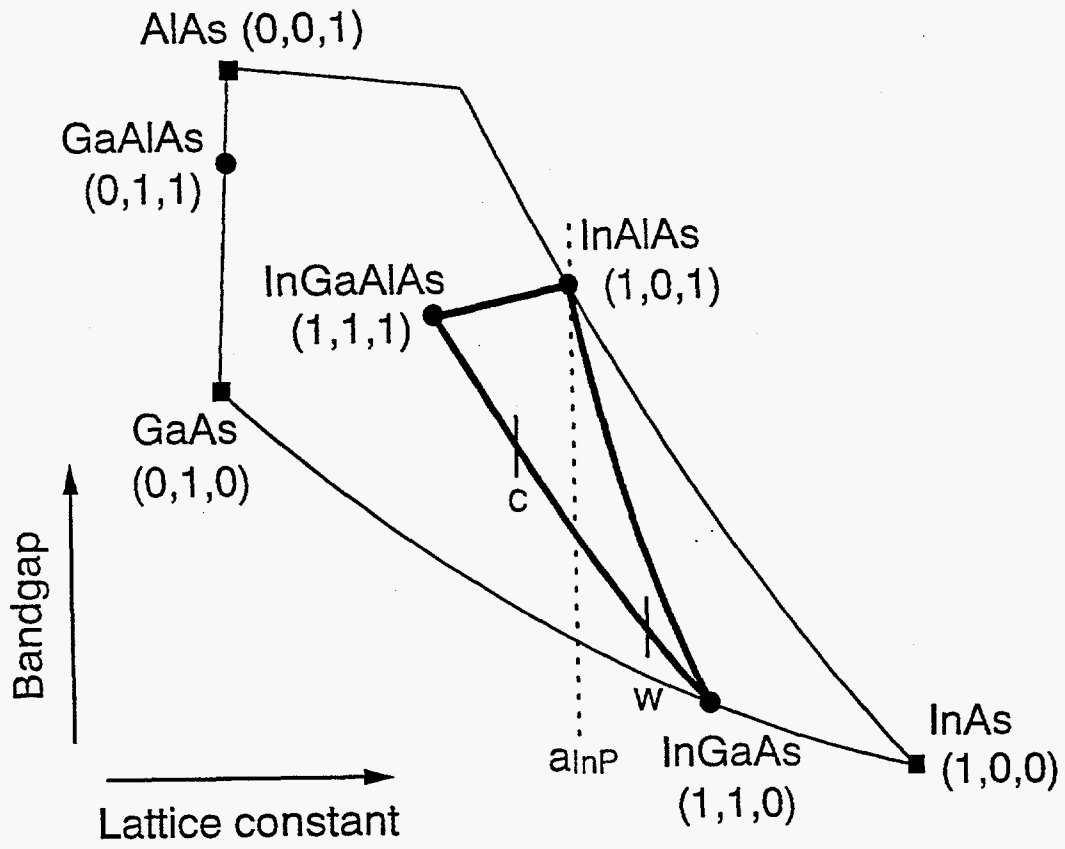


Fig. 1 Schematic illustration of bandgap vs. lattice constant relation for InGaAlAs materials used in present study. Solid squares and circles indicate materials grown with various combinations of shutter openings (states) of the MBE machine, denoted by the shorthand notation (i,j,k), where i, j, k = 0, 1 indicates whether the In, Ga, or Al shutters, respectively, are closed or open. Materials within the dark line may be approximated by pseudo alloys using short-period superlattices from the three states indicated by solid circles. Pseudo-quaternaries along the (1,1,0) - (1,1,1) line are used in the present study to make a strain-balanced separate-confinement quantum-well structure, with well and confinement materials indicated by w and c on the figure.

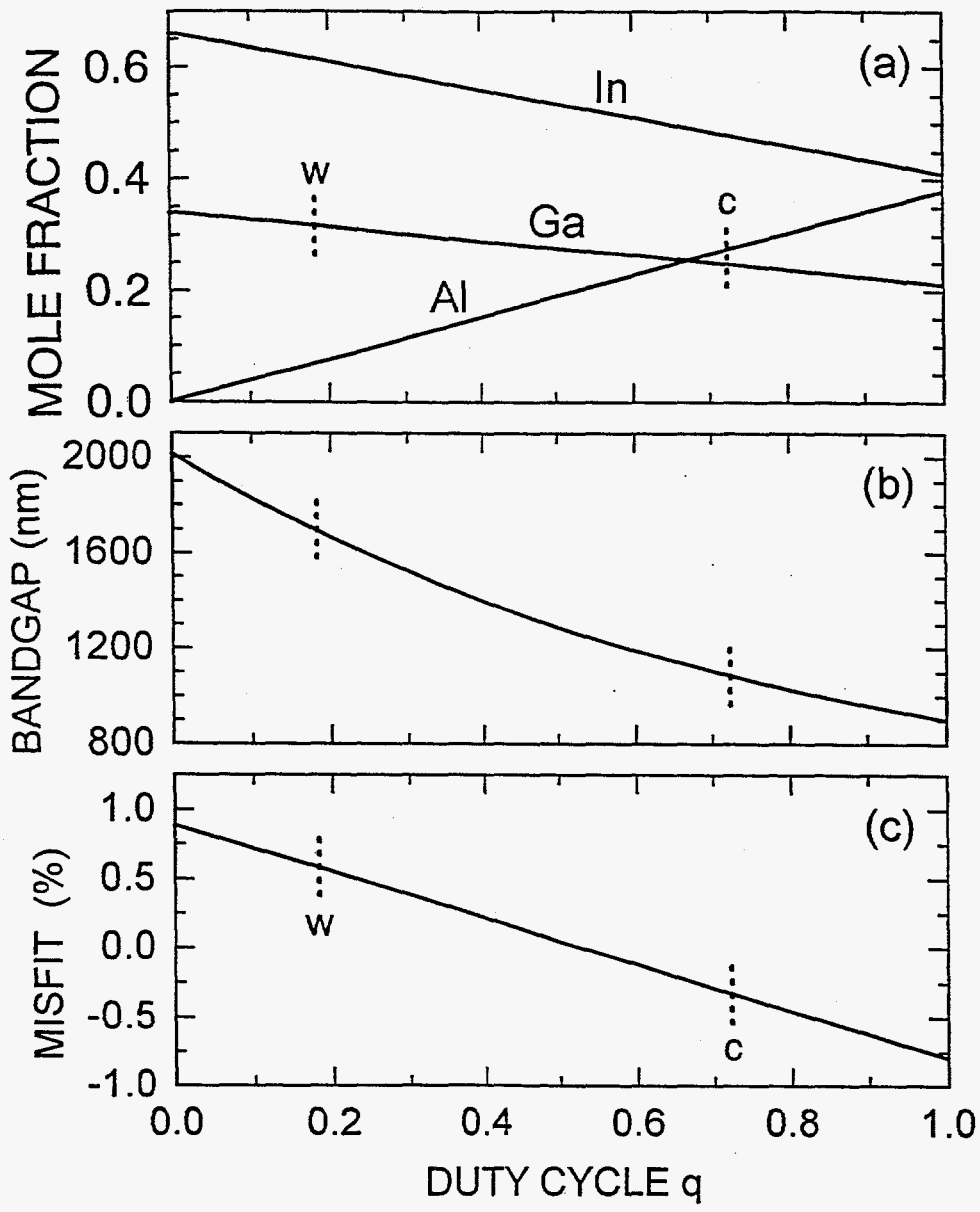


Fig. 2 Properties of the quaternary **Q**=(1,1,q) assuming perfect mixing. Shown are (a) mole fraction of the constituents, (b) bandgap wavelength (bulk, unstrained) and (c) lattice mismatch to InP substrate as a function of q. The quantum-well and separate-confinement materials are indicated by w and c on the figure.

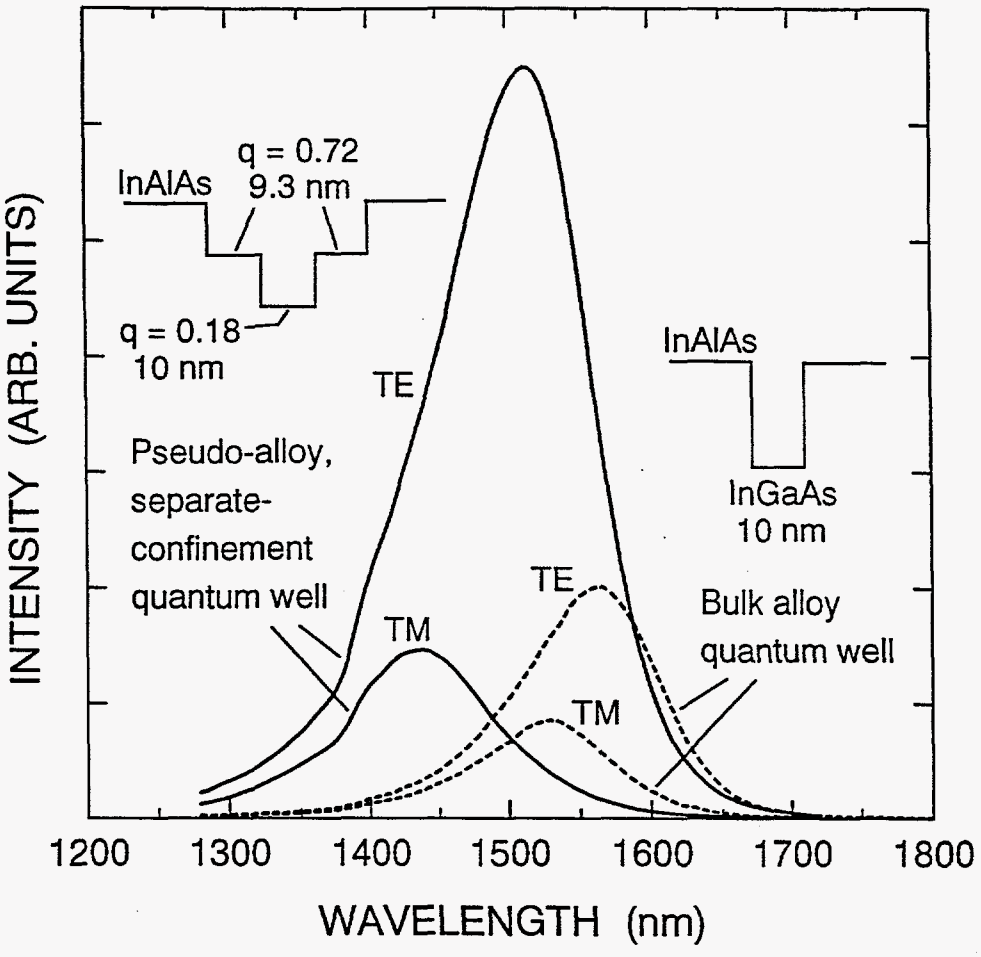


Fig. 3 Electroluminescent emission intensity for the pseudo-alloy, separate-confinement LED and a reference LED having a single, bulk-alloy quantum well. The two quantum-well structures are shown in the inserts. Emission was measured from the edge of the devices with TE or TM polarization as indicated. The pseudo-alloy device is brighter than the reference device, probably because of the separate-confinement layers.

BROADBAND LIGHT-EMITTING DIODES FROM PSEUDO-ALLOY QUANTUM WELLS

I. J. Fritz, J. F. Klem, M. J. Hafich and A. J. Howard
Sandia National Laboratories, Albuquerque, New Mexico 87185-0603

ABSTRACT

We demonstrate broadband light-emitting diodes with output over the entire 1.4 to 1.9 μm wavelength range. These devices employ digital pseudo-alloys, comprising InGaAs/InGaAlAs short-period superlattices, to construct quantum wells, quantum barriers and separate confinement layers with different band gaps and strain.

Broadband light-emitting diodes (BBLEDs) employing quantum wells with different band gaps [1] have many potential applications, for example in sensor systems and customized miniature spectrometers. [2] Growth of these devices is challenging because of the need to incorporate, in a single structure, layers with all the different compositions needed for the various quantum wells, quantum barriers and separate confinement layers. In this Letter we describe a versatile approach to growing devices with multiple InGaAlAs compositions by using pseudo-alloys (short-period superlattices) made from ternary and quaternary layers intergrown in different combinations. We demonstrate this approach by the molecular-beam-epitaxy (MBE) growth of BBLEDs emitting over the 1.4 - 1.9 μm wavelength range.

Our structures were grown on InP substrates using single effusion cells for the In, Ga, Al and As sources. The In:Ga flux ratio was set to grow $In_{0.66}Ga_{0.34}As$ when the shutters of the In, Ga and As cells were open. Similarly, the In:Al flux ratio was set to grow $In_{0.52}Al_{0.48}As$. With large As overpressure, flux rates are approximately additive, so with all shutters open the quaternary compound $In_{0.41}Ga_{0.21}Al_{0.38}As$ is grown. The active regions of our BBLEDs consist of layers of various pseudo-alloys grown by leaving the In and Ga shutters open and switching the Al shutter open and shut with a duty cycle q, thereby producing the effective compound $q\cdot(In_{0.41}Ga_{0.21}Al_{0.38}As)+(1-q)\cdot(In_{0.66}Ga_{0.34}As)$. As q varies from 0 to 1, the band gap wavelength of this pseudo-quaternary varies from 2.0 to 0.9 μm and its mismatch to InP varies from +0.9 to -0.8%.

Figure 1 is a schematic of the active regions of our BBLEDs, consisting of three compressively strained quantum wells (q=0, 0.18 and 0.36) along with strain-compensating barriers and separate confinement layers (q=0.72). These quantum-well structures are located in the p-doped region of InAlAs junction diodes. Cap layers of 2-nm-thick $In_{0.53}Ga_{0.47}$ As were added for contacting. Mesa devices, 0.4 - 1.5 mm square, were defined by standard photolithography and wet chemical etching. Full-surface Ge/Au/Ni/Au substrate metallization and Ti/Pt/Au rings on the mesas provide electrical contacts. The wafers were cleaved along the mesas edges to allow side emission to be studied as well as top emission. Emission spectra were obtained under dc electrical excitation by collecting the output light at f/3.6 and analyzing it with a 0.25 m monochromator followed by a liquid-nitrogen cooled InAs detector. Data were normalized by comparison to a calibrated tungsten-halogen lamp.

In Fig. 2a we show spectra from a triple quantum-well structure with 10-nm-thick wells, 6.7-nm-thick barriers and 21-nm-thick separate-confinement layers. In Fig. 2b we show comparison spectra (labeled SQW1, SQW2 and SQW3) from three diodes with single quantum wells similar to the wells in the triple-well structure. Samples SQW1 and SQW2 have q=0.36 and q=0.18 wells, respectively, and q=0.72 separate-confinement layers, whereas SQW3 is an $In_{0.66}Ga_{0.34}As/In_{0.52}Al_{0.48}As$ quantum well with no separate confinement layers. Figure 2a includes separate spectra for top emission (with no polarization selection) and for side emission (with TM-mode polarization). In the TM spectrum, heavy-hole transitions are forbidden to first order, and we expect the n=1 light-hole transitions (11L) to dominate. For top emission both heavy- and light-hole transitions are allowed, but the n=1 heavy-hole transitions (11H) are expected to dominate, because of the large heavy-light hole splittings due to the compressive strains. We have predicted the positions of the 11H and 11L transitions in our structures by using a tunneling resonance calculation in the two-band envelope-function formalism. The results are shown as vertical bars in Fig. 2a. The predicted line positions agree well with measured features

in the spectra, although the emission from well #1 (near 1.3 μm) is quite weak. Note that this device has useful output between 1.40 and 1.90 μm, a range of 500 nm.

The relatively weak output of well #1 is confirmed in Fig. 2b, where we see that SQW1 requires about four times as much drive current as SQW2 for the same peak output. In addition, much of SQW1's emission is from the n=2 heavy-hole transition (HH2). Thus it appears that well #1's emission in the triple-well structure is weak due to it being inefficient both in capturing carriers and in thermalizing those that it does capture down to the ground state. Studies of carrier capture in quantum wells have shown that the capture and radiation processes depend strongly on the details of the energy level structure of the system. [3] To improve the emission of wells #1 and #2 in our structure, we have grown a structure where their widths are increased to 13 and 11.5 nm, respectively. Calculations show that each well should now confine three electron levels, with the n=3 levels all being near the tops of the wells. In contrast, the structure with 10-nm-thick wells only confines two levels in wells #1 and #2. In Fig. 3 we show the top emission from our second triple-quantum-well structure at three current levels. The emission from the #1 and #2 wells have been significantly enhanced relative to the #3 well, with the second well actually having the strongest output. Further work to improve the balance of the three emission lines is in progress.

Measured angular dependences of top emission at several wavelengths show an approximately cosinusoidal dependence. From this dependence we estimate an integrated top emission intensity of approximately 250 μW. Measurements of current-voltage characteristics on devices of different areas, along with visual observations using an infrared viewer, show that our devices suffer from current crowding effects at operational currents. Preliminary data on edge emission from fully electroded bars suggest that integrated single-surface outputs of ~1 mW can be easily achieved with our structures.

In conclusion, we have demonstrated a method of incorporating compressively strained quantum wells with a range of band gaps into a broadband LED structure. Significant emission has been obtained over a wavelength range of ~500 nm in the near infrared. Our technique should be applicable to other near-infrared wavelength ranges for a variety of sensing and spectroscopic applications.

ACKNOWLEDGMENT

It is a pleasure to acknowledge expert technical assistance from J. L. Dishman, J. Avery and J. M. Sergeant. This work was supported by the United States Department of Energy under Contract DE-AC04-94AL85000.

REFERENCES

1. MOSELEY, A. J., ROBBINS, D. J., MEATON, C., ASH, R. M., NICKLIN, R., BROMLEY, P., BRADLEY, R. R., CARTER, A. C., HONG, C. S., and FIGUEROA, L., "Broadband GaAs/Al_xGa_{1-x} Multi-Quantum Well LED", in Quantum Optoelectronics, 1991, Technical Digest Series, Vol. 7 (Optical Society of America, Washington 1991), pp. 193-196

2. WANG, X., VAUGHAN, D. E., PELEKHATY, V., and CRISP, J., "A novel miniature spectrometer using an integrated acousto-optic tunable filter", *Rev. Sci. Instrum.*, 1994, **65**, pp. 3653-3656

3. BRUM, J. A., and BASTARD, G., "Resonant carrier capture by semiconductor quantum wells", *Phys. Rev. Lett.*, 1986, **33**, pp. 1420-1423

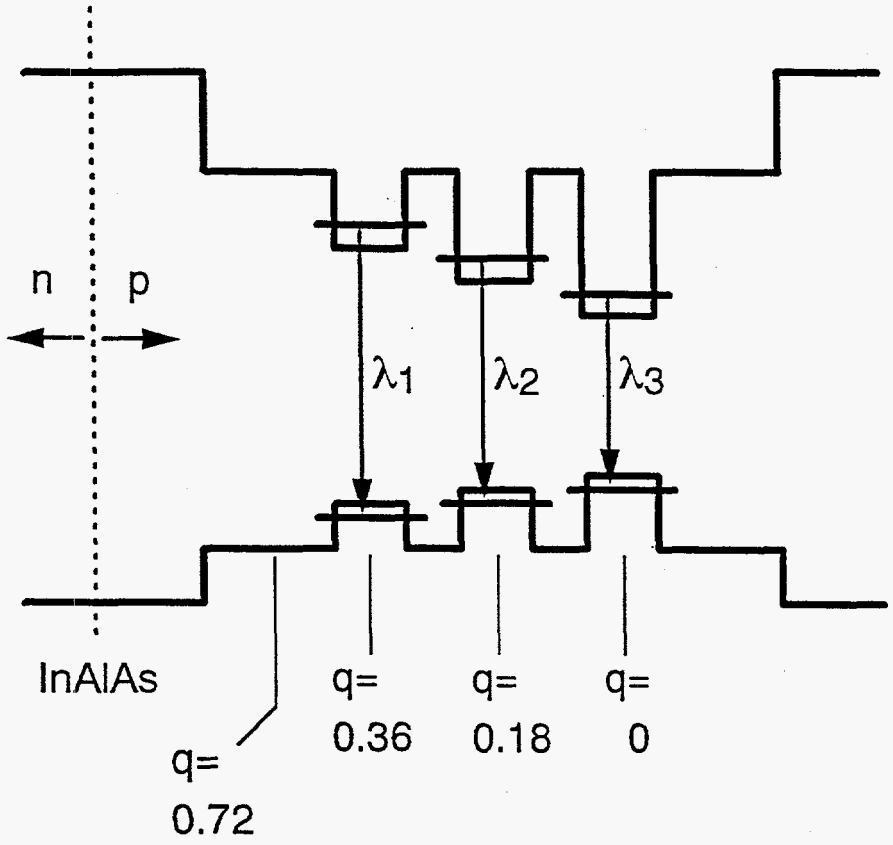


Fig. 1 Schematic of active region of broadband light-emitting diode employing pseudo-quaternary-alloy quantum wells with different band gaps.

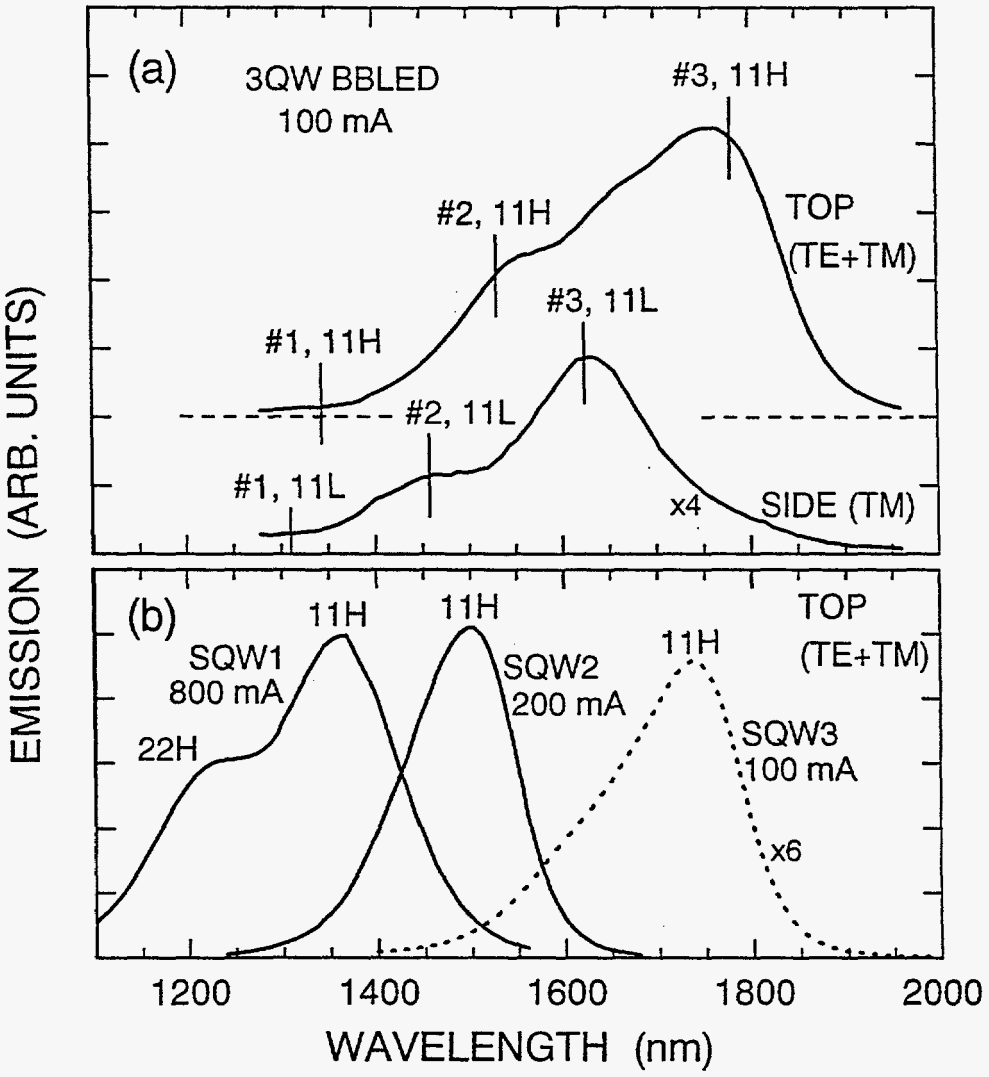


Fig. 2 Emission spectra from: (a) a triple-quantum-well LED and (b) three single-quantum-well LEDs having wells similar to those in the triple-well structure.

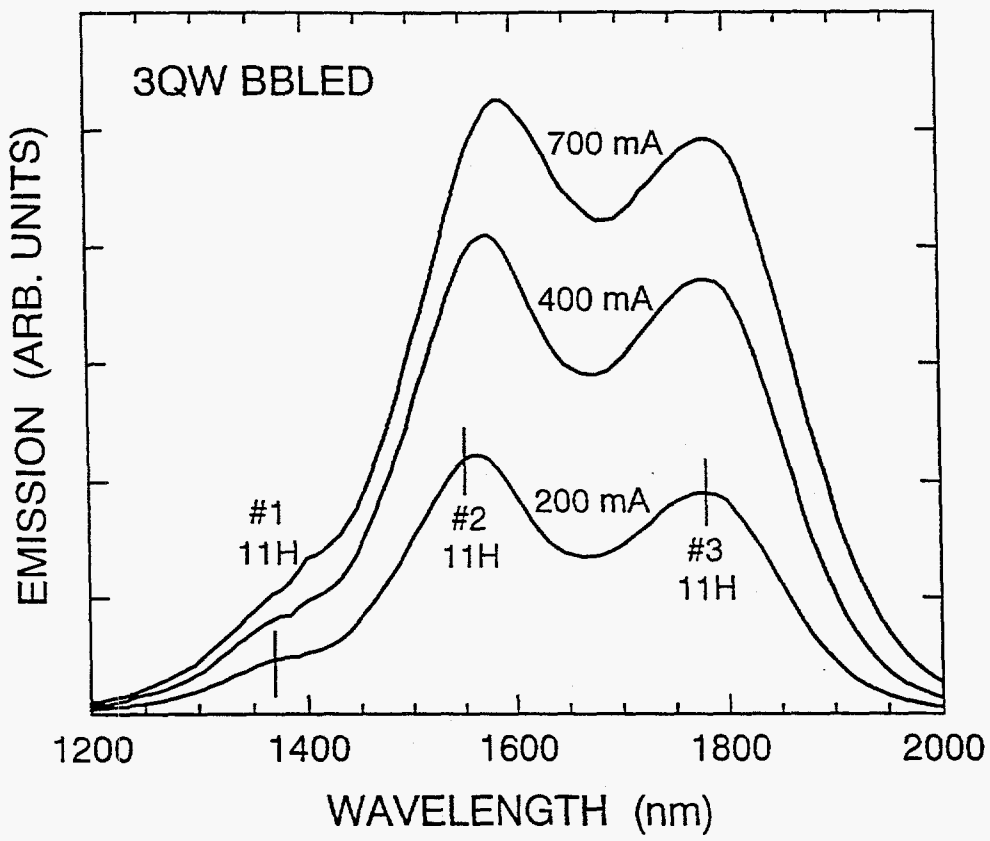


Fig. 3 Top-emission spectra of a triple-quantum-well structure with well widths modified to improve carrier capture and radiation.

Near-IR Reflectance Spectroscopy for the Determination of Motor Oil Contamination in Sandy Loam†

Brian R. Stallard,* Manuel J. Garcia, and Sumanth Kaushik
Sandia National Laboratories, Albuquerque, NM 87185-1405

ABSTRACT

We have investigated the application of near-ir reflectance spectroscopy to the determination of motor oil contamination in sandy loam. Although the present work is concerned with a specific case of contamination, we discuss the possibility of applying the method to other organic contaminants and other types of soil. The spectral region considered was 1600 to 1900 nm which contains the first overtone of the CH stretch. Using a commercial Fourier transform spectrometer together with cross-validated partial least squares data analysis, the one sigma precision for the determination of motor oil in sandy loam was 0.17 wt% (0.13 to 0.26 wt% at the 95% confidence level). The largest contribution to the precision of the determination was sampling error, or inhomogeneity in each sample. Given the precision limit imposed by the sampling error, we found that the performance of the spectrometer could be lowered without impacting the overall precision. In a modeling exercise, adequate performance was obtained with a spectrometer having only 7 spectral channels with a spectral resolution of 10 nm and a spectral noise level of 10^{-3} absorbance units. A design for an inexpensive miniature instrument is presented.

Index Headings: Near-ir spectroscopy; Soil contamination; Chemometrics; Environmental applications of optical spectroscopy; Reflectance spectroscopy; Field instrumentation.

† This work was supported by the U. S. Department of Energy under contract number DE-AC04-94AL85000.

* Author to whom correspondence should be sent.

INTRODUCTION

Detecting organic chemical contamination in soil is a major concern as environmental laws are applied to present and past industrial activities. Standard procedures require that the samples be taken from the suspected site to the laboratory. Determination of organic contaminants is typically performed by solvent extraction and gas chromatography, with detection limits for important classes of contaminants often in the ppm range.[1] After the appropriate samples are collected and lab results completed, a map of the suspected site is made. The overall process may be inefficient because it is often difficult for the scientists at the site to acquire the appropriate samples. Oversampling increases the cost of the laboratory work by increasing the number of negative results. Undersampling increases the cost of the field work by increasing the number of site visits required to complete the map. A field instrument for preliminary screening of samples would clearly improve the efficiency of the process.

We have investigated near-ir reflectance spectroscopy as a method for the determination of motor oil in sandy loam. Although the present work is concerned with a specific case of contamination, we discuss the possibility of applying the method to other organic contaminants and other types of soil. Related work has been reported, but includes little quantitative results.[2] In the present work, a measurement precision of 0.17 wt% was demonstrated. A laboratory spectrometer was used, though existing portable spectrometers would, no doubt, be adaptable to this task. We propose to develop an inexpensive miniature spectroscopic instrument for this application. The performance requirements and a suggested design for the instrument are discussed.

EXPERIMENTAL

All spectra were collected with a Nicolet 800 Fourier transform interferometer, fitted with a Spectra-Tech "seagull" reflectance accessory. The angle of the incident beam was 45°. Both the specular and diffuse reflected beams were collected by a wide angle mirror. An InSb detector was utilized and the resolution was set to 1.2 nm (4 cm^{-1}) at 1750 nm.

The sandy loam, the principal type of soil used in this work, was collected in an open area at Sandia. The potting soil was purchased at a local nursery. The garden soil was collected at the home of one of the authors. The soil samples were sized by collecting the portion retained between two sieves of hole sizes 0.71 and 0.425 mm. For the calibration samples, a quantity of sandy loam was sieved, dried, and equilibrated with the laboratory environment. Portions of about 20 g were weighed repeatedly over about 30 hours to ensure stability. Each sample was then taken through the following procedure in a serial fashion. The soil sample was weighed and a predetermined number of drops of oil were added. A second weighing gave a quantitative measure of the weight of oil added. The sample was mixed by hand with a spatula for five minutes and then placed in the spectrometer. After a two minute delay, the scanning was initiated. Care was taken to hold precisely to this timing sequence. The sample ordering was randomized according to the amount of oil added. The oil was Pennzoil 10-30 purchased from a commercial outlet.

The number of scans was 350 (about 70 seconds) for both the sample and background. All spectra, whether acquired in a transmission or reflection mode, were ratioed to the appropriate background and displayed on a scale of absorbance units. The transmission mode spectra in Fig. 1 were ratioed to an empty beam path. The reflectance mode spectra in Fig. 1 were ratioed to a gold mirror. The reflectance mode spectra used in the calibration of Fig. 2 were ratioed to a soil sample containing no motor oil.

Cross-validated partial least squares (PLS) data analysis[3-5] was used to determine the standard error of prediction (SEP) for the calibration samples. Calculations were performed on a personal computer using Lab Calc software and a program written in Array Basic by D. M. Haaland and D. Melgaard at Sandia. The SEP calculated through cross-validation is an estimate of the one sigma precision of the determination. The cross-validation procedure consists of constructing a series of models where a single calibration point is excluded for each model. The series is used to predict the concentrations of the excluded points, in turn, as if they were unknown. Finally, the predicted values of the excluded points are plotted versus their reference values. For perfect prediction, the data lie on a 45° line with zero intercept. The SEP is the root-mean-square of the deviations from the 45° line.

RESULTS AND DISCUSSION

General Spectral Data. Figure 1 shows several pertinent spectra in the region of the first overtone of the CH stretch. Figure 1a is the absorption spectrum of a thin film of water between two BaF_2 windows. Clearly, the near-ir band of water in this region does not interfere with the CH bands seen in subsequent portions of the figure. Figure 1b is the absorption spectrum of a thin film of motor oil between two KBr windows. Figure 1c is a reflectance spectrum of motor oil mixed with a sample of sandy loam. The oil content is about 1 wt%. Note that the reflectance spectrum is essentially identical to the true absorption spectrum. Sandy loam has little or no organic material which could mask the spectrum of the oil. However, some types of soil have measurable amounts of organic material. Figure 1d is a reflectance spectrum of a commercial potting soil which contains considerable organic material. The peak height of the interfering spectrum is similar to that of the soil containing about 1 wt% motor oil. Figure 1e is the reflectance spectrum of soil from a garden. The peak height of the interfering spectrum is about one fourth of that in Fig. 1d. The potting soil is probably a worst case for spectral interference from organic material, while the garden soil is probably more typical of organic rich natural soils. Figures 1d and 1e will be helpful in our consideration of spectral interferences below.

PLS Calibration and Estimated Precision. Figure 2 is a cross-validated PLS calibration curve. The samples used for the spectra were prepared using the same soil as in Fig. 1c. An SEP of 0.17 wt% was obtained with a two factor PLS model. With 17 samples the range for the SEP is 0.13 to 0.26 wt%, at the 95% confidence level. This calculation is accomplished in the usual fashion assuming a χ^2 distribution for the variance (i.e. SEP^2). Figure 3 shows two plots that explore the origins of the variance of the experiment. In Fig. 3a, spectra were scanned repeatedly using a single sample. The two factor PLS model constructed from the data of Fig. 2 was used to predict the wt% motor oil in each of the repeated runs. The precision of these determinations is best specified by the standard error of estimate (SEE) which is the standard deviation after the

trend modeled by the solid line is removed. The SEE is only 0.005 wt%, showing that the spectrometer is capable of a high precision determination of the motor oil content in the sandy loam. A loss of motor oil over time is evident in the negative slope of the linear fit to the data. Converting the slope to a time scale we find that the sample is losing 0.03 units of wt% per hour. Figure 3b is a repeated determination where a fresh portion of the 20 g sample was placed in the sample holder for each run. As in Fig. 3a, the model of Fig. 2 was used for the prediction. The SEE after a linear fit is 0.14 wt%. This error, which results from the inhomogeneity of the sample, will be referred to as sampling error. It accounts for nearly all of the variance represented by the SEP of 0.17 wt% derived from Fig. 2. Again we notice a negative slope. The loss of oil in Fig. 3b is 0.07 units of wt% per hour. Although similar to the loss in Fig. 3a, the difference is unexplained.

It is important to remember that the precision of the determination in Fig. 2 is dominated by the sampling error and is perhaps 30 times worse (0.14 wt% / 0.005 wt%, from Fig. 3) than the inherent capabilities of the instrument.

Modeling Other Soil Types. The sandy loam of Fig. 2 contains no organic matter detectable by the present method. As we have seen in Fig. 1, other types of soil contain organic matter that may interfere spectrally with the determination of the motor oil content. Therefore, a simple modeling study is presented to assess the feasibility of performing the determination in soils other than sandy loam.

Spectral interference was modeled by adding a randomized portion of the spectrum in Fig. 1d, the commercial potting soil, to each spectrum in the calibration set. Choosing a soil with very high organic content increases the difficulty of the challenge. Recall that the spectrum of a more typical organic rich soil, garden soil in Fig. 1e, may have a peak height of only one fourth that of potting soil. Before each spectral addition, the interfering spectrum was multiplied by a normally distributed random number having a mean of unity and a sigma of 0.25. The result is a simulated set of spectra referred to as the second calibration set. The original set of spectra without the added spectral interference is referred to as the first calibration set. The cross-validated SEP for the second calibration set was 0.17 wt%, the same as for the first.

Although this modeling exercise indicates that interference from other organic matter in the soil may not be troublesome, experimental results for a wide variety of field samples will be required to have complete confidence in this conclusion. The model lets an organic interferant vary randomly but stipulates that there is only one type of interferant. This may be realistic for mapping a small contaminated area, but optimistic for mapping a large area. As chemometricians remind us, the key to building a robust PLS calibration model is to include in the training set the full range of variations, including interferants, that will be seen in the unknown samples. In other words, avoid extrapolation but have confidence in interpolation. Hence, the ultimate success of this method in the face of large and varying spectral interferences depends on the attention that the user gives to building an adequate training set.

Specifications for a Field Instrument. Because of the dominance of sampling error, the high resolution and excellent signal-to-noise ratio (SNR) of the laboratory spectrometer are underutilized. To investigate the minimal performance required, the following design criteria are discussed: spectral resolution, number of spectral channels, and SNR.

Both calibration sets were deresolved (by averaging neighboring data points) to a spectral resolution of about 10.8 nm (at 1750 nm) giving 21 spectral channels over the region of interest. This resolution was chosen since it is readily achieved by the interference filter array that is discussed in the design of the new instrument below. If this spectral resolution proves adequate, there is no motivation to go to a higher resolution. A cross-validated PLS calibration, using the low resolution spectra, produced SEPs of 0.17 wt% and 0.18 wt% for the first and second calibration sets, respectively. This compares to an SEP of 0.17 wt% for both of the high resolution calibration sets, and demonstrates that lowering the resolution does not appreciably reduce the precision of the determination.

The next concern is the number of spectral channels. Figure 4a shows the trend in going from three channels to the full 21 channels, using the low resolution version of the second calibration set. In all cases, a two factor PLS model was adopted (except, of course, for the three channel case, where only one factor is available after fitting a linear baseline). For a given number of channels the best combination is plotted. This is, in general, two channels to establish the baseline, plus a few channels near the center of the absorption bands. Figure 4a shows that a few, well chosen, channels produce a lower SEP, and hence a superior result, than the full spectrum. Apparently, some channels add more noise than information. Based on Fig. 4a, the instrument can have as few as 4 spectral channels. However, 5 to 10 spectral channels may be more sensible in order to retain flexibility in applying the instrument to wide variety of problems.

A final task is to examine the spectral noise requirements. From Fig. 1, we see that the spectral noise in the commercial spectrometer is imperceptible even when plotting these weak absorption bands. In fact, the Nicolet 800, under the conditions of the experiment, had a spectral noise of about 10^{-5} absorbance units. Taking the low resolution version of the second calibration set, we added normally distributed noise to simulate a poorer performing instrument. Figure 4b shows the effect on the SEPs of adding noise to the 7 channel calibration set of spectra. The x scale is four times the sigma value (i.e. two times the full width at 1/e of the maximum) of the added noise, in absorbance units. The fact that the SEPs appear to improve slightly with small additions of noise is unexpected and presumably not statistically significant. The main deduction from Fig. 4b is that, if the noise is kept under about 10^{-3} absorbance units, it does not degrade the SEP. This noise level is equivalent to about 0.14 wt% of oil, if we take from Fig. 1c that a peak height of about 7×10^{-3} absorbance units corresponds to an oil content of about 1 wt%.

Design for a Field Instrument. From the last subsection we conclude that a practical instrument should have about 7 channels with a spectral resolution of about 10 nm, and a spectral noise level less than 10^{-3} absorbance units. Future work with other organic contaminants and other types of soil may uncover the need to modify this design specification. Figure 5 shows the concept for a proposed instrument. The source has been developed at Sandia especially for near-ir sensors. It is a multiple quantum well broad band diode with an output which can be tailored to the present needs of 1.6 to 1.9 μm.[6] Its advantages compared to a more traditional tungsten source are high efficiency, high brightness, and ease of modulation (via the applied current) for phase sensitive detection, if required. The wavelength selective component is also a development from Sandia.[7] It is a Fabry-Perot filter array, where the individual passbands are tuned by sub-wavelength lithography which adjusts the refractive index of the material between the mirrors. The advantage of this device is that the pixels may be made very small and spectrally flat. Also, two dimensional arrays may be constructed with arbitrarily chosen passbands among the pixels.

The detector itself is a two dimensional PbS array, where the pixels are 100 μm squares. The optical portion of the instrument could be as small as a few centimeters on a side. It will require no detector cooling, no purging, and can be operated with battery power.

Some Practical Issues. Further development will be required to insure that this is a robust technique in field applications. In the beginning, each contaminated site will require a new calibrated training set, or at least a verification that the type of soil contains no spectral interferant. Eventually, it may be possible to develop a universal and transferable calibration. As mentioned earlier, successful application of the technique in the face of spectral interferants will depend on constructing a training set of truly representative samples. Another important question concerns the specific organic compounds that can reliably be determined. No doubt, a large number can be detected since many compounds of interest have CH overtones in the spectral region 1600 to 1900 nm. However, we do not know how well the components of a mixed contamination can be distinguished? Also, sample drying issues must be explored. What does a wt% calibration mean if the amount of water is variable from sample to sample? There are certainly other unresolved issues, but this short summary gives some direction for future research.

CONCLUSIONS

We have shown that near-ir reflectance spectroscopy can be used to determine the motor oil content of contaminated sandy loam. The one sigma precision of the determination is 0.17 wt% motor oil, where the principal contribution to the precision is the inhomogeneity of the sample. A small portable instrument has been proposed having 7 spectral channels, each with a resolution of 10 nm, and a spectral noise of 10^{-3} absorbance units. Such an instrument would have a similar precision, in the present application, as the laboratory instrument.

ACKNOWLEDGMENTS

We acknowledge the support of the Laboratory Directed Research and Development Program at Sandia National Laboratories. Z. Wade, a local high school student, helped acquire and process the data.

REFERENCES

1. D. Barcelo, edt., *Environmental Analysis, Techniques, Applications, and Quality Assurance,* (Elsevier, New York, 1993).

2. I. Schneider, G. Nau, T. V. V. King, and I. Aggarwal, *IEEE Photonics Tech. Lett.* 7, 87 (1995).

3. M. P. Fuller, G. L. Ritter, and C. S. Draper, *Appl. Spect.* **42**, 217 (1988).

4. D. M. Haaland, R. G. Easterling, and D. A. Vopicka, *Appl. Spect.* **39**, 73 (1985).

5. E. V. Thomas and D. M. Haaland, *Anal. Chem.* **62**, 1091 (1990).

6. I. J. Fritz, private communication.

7. S. Kaushick and B. R. Stallard, *SPIE* **2532**, page number not yet assigned (1995).

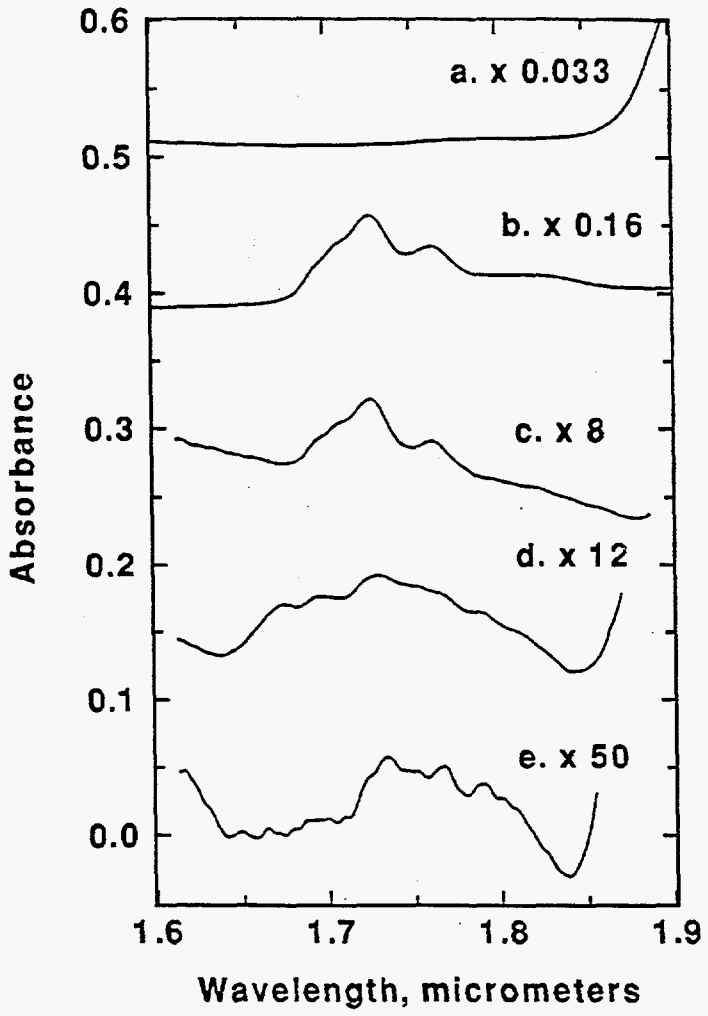


Fig. 1 **(a)** A spectrum of liquid water taken in transmission mode. **(b)** A spectrum of motor oil taken in transmission mode. **(c)** A spectrum of about 1 wt% motor oil in sandy loam taken in reflectance mode. **(d)** A spectrum of potting soil taken in reflectance mode. **(e)** A spectrum of garden soil taken in reflectance mode. The last two spectra have been leveled by subtracting a linear baseline. The y-scale for each subfigure has been multiplied by the indicated factor relative to the y scale shown.

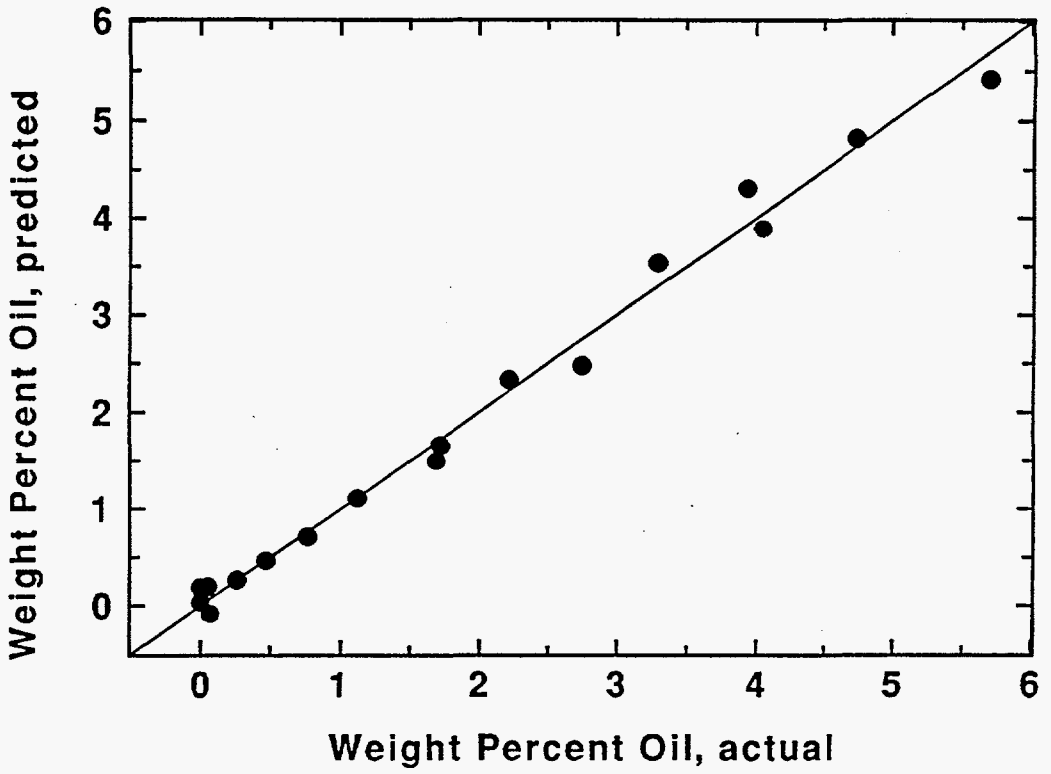


Fig. 2 A cross-validated PLS calibration plot of the motor oil content of sandy loam. The SEP is 0.17 wt%. The spectrum for each point is similar to Fig. 1c.

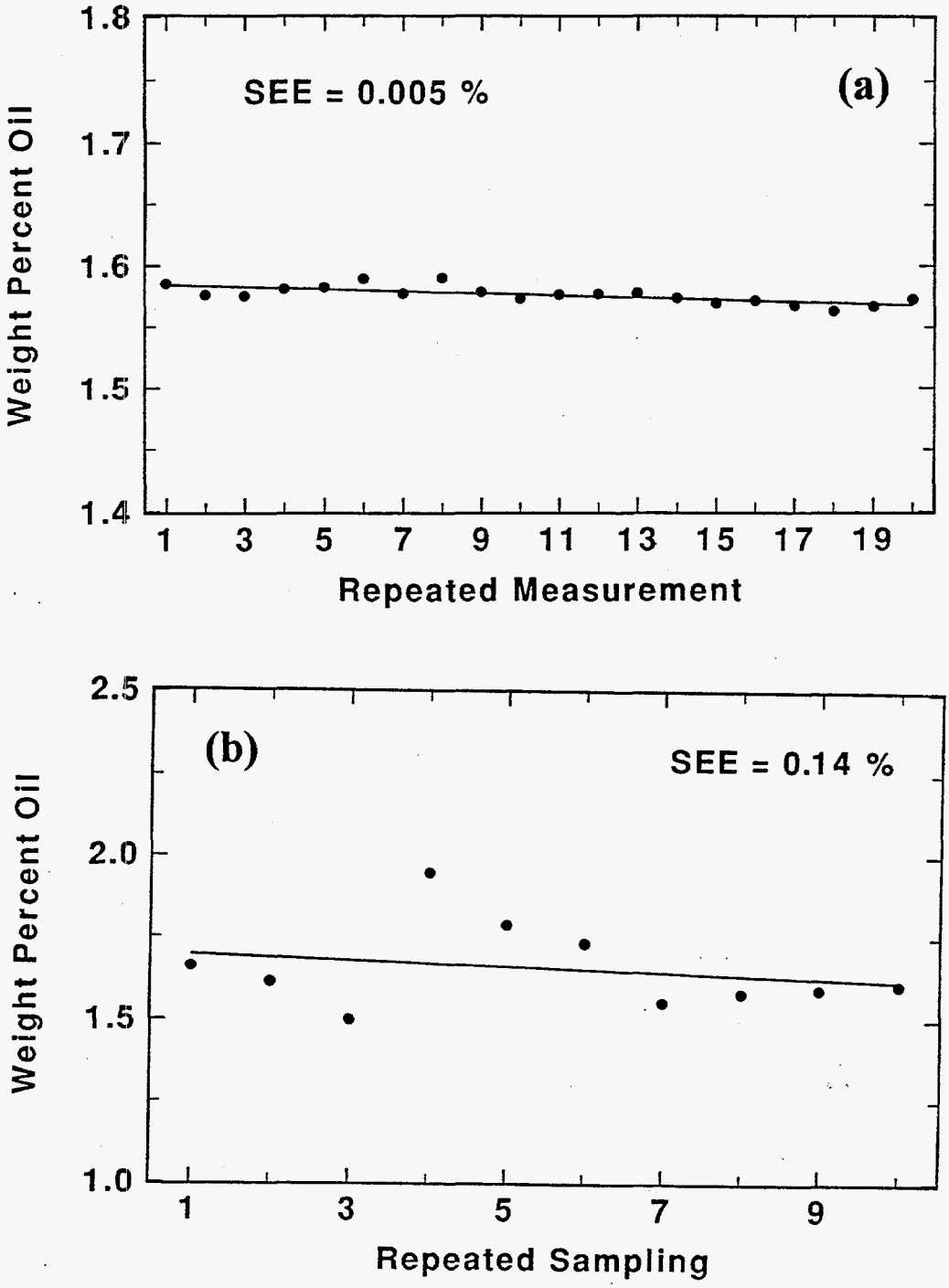


Fig. 3 **(a)** The predicted concentration of motor oil in sandy loam for repeated measurements on the same sample. The SEE is 0.005 wt%. **(b)** The predicted concentration of motor oil in sandy loam for repeated small samples taken from the same large sample. The SEE is 0.14 wt%. In both (a) and (b), the predictive model was derived from the data of Fig. 2.

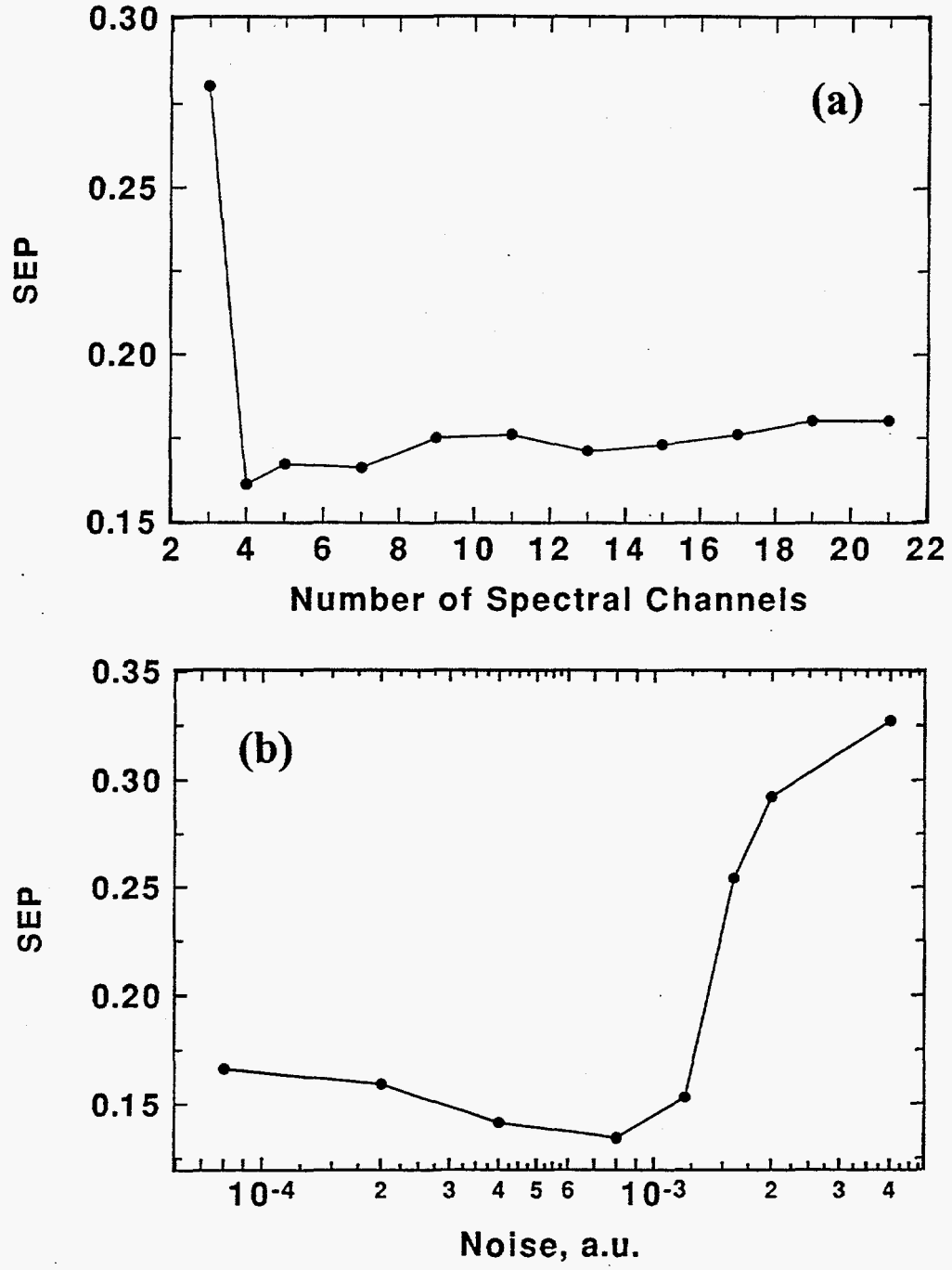


Fig. 4 **(a)** The change in SEP as a function of the number of channels in the calibration set of spectra. **(b)** The change in SEP as a function of noise added to the 7 channel calibration set. The spectra, in both cases, are versions of the second calibration set (i.e. including the spectral interference of Fig. 1d) that have been deresolved to a resolution of about 10.8 nm.

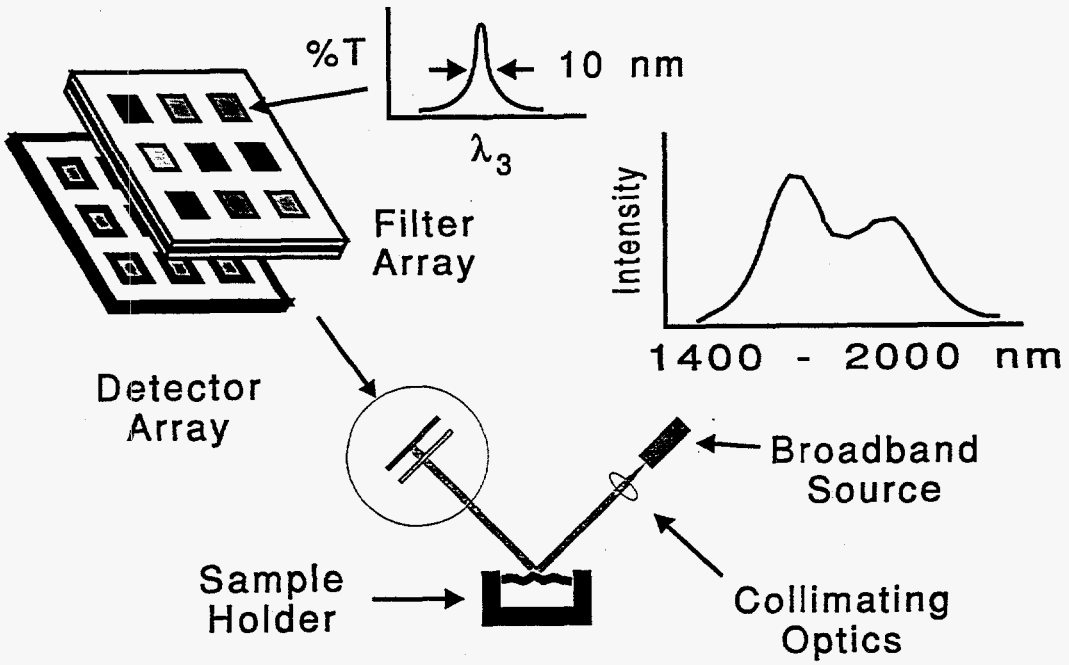


Fig. 5 A schematic drawing of the optical portion of a small portable instrument suggested for measuring organic contamination in soil samples.

Near-IR Versus Mid-IR, Separability of Three Classes of Organic Compounds

Brian R. Stallard

Sandia National Laboratories, Albuquerque, NM 87185-1411

ABSTRACT

Recently there has been a surge of interest in spectroscopic sensors operating in the near-IR. Yet, it is recognized that the mid-IR contains more spectral information. The general question addressed in this paper is how much specificity is lost in choosing the near-IR over the mid-IR for sensor applications. The example considered is the separability among three classes of organic compounds: alkanes, alcohols, and ketones/aldehydes. We use spectra from two sources: the Hummel polymer library (mid-IR) and the library of Buback and Vögele (near-IR). This is the first paper on class separability to make use of this new near-IR library, available in digital form only since July 1995. Five spectral regions are considered: region 5, 10500 to 6300 cm^{-1}; region 4, 7200 to 5200 cm^{-1}; region 3, 5500 to 3800 cm^{-1}; region 2, 3900 to 2500 cm^{-1}; and region 1, 2500 to 500 cm^{-1}. Class separability is explored both qualitatively and quantitatively using principal component scatter plots, linear discriminant analysis, Bhattacharyya distances, and other methods. We find that the separability is greatest in region 1 and least in region 2, with the three near-IR regions being intermediate. Furthermore, we find that, in the near-IR, there is sufficient class separability to insure that organic compounds of one class can be determined in the midst of interference from the other classes.

Index Headings: Near-IR spectroscopy; IR spectroscopy; Chemometrics; Class separation.

INTRODUCTION

Recently there has been a surge of interest in spectroscopic sensors operating in the near-IR.[1-6] Yet, historically the mid-IR has received much more attention from applied spectroscopists. This paper discusses the relative specificity available in these two spectral regions. The general problem considered is the detection or determination of organic compounds in the presence of spectral interferences from related classes of compounds.

Fabrication techniques that have revolutionized the design and manufacture of electronic components are now being applied to optical sensors.[7,8] This potential for cheap and small devices appears much greater in the near-IR than the mid-IR. We expect this to be a major driving force to increase the interest in near-IR sensors. The near-IR spectral region has additional significant advantages.[5,6,9] Detectors are more sensitive and cooling is less of an issue. Sources are brighter and more efficient. Pathlengths for liquid samples are more convenient (about 1 to 5 mm versus 10 to 50 μm). Also, diffuse reflectance, which is very popular in a number of industrial applications, is more efficient with the shorter wavelengths.

On the other hand, a well known rule-of thumb states that the absorption cross-sections for molecular vibrations decrease from the fundamental (in the mid-IR) by one decade for each order in the overtone spectrum. This issue will be ignored in the present work, since it can readily be added to the model when the need arises. As a further justification, we note that, for a similar cost, near-IR sensors generally have a higher signal-to-noise ratio than mid-IR sensors.

A second problem in working in the near-IR is that there is not as much unique spectral information. Classes of compounds may be less distinguishable and spectral backgrounds may be more likely to interfere. The importance of this consideration has been diminished by modern multivariate spectral analysis. Nevertheless, we intuitively feel (and correctly so) that there is more information in the well known fingerprint region (1800 to 400 cm^{-1}) than in the near-IR. The purpose of this paper is to determine how much specificity is lost in choosing the near-IR over the mid-IR for sensor applications. This information is needed by the design engineer who may find it cheaper to build a near-IR sensor, but must assess the loss in specificity that this choice will entail.

This paper considers only broadband spectroscopic sensors which might be employed to detect large molecules in the gas phase or any size molecules in the condensed phase. Gas phase detection of small molecules with narrow spectral features is not discussed since the critical considerations are distinctly different.

Recently the first large digital spectral library of near-IR spectra covering the range 10500 to 3800 cm^{-1} has become available.[10] This paper is the first to use the new spectral library to answer questions regarding class separability.

MATERIALS AND METHODS

Spectra were drawn from two libraries. The Hummel polymer library, acquired from Nicolet Analytical Instruments, covers the spectral region 3800 to 500 cm^{-1}, while the library of Buback and Vögele, acquired from Chemical Concepts, covers the spectral region 10500 to 3800 cm^{-1}.

The increment between data points was set to a constant 16 cm^{-1} for the combined regions. A simple average of neighboring points was used to reduce the resolution when required.

Data analysis was accomplished principally with the Windows version of the statistical program S-Plus, from MathSoft. Distributed routines were used where possible, yet a certain amount of programming was required. The partial least squares (PLS) calculations were performed with software developed by David M. Haaland and written in the array basic language to be run in the environment of GRAMS/386 from Galactic Industries.

RESULTS AND DISCUSSION

The classes and spectral regions. Three classes of compounds were drawn from the two spectral libraries: alkanes, alcohols, and ketones/aldehydes. To account fully for the within-class variance there should ideally be a large number of spectra to represent each class. In this case, the number was limited since only compounds found in both libraries could be considered. Table 1 contains a list of the compounds whose spectra were used.

The complete spectrum of each compound was partitioned into the five spectral regions: region 5, 10500 to 6300 cm^{-1}; region 4, 7200 to 5200 cm^{-1}; region 3, 5500 to 3800 cm^{-1}; region 2, 3900 to 2500 cm^{-1}; and region 1, 2500 to 500 cm^{-1}. Within each region the spectra were scaled to give an absorbance range from 0 to 1. Figure 1 contains the mean spectrum for each of the three classes in the five spectral regions. Studying this figure gives an intuitive feel for the separability among the classes. The rest of the paper is concerned with methods to explore and quantify the separability.

Dimension reduction and visualization. One way to visualize the class separation is in a scatter plot, after the dimension of the spectra have been reduced from several hundred to only three. A popular method of reducing the dimension of a data set is known as principal component analysis (PCA).[11-13] The vector corresponding to the direction of maximal variance is determined for the full dimensional data set. Subject to the constraint of orthonormality, the procedure is repeated for the direction of next maximal variance as many times as suits the problem; usually until an acceptable fraction of the total variance is captured. The spectra are then represented by their expansion coefficients (known as scores) in the new set of basis vectors (known as factors, or specifically PC1, PC2, etc.). Figure 2 contains scatter plots of the scores of the first three PCs for all the classes and regions. As expected, region 1 shows the best class separability. Region 2 shows perhaps the least, with the other regions being more-or-less intermediate cases. The lack of concreteness in these deductions motivates us to consider more quantitative approaches.

Linear discriminant analysis. One way to quantify separability between two classes is known as linear discriminant analysis (LDA).[13] In this approach the class identities of the spectra must be indicated at the outset of the calculation. A single axis is determined which maximizes the separation between the two classes. A limitation of LDA is that the covariance matrix encountered in the calculation must be non-singular. This means that the computation requires at least as many spectra as the dimension of the spectral space. But statisticians warn that a high dimensional covariance matrix (a few hundred if we use the full spectra) is not truly valid unless there are an impractically large number of samples. Since the number of samples is limited, we

chose the alternative of reducing the dimension of the data set. PCA is appropriate for accomplishing this necessary first step (reference 14 offers another approach). Figure 3 shows two examples of LDA applied to a five dimensional data set. The result is essentially identical when a three dimensional data set is used. LDA has further reduced the representation of the data to a single dimension. We see in Fig. 3 that alkanes and alcohols are much more separated in region 1 than in region 2. This degree of separation may be quantified into a *separation metric* by ratioing the absolute value of the difference of the means to the square root of the pooled variance. This approach, of course, implies an assumption of normality, which is not likely to be strictly valid. Nevertheless, the assumption is convenient and probably has little deleterious effect on the validity of our deductions. Figure 4 plots the LDA based separation metric for the five spectral regions and the three pairwise class comparisons (#1 alkanes versus alcohols; #2 alkanes versus ketones/aldehydes; #3 alcohols versus ketones/aldehydes). The numerical data contained in Fig. 4 are also presented in Table 2a. Again region 1 is shown to have the best class separation. We see clearly that regions 3, 4, and 5 have sufficient separation to be considered for sensor applications. Region 2 is the least useful. Even so, region 2 is very popular (perhaps inappropriately so) for the detection of organic compounds.

Other measurements of class separation. Table 2 includes data for three additional separation metrics with no accompanying figure such as Fig. 4. All are calculated after reducing the data set to a dimension of three via PCA. We find that all metrics show similar trends to Fig. 4 and Table 2a.

Table 2b contains results for the first term of the Bhattacharyya distance.[15] We are familiar in one dimension with the approach of normalizing the difference between the class means to the within-class standard deviation or variance (see discussion of Fig. 3, above). The Bhattacharyya distance is the rigorous generalization of this calculation, assuming normality, to higher dimensions. Like LDA it makes use of covariant matrices that must be non-singular.

$$B = \frac{1}{8}(M_2 - M_1)^T \left[\frac{\Sigma_1 + \Sigma_2}{2}\right]^{-1} (M_2 - M_1) \qquad (1)$$

$$+\frac{1}{2}\ln\frac{\left|\frac{\Sigma_1 + \Sigma_2}{2}\right|}{\sqrt{|\Sigma_1||\Sigma_2|}},$$

where the subscripts 1 and 2 designate the classes, superscript T indicates transpose, |--| indicates the determinant, M_i is the mean vector (i.e., spectrum), and Σ_i is the covariance matrix. The estimated covariance matrix is calculated according to:

$$\Sigma_i = \frac{1}{N-1}\sum_{k=1}^{N}(\mathrm{X}_k - M_i)(\mathrm{X}_k - M_i)^T,$$

where X_k is the individual vector representing the kth spectrum. The first term in Eq. 1 is the distance between two classes normalized to their variance, while the second term captures the

component of class separation related to unequal within-class variances. The second term is interesting because it reminds us that classes may have the same mean but be distinguishable to a certain degree by a difference in their distribution. However, in the present application there is a problem with the second term. The samples are too few to provide an estimate of the second term that is different from zero, in the sense of statistical significance. It is better to set the second term to zero (i.e. assume equal variances) than assume that the separability indicated by the second term is real. Hence, we include only the first term of the Bhattacharyya distance in Table 2b. Note that this term is similar to, but not the same as the Mahalanobis distance.[12]

Table 2c contains results for one type of scatter matrix calculation[15] which is defined by the following equation:

$$SM_1 = \ln\left|\frac{S_W}{S_M}\right|, \qquad (2)$$

where S_W is the within-class scatter matrix which is simply the mean of the class covariance matrices, weighted by the number of samples in each class, and S_M is the mixture scatter matrix which is simply the covariance matrix of all the samples regardless of their class assignments.

Table 2d contains results for a second type of scatter matrix calculation[15] which is defined by the following equation:

$$SM_2 = \text{trace}\left(\frac{S_W}{S_M}\right). \qquad (3)$$

The approach in Eqs. 2 and 3 is to ratio the within-class variance to total variance. Then reduce the resulting matrix to a single number by employing a measure of the size of the matrix. Both the determinant and the trace can serve this function. The general trends of Tables 2a and 2b are repeated when Eqs. 2 or 3 are used for the separation metric.

PLS prediction as an alternative approach. The method of PLS[11] is primarily aimed at prediction rather than class separability. However, PLS results may be folded into the present work by focusing on the precision of determining a component from one class when it is present in a large and varying background of the other classes. As the separation metric, we use the inverse of the cross validated standard error of prediction (SEP), which is an estimate of the one sigma precision of the determination. The cross validation procedure consists of constructing a series of predictive models where a single calibration point is excluded from each model. The series is used to predict the concentrations of the excluded points, in turn, as if they were unknown. Finally, the predicted values of the excluded points are compared to their known values. The SEP is the root-mean-square of these differences.

We chose the problem of determining the concentration of hexane in a background of alcohols and ketones/aldehydes. Twenty five randomized backgrounds were created by averaging the spectra of these two classes with a random weighting from 0 to 1. The spectrum of hexane was scaled by a known (i.e., not randomized) factor between 0.01 and 0.1 and added to one of the set of randomized background spectra. A cross validated PLS calibration was done with these 25 simulated samples. The SEPs for the five spectral region were: region 1 = 0.005, region 2 =

0.016, region 3 = 0.008, region 4 = 0.010, and region 5 = 0.016, where the units are fractions of the reference hexane spectrum which is scaled from 0 to 1 absorbance units. Lower SEP values indicate better performance. To compare these results to Table 2 it is convenient to define the separation metric as $(10*SEP)^{-1}$. The PLS results, in terms of this separation metric are listed in the first column of Table 3. The trends are consistent with Table 2. The PLS approach has the desirable feature that the degree to which one region is superior to another is clearly quantified. For example, the expected precision of the determination in region 3 is two times better than in region 2.

Also, included in Table 3 are the results when noise is added to the set of 25 modeled spectra. We see that the separation metric is diminished by noise, but not by the same proportion in each spectral region. This is an important observation for practical engineering decisions. As a matter of related interest, we have found that adding a reasonable amount of spectral noise to the calculations of Table 2 has essentially no effect on the numerical results. This is because the noise induced variance is small compared to the other variances considered in the calculation.

There is a subtlety buried in Table 3 that deserves comment. The SEPs, calculated by the PLS calibration runs, depend on the number of factors retained in the model. In Table 3 the number of factors decreases from 5 to 3 with increasing noise, according to the following rule: that the number of factors is lowered until the SEP for the full-calibration (i.e., no samples left out for cross validation) divided by the SEP for the cross validated calibration is greater than 0.60. This avoids overfitting of the data. The specific number, 0.60, in our experience, is appropriate for about 25 samples. The rigorous justification for this procedure is still a research topic.

Additional comments. Considering Tables 2 and 3, we see that the various types of separation metrics yield the same trends. Some metrics seem to have a non-linear scale and tend to exaggerate the changes from region to region. The general finding is that the separability of classes is greatest in region 1 and least in region 2, with the others being intermediate. One result that goes against the general rule is that the specificity for alcohols is greatest in region 5.

Table 2a has one of the more convenient quantitative interpretations. Class means are different by the tabulated number of units of standard deviation. A three standard deviation separation already gives very little overlap for normally distributed classes. Hence, even the most overlapped of classes in Table 2a are actually well separated. This is a bit counter-intuitive since, in Fig. 1, the classes have quite similar spectra in some regions. Nevertheless, multivariate analysis is capable of making distinctions that are not clearly evident to the eye. Hence, we anticipate that that near-IR spectroscopic sensors will, for most applications, have the specificity required to determine organic compounds in the midst of interfering backgrounds. This assumes that the sensor can be constructed to take advantage of multivariate data analysis and is not limited to one or two wavelengths. With this multivariate approach, building a sensor to function in the fingerprint region (region 1) may not be a good engineering choice. The instrumentation is expensive and the added specificity may not be required. In addition, the CH region (region 2) is usually a poor choice since the specificity is low, yet the cost is high. Our analysis shows that the near-IR spectral region offers reasonable specificity for the type of problem considered. Assuming that the promise of low cost sensors is realized, applications in the near-IR should continue to increase.

ACKNOWLEDGMENTS

This work was supported by the U. S. Department of Energy under contract number DE-AC04-94AL85000. Discussions with Edward V. Thomas, a statistician at Sandia National Laboratories, were beneficial. Also, Jason Harper, a high school intern, contributed by producing a list of compounds which are common to both spectral libraries.

REFERENCES

1. A. Fong and G. M. Hieftje, *Appl. Spectrosc.* **49**, 1261 (1995).

2. I. Schneider, G. N. Trude, V. V. King, and I. D. Aggarwal, *IEEE Photonics Technology Letter* 7, 1041 (1995).

3. A. S. Bonanno and P. R. Griffiths, *Appl. Spectrosc.* **49**, 1598 (1995).

4. B. R. Stallard, M. J. Garcia, and S. Kaushik, *Appl. Spectrosc.* **50**, 334 (1996).

5. I. Murray and I. A. Cowe, edt., *Making Light Work: Advances in Near Infrared Spectroscopy,* (VCH, Weinheim, 1992).

6. P. Williams and K. Norris, edt., *Near-Infrared Technology in the Agricultural and Food Industries,* (American Association of Cereal Chemists, St. Paul, 1987).

7. P. V. Lambeck, *Sensors and Actuators* B **8**, 103 (1992).

8. D. P. Saini and S. L. Coulter, *Photonics Spectra* **30**, 91 (March 1996).

9. E. Stark, *SPIE* **1575**, 70 (1991).

10. M. Buback and H. P. Vögele, *FT-NIR Atlas,* (VCH, Weinheim, 1993). Available in digital form since July 1995 from Chemical Concepts, Weinheim, Germany.

11. H. Martens and T. Naes, *Multivariate Calibration,* (John Wiley & Sons, Chichester, England 1989).

12. D. L. Massart, B. G. M. Vandeginste, S. N. Deming, Y. Michotte, and L. Kaufman, *Chemometrics: a Textbook,* (Elsevier, Amsterdam, 1988).

13. R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis,* (Prentice-Hall, Englewood Cliffs, New Jersey 1992).

14. P. Jonathan, W. V. McCarthy, and A. M. I. Roberts, *J. of Chemometrics* **10**, 189 (1996).

15. K. Fukunaga, *Introduction to Statistical Pattern Recognition,* (Academic Press, San Diego, 1990).

Table 1. The compounds whose spectra were used in this study.

ALKANES	ALCOHOLS	KETONES/ALDEHYDES
2,2,4-Trimethylpentane	2-Methyl-1-propanol	3-Heptanone
Cyclohexane	2-Butanol	2-Heptanone
Hexane	2,4-Dimethyl-3-pentanol	Octanal
Decane	1-Butanol	Hexanal
Dodecane	Methanol	Pentanal
Heptane	2-Methyl-2-propanol	2-Methylbutyraldehyde
Nonane	3--Methyl-1-butanol	3-Pentanone
Undecane	1-Octanol	Cycloheptanone1
Cyclooctane	2-Methyl-1-butanol	1,4-Dimethylcyclohexanone
Methylcyclohexane	2-Octanol	4-Methyl-2-pentanone
Cyclopentane	1-Heptanol	
1,2-Dimethylcyclohexane	Cycloheptanol	
Hexadecane	2-Methylcyclohexanol	
	3,3-Dimethyl-1-butanol	
	1-Hexanol	

Table 2. Summary of results for the various class separation metrics defined in the text. Larger numbers always mean more separation, but the precise nature of the scale may be poorly defined (see text). The classes are compared pairwise where a = alkanes, b = alcohols, and c = ketones/aldehydes.

A. Separation metrics based on LDA.

	Class Pairs		
Spectral Regions	a - b	a - c	b - c
1. 500 - 2500 cm^{-1}	13	13	15
2. 2500 - 3900 cm^{-1}	4.6	2.3	4.5
3. 3800 - 5500 cm^{-1}	6.5	4.3	6.5
4. 5200 - 7200 cm^{-1}	3.3	4.5	6.1
5. 6300 - 10500 cm^{-1}	17	2.8	9.0

B. Separation metrics based on the first term of the Bhattacharyya distance, Eq. 1.

	Class Pairs		
Spectral Regions	a - b	a - c	b - c
1. 500 - 2500 cm^{-1}	11	20	7.2
2. 2500 - 3900 cm^{-1}	2.1	0.5	2.2
3. 3800 - 5500 cm^{-1}	2.7	1.4	1.8
4. 5200 - 7200 cm^{-1}	1.3	2.3	3.8
5. 6300 - 10500 cm^{-1}	14	0.8	9.3

C. Separation metrics based on scatter matrices in Eq. 2.

	Class Pairs		
Spectral Regions	a - b	a - c	b - c
1. 500 - 2500 cm^{-1}	3.3	4.0	2.8
2. 2500 - 3900 cm^{-1}	1.8	0.9	1.7
3. 3800 - 5500 cm^{-1}	2.0	1.5	1.6
4. 5200 - 7200 cm^{-1}	1.4	2.0	2.3
5. 6300 - 10500 cm^{-1}	3.5	1.2	3.0

D. Separation metrics based on scatter matrices in Eq. 3.

	Class Pairs		
Spectral Regions	a - b	a - c	b - c
1. 500 - 2500 cm^{-1}	27	50	17
2. 2500 - 3900 cm^{-1}	7.5	4.4	7.1
3. 3800 - 5500 cm^{-1}	8.9	6.3	6.8
4. 5200 - 7200 cm^{-1}	5.8	8.6	11
5. 6300 - 10500 cm^{-1}	32	5.1	22

Table 3. The class separation metric based on SEPs obtained from PLS analysis on a simulated data set described in the text. The analyte is hexane in a random background of both alcohols and ketone/aldehydes. In all but the first column, noise has been added.

	Percent Added Noise			
Spectral Regions	0	1	2	4
1. 500 - 2500 cm^{-1}	20	20	14	8.3
2. 2500 - 3900 cm^{-1}	6.3	5.5	5.3	5.3
3. 3800 - 5500 cm^{-1}	13	7.7	7.7	4.5
4. 5200 - 7200 cm^{-1}	10	6.3	4.8	4.3
5. 6300 - 10500 cm^{-1}	6.3	6.3	4.8	4.3

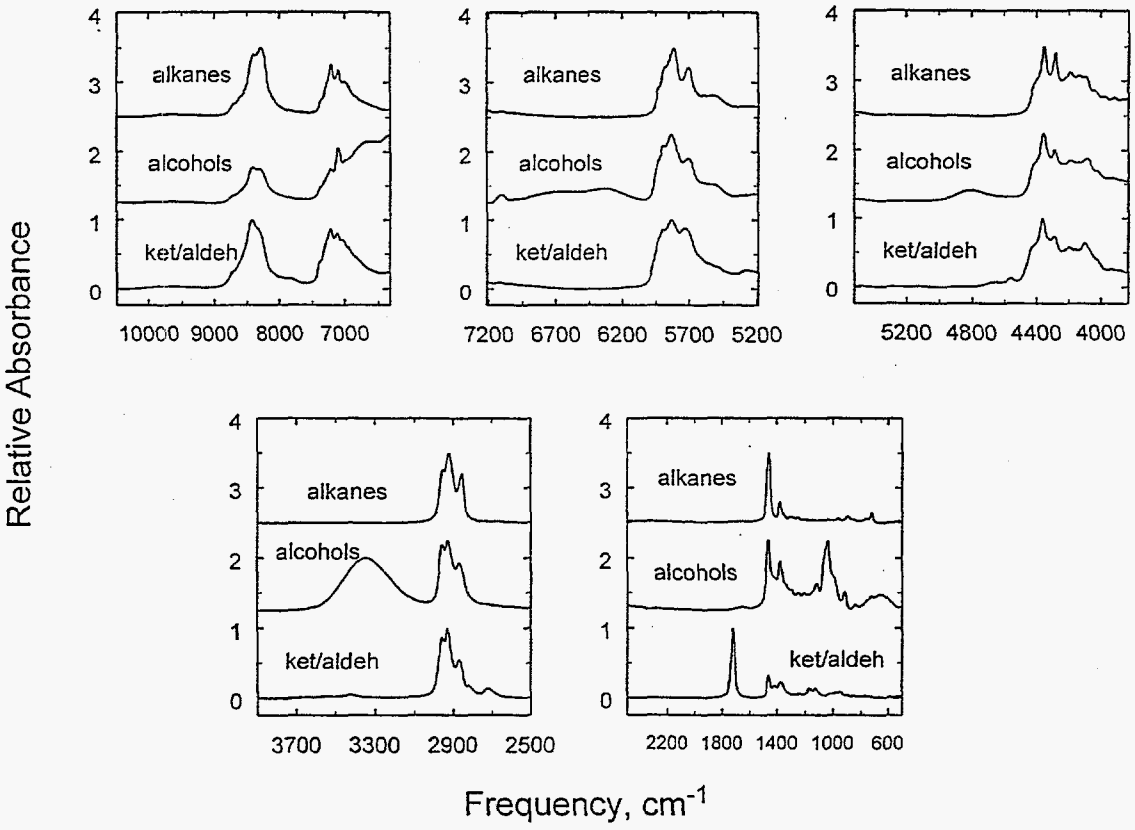


Fig. 1 Plots of the mean spectrum for each class in each spectral region. Spectra have been scaled separately in each region so that the absorbance values are from 0 to 1. Spectra have been offset for clarity.

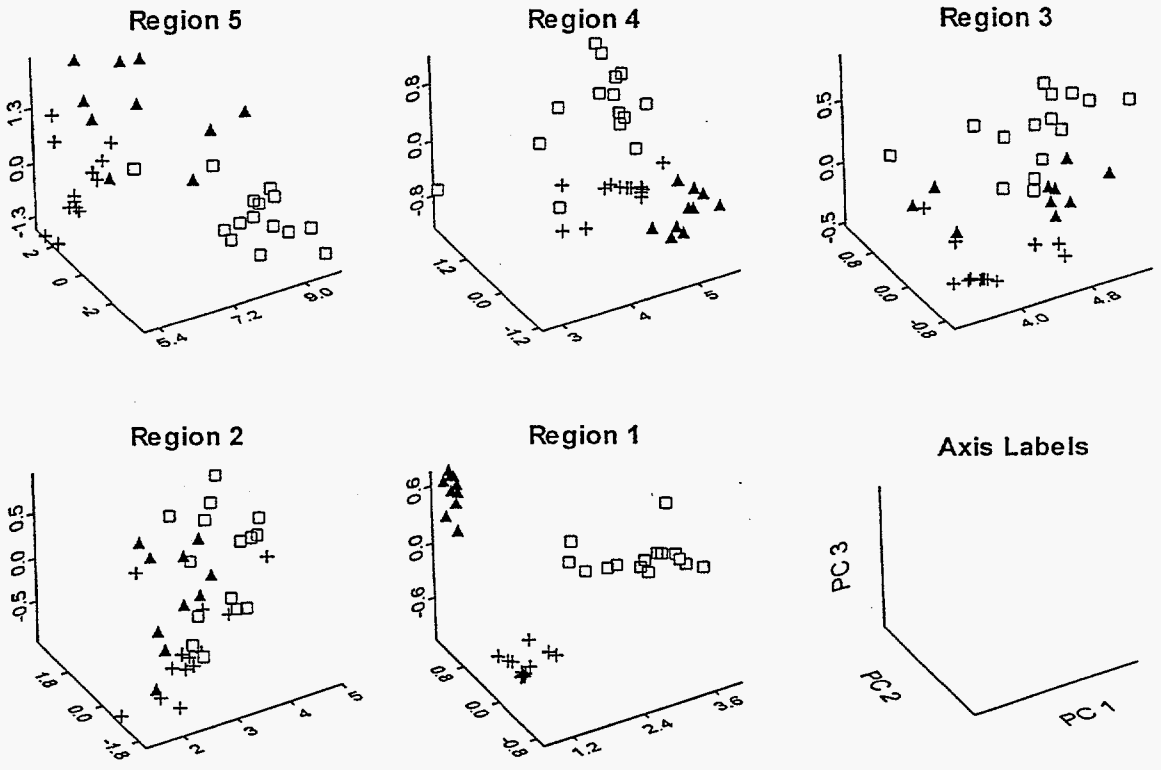


Fig. 2 Three dimensional scatter plots of the scores of the first three PCs in each of the five spectral regions. The regions are specified in the text and have the same position in Fig. 1 as in this figure. Each point represents a spectrum where + = alkane, □ = alcohol, and ▲ = ketone/aldehyde.

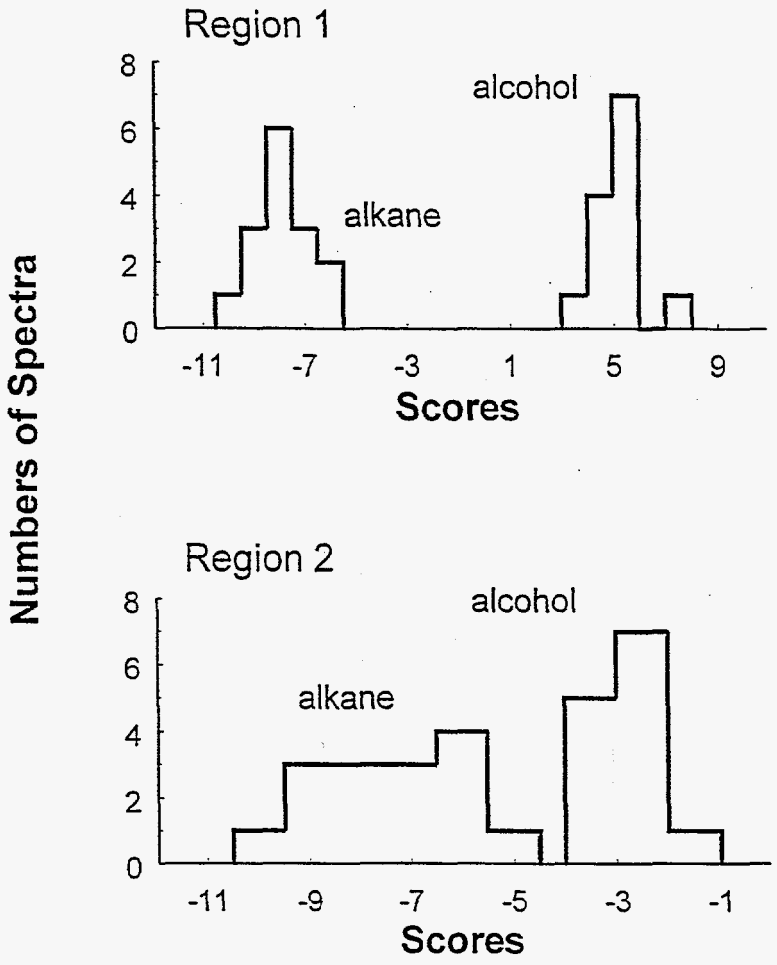


Fig. 3 Two examples of the one-dimensional distribution of scores that are produced through LDA. Complete LDA results are contained in Fig. 4 and Table 2a.

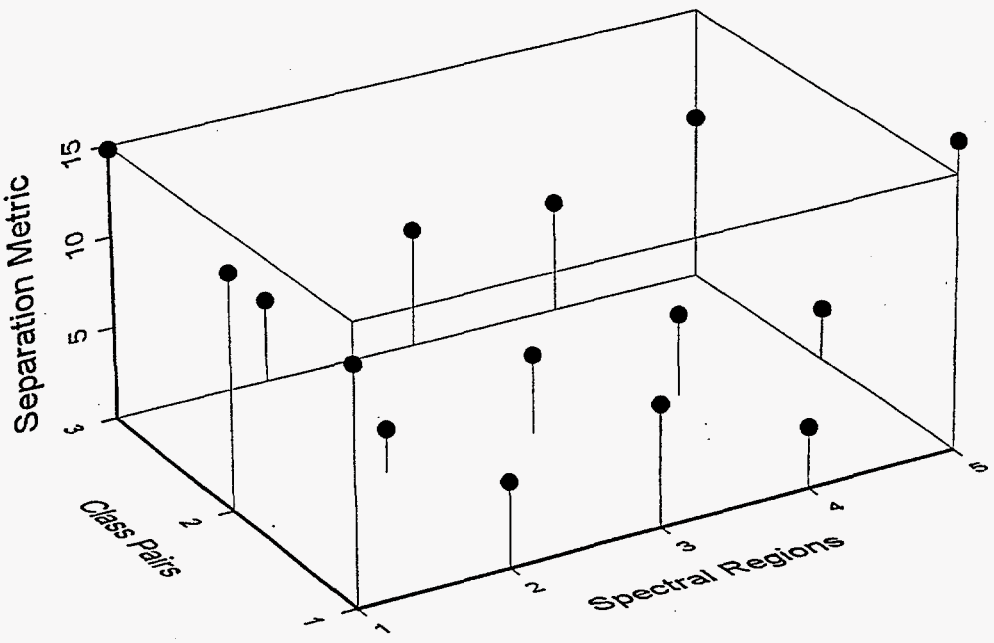


Fig. 4 A three dimensional plot showing the class separation metric data derived from LDA (as explained in the text) for all spectral regions and pairwise class comparisons. The class pair labels have the following meaning: #1 = alkanes versus alcohols, #2 = alkanes versus ketone/aldehyde, and #3 = alcohols versus ketone/aldehydes.

Distribution:

20	MS	0603	I. J. Fritz, 1312
1		0603	J. F. Klem, 1314
1		0603	M. J. Hafich, 1314
1	0820	~~0603~~	H. P. Hjalmarson, ~~1313~~ 9232
1		0603	A. Owyoung, 1304
1		1411	B. R. Stallard, 1823
1		1411	M. J. Garcia, 1823
1		1411	S. Kaushik, 1823
1		9018	Central Technical Files, 8940-2
5		0899	Technical Library, 13414
2		0619	Review and Approval Desk, 12690 for DOE/OSTI