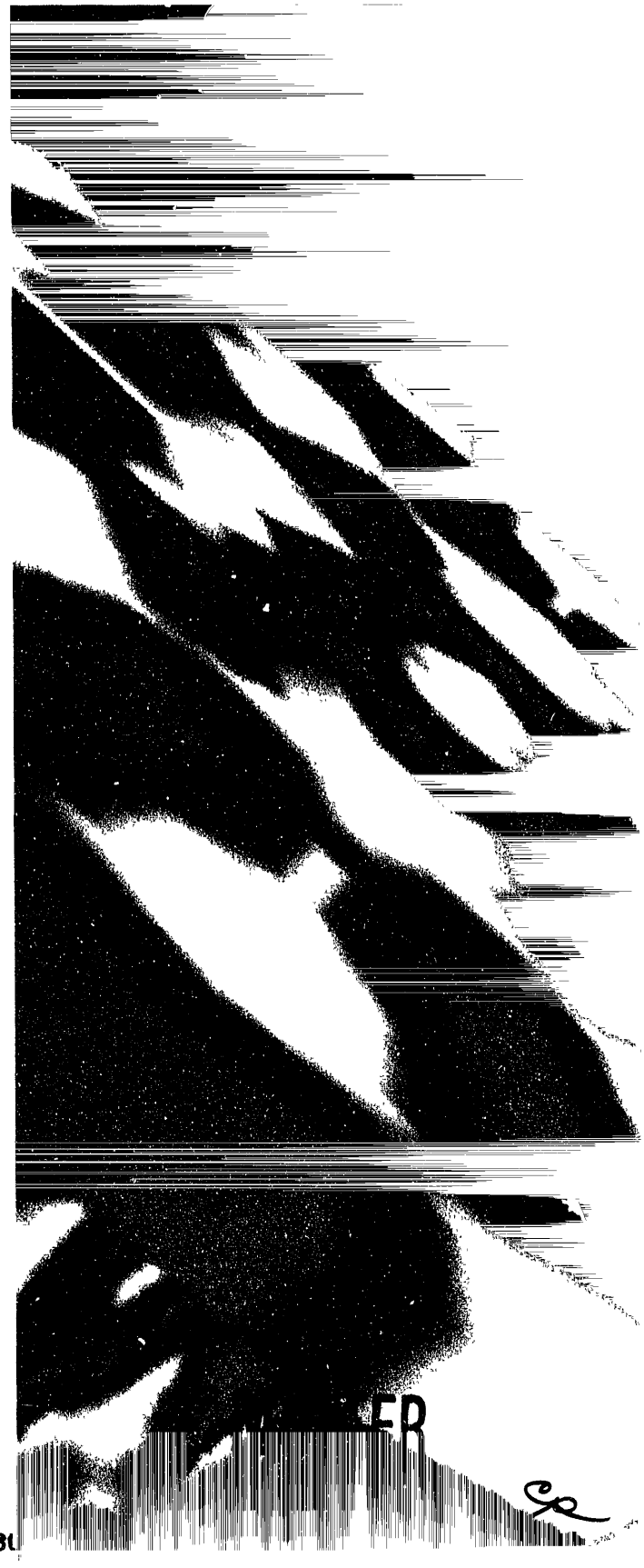
Engineering Research Development and Technology



Thrust Area Report FY92

Manuscript Date March 1993
Distribution Category UC-706
Lawrence Livermore National Laboratory
UCRL 53868-92

Introduction
Roger W. Werne, Associate Director for Engineering and Technology Transfer

1. Computational Electronics and Electromagnetics

Overview
John F. DeFord, Thrust Area Leader

Parallel Computers and Three-Dimensional Computational Electromagnetics
Niel K. Madsen 1-1

Computational Integrated Photonics
Raymond J. Hawkins, Jeffery S. Kallman, and Richard W. Ziolkowski 1-7

Analysis of High-Average-Power, Millimeter-Wave Microwave Components and Induction Linear Accelerator Modules
Clifford C. Shang, John F. DeFord, and Malcolm Caplan 1-13

Electromagnetic Modeling and Experiments for Dispersive Media
Scott D. Nelson and Carlos A. Avalle 1-21

Band Gap Engineering for Infrared Detectors
J. Brian Grant 1-25

2. Computational Mechanics

Overview
Gerald L. Goudreau, Thrust Area Leader

Solution Strategies: New Approaches for Strongly Nonlinear Quasistatic Problems Using DYNA3D
Robert G. Whirley and Bruce E. Engelmann 2-1

Enhanced Enforcement of Mechanical Contact: The Method of Augmented Lagrangians
Bradley N. Maker and Tod A. Laursen 2-7

ParaDyn: New Generation Solid/Structural Mechanics Codes for Massively Parallel Processors
Carol G. Hoover, Anthony J. De Groot, James D. Maltby, and Robert G. Whirley 2-11

Composite Damage Modeling
Edward Zywicz 2-15

HYDRA: A Flow Solver for Three-Dimensional, Transient, Incompressible Viscous Fluid
Mark A. Christon 2-19

Development and Testing of the TRIM3D Radiation Heat Transfer Code
James D. Maltby 2-23

A Methodology for Calculating the Seismic Response of Critical Structures
David B. McCallen, Francois E. Heuze, Lawrence J. Hutchings, and Stephen P. Jarpe 2-27

Reinforced Concrete Damage Modeling
Sanjay Govindjee and Gregory J. Kay 2-35

3. Diagnostics and Microelectronics

Overview
Joseph W. Balch, Thrust Area Leader

Novel Photonic Detectors
Raymond P. Mariella, Jr., Gregory A. Cooper, Sol P. Dijaili, Robert Chow, and Z. Liliental-Weber 3-1

Wideband Phase Modulator
Charles F. McConaghy, Sol P. Dijaili, and Jeffrey D. Morse 3-5

Optoelectronic Terahertz Beam System: Enabling Technologies
Jeffrey D. Morse 3-9

Fabrication of Microelectrode Electrochemical Sensors
Dino R. Ciarlo, Jackson C. Koo, Conrad M. Yu, and Robert S. Glass 3-13

Diamond Heatsinks
Dino R. Ciarlo, Jick H. Yee, Gizzing H. Khanaka, and Erik Randich 3-15

Advanced Micromachining Technologies
Wing C. Hui 3-19

Electrophoresis Using Silicon Microchannels
Jackson C. Koo, J. Courtney Davidson, and Joseph W. Balch 3-21

4. Emerging Technologies

Overview
Shin-yee Lu, Thrust Area Leader

Tire, Accident, Handling, and Roadway Safety
Roger W. Logan 4-1

EXTRANSYT: An Expert System for Advanced Traffic Management
Rowland R. Johnson 4-9

Odin: A High Power, Underwater, Acoustic Transmitter for Surveillance Applications
Terry R. Donich, Scott W. McAllister, and Charles S. Landram 4-13

Passive Seismic Reservoir Monitoring: Signal Processing Innovations
David B. Harris, Robert J. Sherwood, Stephen P. Jarpe, and David C. DeMartini 4-17

Paste Extrudable Explosive Aft Charge for Multi-stage Munitions
Douglas R. Faux and Russell W. Rosinsky 4-21

A Continuum Model for Reinforced Concrete at High Pressures and Strain Rates
Kurt H. Sinz 4-23

Benchmarking of the Criticality Evaluation Code COG
William R. Lloyd, John S. Pearson, and H. Peter Alesso 4-27

Fast Algorithm for Large-Scale Consensus DNA Sequence Assembly
Shin-yee Lu, Elbert W. Branscomb, Michael E. Colvin, and Richard S. Judson 4-29

Using Electrical Heating To Enhance the Extraction of Volatile Organic Compounds from Soil
H. Michael Buettner and William D. Daily 4-31

5. Fabrication Technology

Overview
Kenneth L. Blaedel, Thrust Area Leader

Fabrication of Amorphous Diamond Coatings
Steven Falabella, David M. Sanders, and David B. Boercker 5-1

Laser-Assisted Self-Sputtering
Peter J. Biltoft, Steven Falabella, Steven R. Bryan, Jr., Ralph F. Pombo, and Barry L. Olsen 5-5

Simulation of Diamond Turning of Copper and Silicon Surfaces
David B. Boercker, James Belak, and Irving F. Stowers 5-7

6. Materials Science and Engineering

Overview
Donald R. Lesuer, Thrust Area Leader

Processing and Characterization of Laminated Metal Composites
Chol K. Syn, Donald R. Lesuer, and O.D. Sherby 6-1

Casting Process Modeling
Arthur B. Shapiro 6-7

Characterizing the Failure of Composite Materials
Scott E. Groves, Roberto J. Sanchez, William W. Feng, Albert E. Brown, Steven J. DeTeresa, and Richard E. Lyon 6-11

Fiber-Optic Raman Spectroscopy for Cure Monitoring of Advanced Polymer Composites
Richard E. Lyon, Thomas M. Vess, S. Michael Angel, and M.L. Myrick 6-17

Modeling Superplastic Materials
Donald R. Lesuer, Chol K. Syn, Charles S. Preuss, and Peter J. Raboin 6-23

7. Microwave and Pulsed Power

Overview
E. Karl Freytag, Thrust Area Leader

Pulsed Plasma Processing of Effluent Pollutants and Toxic Chemicals
George E. Vogtlin 7-1

Ground Penetrating Imaging Radar for Bridge Inspection
John P. Warhus, Scott D. Nelson, Jose M. Hernandez, Erik M. Johansson, Hua Lee, and Brett Douglass 7-5

High-Average-Power, Electron Beam-Controlled Switching in Diamond
W. Wayne Hofer, Don R. Kania, Karl H. Schoenbach, Ravindra Joshi, and Ralf P. Brinkmann 7-13

Testing of CFC Replacement Fluids for Arc-Induced Toxic By-Products
W. Ray Cravey, Wayne R. Luedtka, Ruth A. Hawley-Fedder, and Linda Foiles 7-19

Applying Statistical Electromagnetic Theory to Mode Stirred Chamber Measurements
Richard A. Zacharias and Carlos A. Avalle 7-23

Magnetically Delayed Low-Pressure Gas Discharge Switching
Stephen E. Sampayan, Hugh C. Kirbie, Anthony N. Payne, Eugene Lauer, and Donald Prosnitz 7-27

8. Nondestructive Evaluation

Overview
Satish V. Kulkarni, Thrust Area Leader

Fieldable Chemical Sensor Systems
Billy J. McKinley and Fred P. Milanovich 8-1

Computed Tomography
Harry E. Martz, Stephen G. Azevedo, Daniel J. Schneberk, and George P. Roberson 8-5

Laser Generation and Detection of Ultrasonic Energy
Graham H. Thomas 8-23

9. Remote Sensing, Imaging, and Signal Engineering

Overview
James M. Brase, Thrust Area Leader

Vision-Based Grasping for Autonomous Sorting of Unknown Objects
Shin-yee Lu, Robert K. Johnson, and Jose E. Hernandez 9-1

Image-Restoration and Image-Recovery Algorithms
Dennis M. Goodman 9-7

View: A Signal- and Image-Processing System
James M. Brase, Sean K. Lehman, Melvin G. Wieting, Joseph P. Phillips, and Hanna Szoke 9-11

VISION: An Object-Oriented Environment for Computer Vision and Pattern Recognition
Jose E. Hernandez and Michael R. Buhl 9-15

Biomedical Image Processing
Laura N. Mascio 9-21

Multisensor Data Fusion Using Fuzzy Logic
Donald T. Gavel 9-23

Adaptive Optics for Laser Guide Stars
James M. Brase, Kenneth Avicola, Donald T. Gavel, Kenneth E. Waltjen, and Horst D. Bissinger 9-27

Introduction

The mission of the Engineering Research, Development, and Technology Program at Lawrence Livermore National Laboratory (LLNL) is to develop the technical staff and the technology needed to support current and future LLNL programs.

To accomplish this mission, the Engineering Research, Development, and Technology Program has two important goals: (1) to identify key technologies and (2) to conduct high-quality work to enhance our capabilities in these key technologies.

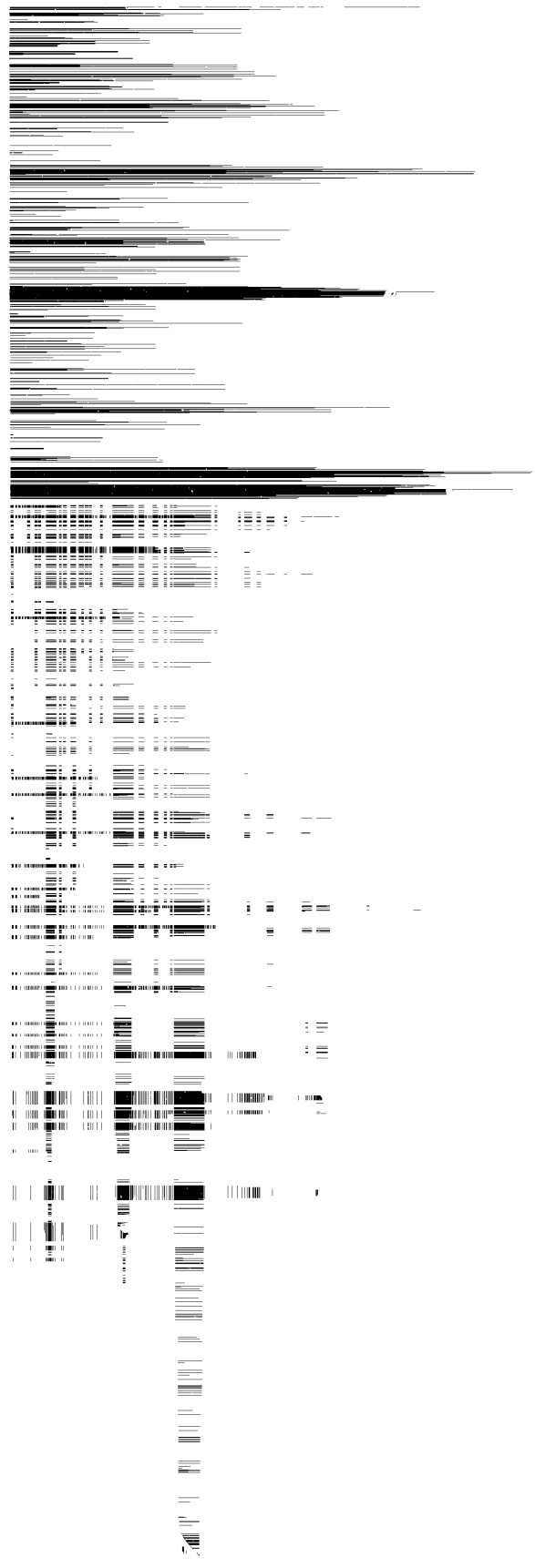
To help focus our efforts, we identify technology *thrust areas* and select technical leaders for each area. The thrust areas are integrated engineering activities and, rather than being based on individual disciplines, they are staffed by personnel from Electronics Engineering, Mechanical Engineering, and other LLNL organizations, as appropriate.

The thrust area leaders are accountable to me for the quality and progress of their activities, but they have sufficient latitude to manage the resources allocated to them. They are expected to establish strong links to LLNL program leaders and to industry; to use outside and inside experts to review the quality and direction of the work; to use university contacts to supplement and complement their efforts; and to be certain that we are not duplicating the work of others. The thrust area leader is also responsible for carrying out the work that follows from the Engineering Research, Development, and Technology Program so that the results can be applied as early as possible to the needs of LLNL programs.

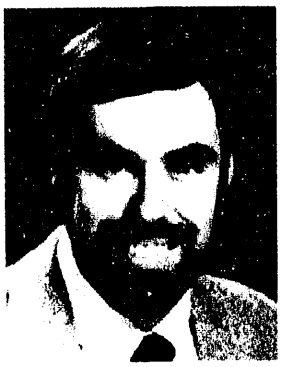
This annual report, organized by thrust area, describes activities conducted within the Program for the fiscal year 1992. Its intent is to provide timely summaries of objectives, theories, methods, and results. The nine thrust areas for this fiscal year are: Computational Electronics and Electromagnetics; Computational Mechanics; Diagnostics and Microelectronics; Emerging Technologies; Fabrication Technology; Materials Science and Engineering; Microwave and Pulsed Power; Nondestructive Evaluation; and Remote Sensing and Imaging, and Signal Engineering.

Readers desiring more information are encouraged to contact the individual thrust area leaders or authors.

Roger W. Werne
Associate Director for Engineering and Technology Transfer



Computational Electronics and Electromagnetics



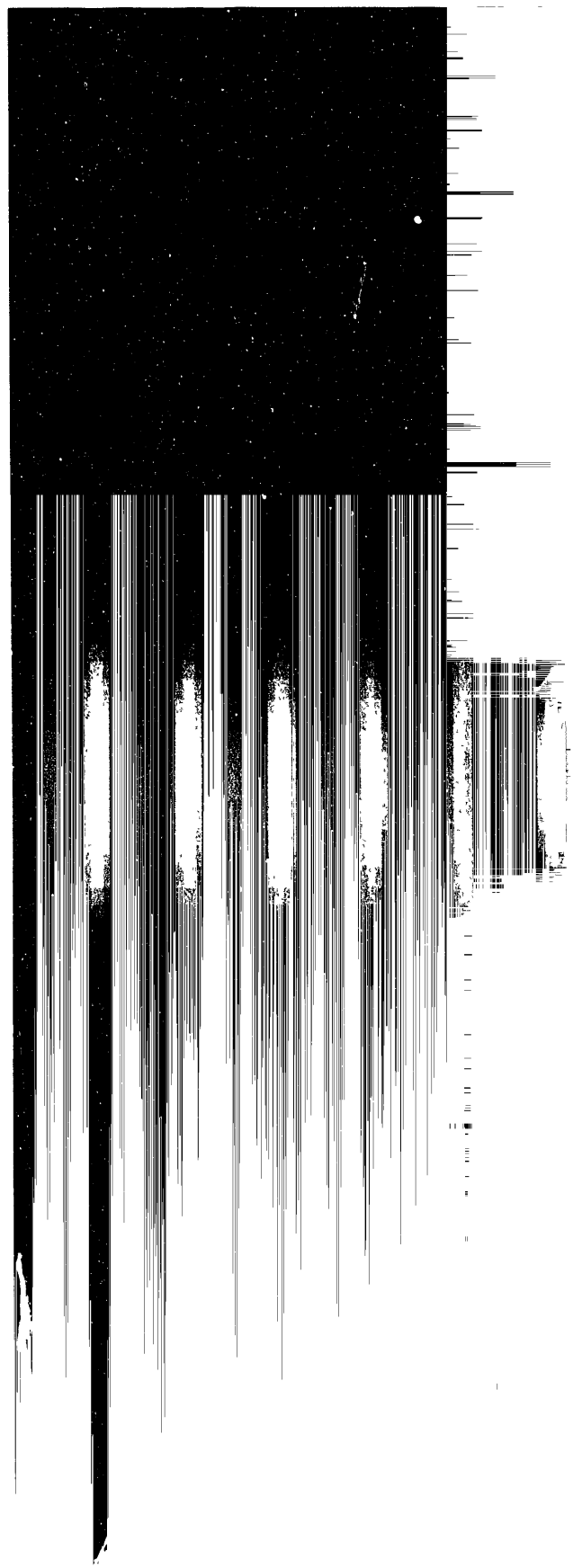
The Computational Electronics and Electromagnetics thrust area is a focal point for computer modeling activities in electronics and electromagnetics in the Electronics Engineering Department of Lawrence Livermore National Laboratory (LLNL). Traditionally, we have focused our efforts in technical areas of importance to existing and developing LLNL programs, and this continues to form the basis for much of our research. A relatively new and increasingly important emphasis for the thrust area is the formation of partnerships with industry and the application of our simulation technology and expertise to the solution of problems faced by industry.

The activities of the thrust area fall into three broad categories: (1) the development of theoretical and computational models of electronic and electromagnetic phenomena, (2) the development of useful and robust software tools based on these models, and (3) the application of these tools to programmatic and industrial problems. In FY-92, we worked on projects in all of the areas outlined above. The object of our work on numerical electromagnetic algorithms continues to be the improvement of time-domain algorithms for electromagnetic simulation on unstructured conforming grids. The thrust area is also investigating various technologies for conforming-grid mesh generation to simplify the application of our advanced field solvers to design problems involving complicated geometries. We are developing a major code suite based on the three-dimensional (3-D), conforming-grid, time-domain code DSI3D. We continue to maintain and distribute the 3-D, finite-difference time-domain (FDTD) code TSAR, which is installed at several dozen university, government, and industry sites. Also, during this past year we have begun to distribute our two-dimensional FDTD accelerator modeling code AMOS, and it is presently being used at several universities and Department of Energy accelerator laboratories. Our principal applications during FY-92 were accelerator components, microwave tubes, photonics, and the evaluation of electromagnetic interference effects in commercial aircraft.

Included in this report are several articles that discuss some of our activities in more detail. The topical areas covered in these articles include computational integrated photonics, the application of massively parallel computers to time-domain modeling, analysis of pulse propagation through concrete for bridge inspection, accelerator component modeling, and the development of tools for semiconductor bandgap calculations.

John F. DeFord
Thrust Area Leader

Section 1



1. Computational Electronics and Electromagnetics

Overview

John F. DeFord, Thrust Area Leader

Parallel Computers and Three-Dimensional Computational Electromagnetics

Niel K. Madsen **1-1**

Computational Integrated Photonics

Raymond J. Hawkins, Jeffery S. Kallman, and Richard W. Ziolkowski **1-7**

Analysis of High-Average-Power, Millimeter-Wave Microwave Components and Induction Linear Accelerator Modules

Clifford C. Shang, John F. DeFord, and Malcolm Caplan **1-13**

Electromagnetic Modeling and Experiments for Dispersive Media

Scott D. Nelson and Carlos A. Avalle **1-21**

Band Gap Engineering for Infrared Detectors

J. Brian Grant **1-25**

Parallel Computers and Three-Dimensional Computational Electromagnetics

Niel K. Madsen
Engineering Research Division
Electronics Engineering

We have continued to make progress in our ability to use massively parallel processing (MPP) computers to solve large, computational, electromagnetics problems. In FY-92, our primary emphasis has been to produce a message-passing version of the preprocessor, PREDSI3D. In addition, the execution module DSI3D has been ported to other parallel machines: the BBN Butterfly, the Thinking Machines CM-5, and the Kendall Square KSR-1 machine. Our DSI3D algorithm and code, together with the ever more capable MPP computers, give us a unique opportunity for significant new contributions to three-dimensional electromagnetic modeling. Two recent applications of DSI3D are presented: (1) full-wave analyses of very-high-frequency optical signals propagating in a weakly guided optical fiber cable; and (2) study of the behavior of whispering-gallery-mode microdisk lasers.

Introduction

The solution of physical problems whose behavior is governed by Maxwell's equations has been of considerable interest for many years. The propagation of electromagnetic (EM) signals, such as microwaves for communication or radar pulses for the detection of aircraft, are two examples of such problems that have been studied over long periods of time. More recently, other areas such as the design of integrated photonics devices; the design and analysis of electronic interconnects for integrated circuits; and the full-wave analysis of microdisk or thumbtack lasers have been studied by numerically solving Maxwell's equations.

The computational tasks for accurately modeling three-dimensional (3-D) problems that are electromagnetically large are very challenging. Two limitations that have been real impediments to the successful solution for these problems are (1) the lack of good, numerical EM algorithms for dealing with problems with complicated, irregular, and nonorthogonal geometries; and (2) the speed and capacity of even the largest and fastest supercomputers.

Present-day supercomputers, such as the Cray-YMP, limit full-wave finite difference or finite volume Maxwell's solutions to computational volumes smaller than about $10^4\lambda^3$, where λ is the wavelength of the EM radiation of interest. For a radar cross-section (RCS) calculation, this limits one to the analysis of scattering from only a small portion of an aircraft fuselage at the upper end of the low-frequency regime, thus neglecting the intra-structure coupling effects that can be important under some conditions. The calculation of the RCS of a complete aircraft, which may be of size $100\,\lambda$ in each of three dimensions, may require as many as 10^9 grid or mesh elements. Problems of this extremely large size clearly will require computers with capabilities that are far beyond those of current supercomputers.

New massively parallel processing (MPP) computers have emerged as the most attractive approach for increasing our computing capabilities to the levels required by large, 3-D, EM simulations. Though still evolving rapidly and not as yet completely viable as production computers, they have demonstrated computational speeds that can no longer be ignored.

With their very distributed nature (memory and CPU's) and lack of sophisticated software development tools, MPP computers present new computing challenges in and of themselves. Large

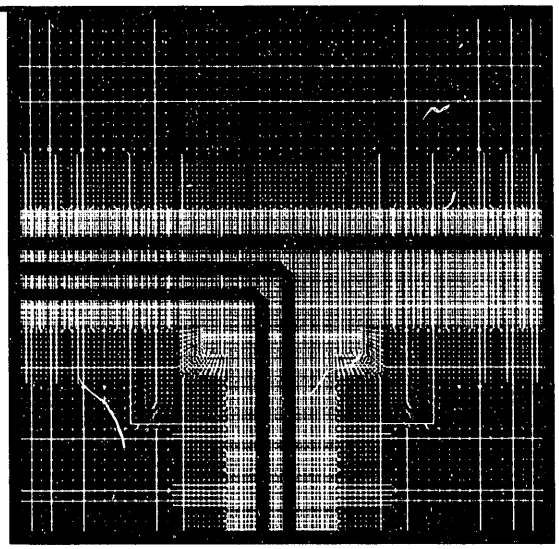


Figure 1. A two-dimensional slice of the 3-D nonorthogonal and unstructured grid used to solve an electronic interconnect problem with three stripline conductors.

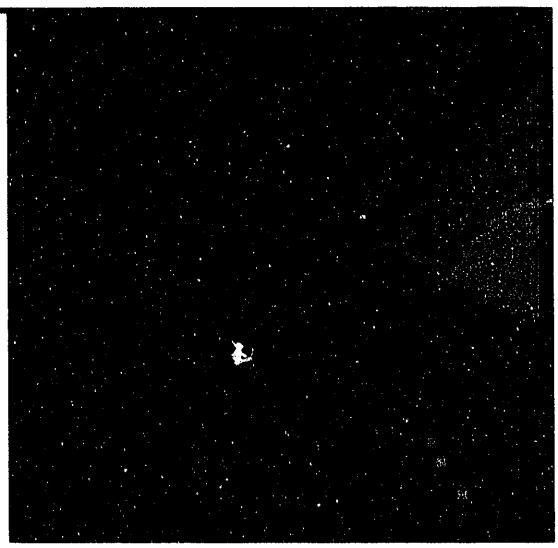


Figure 2. A two-dimensional slice showing the automatic partitioning produced by the recursive spectral bisection method for partitioning the electronic interconnect problem.

simulation problems must be broken into smaller subpieces that can be handled by the individual processors. As a result of this decomposition, data must be efficiently communicated among the processors as required by the numerical algorithm. In the next section, we will describe a new technique that can automatically decompose a problem into smaller subpieces, and also seems to be effective at minimizing the amount of required communication. We will also discuss the performance of our new parallel EM software on one of these newer parallel computers.

Finally, we will show some sample results for two optical applications problems and conclude by indicating our future development directions and thoughts.

Progress

With our development over the past several years of the new discrete surface integral (DSI) methods,[1] the first of the two modeling limitations listed above has been completely overcome, i.e., our new algorithm (implemented in the code DSI3D) has proven to be robust, reliable, and accurate in solving EM problems with complicated and irregular geometries.

The second limitation has been the primary subject of our work for the past two years. Previously, we produced and tested a parallel version of the DSI3D execution module that performs quite efficiently on distributed-memory parallel computers such as the Intel iPSC/860.

Parallel Computation Issues

Recognizing the ultimate physics limitations of trying to speed up traditional serial processing computers, computer manufacturers have begun to design and build MPP computers with hundreds and even thousands of independent processors. These processors are capable of performing hundreds or thousands of arithmetic computations at the same time.

Typically, these MPP computers are distributed-memory computers, i.e., they have very large total amounts of memory, but each processor has rapid direct access to only a small subset of the total memory. For a processor to obtain access to data not residing in its own memory, some form of communication or message passing among processors is required. This distribution of memory presents new complexities when one desires to solve very large problems. Ultimately, the computer's operating system, compilers, and other software tools will automatically take care of these additional complexities. At present, however, all of these software tools are in a state of infancy, and so solving very large EM problems remains a challenging task.

The efficient partitioning or distribution of the computational tasks and data across the computer's memory and processors is an area of high interest. A good partitioning of a problem among multiple processors should satisfy at least two criteria: (1) the partitioning of a problem should produce subpieces of approximately equal size; and (2) the boundaries between the pieces should be as small as possible. The first requirement is

imposed to make sure that each processor has about the same amount of computational work to perform; the second requirement is set to try to minimize the amount of inter-processor data communication.

For problems with a regular structure, it is often the case that an efficient partitioning of the problem is visually obvious. However, for unstructured grids with little predictable structure, a good partitioning is rarely obvious and can present a formidable problem. Last year, we reported our initial experience with a very promising new approach. Others[2] have recently developed a 'recursive spectral bisection' method that seems to meet both of the above criteria. It is based upon the construction of the Laplacian matrix of the dependency graph of the algorithm being used. A dependency graph is produced by linking together variables that depend upon each other, through the underlying solution algorithm (in our case the DSI3D algorithm). The partitioning is accomplished by finding the second eigenvalue of the Laplacian matrix and an associated eigenvector, which is referred to as the 'Fiedler vector.' The median value of the entries of the Fiedler vector is computed, and variables that are associated with Fiedler vector entries that are greater than the median value form one piece of the partition, and those less than the median value form the other partition piece. This process can then be applied recursively to partition the entire problem into the desired number of pieces, which must be an integral power of 2.

We have further tested this new, recursive, spectral bisection technique and have found it to be quite effective at meeting the desired criteria for a good partitioning, even for very large problems. **Figure 1** shows a two-dimensional cross-section of a 3-D, unstructured interconnect grid; **Fig. 2** shows the automatically derived partitioning of this grid cross section into 16 colored subpieces (shades of gray in this rendition).

In addition to the partitioning of an EM problem for MPP solution, there is also the difficulty of producing a version of the unstructured grid code, DSI3D, that runs efficiently on a MPP computer. The primary challenge is to design and implement the passing of data among processors, so that it consumes a small amount of time compared with the time required for computing the solution components. The DSI3D code is really separated into two subpieces: a preprocessing piece, PREDSI3D, which takes the primary grid and the DSI algorithm and derives a dependency graph and update coefficients; and an execution piece, DSI3D,

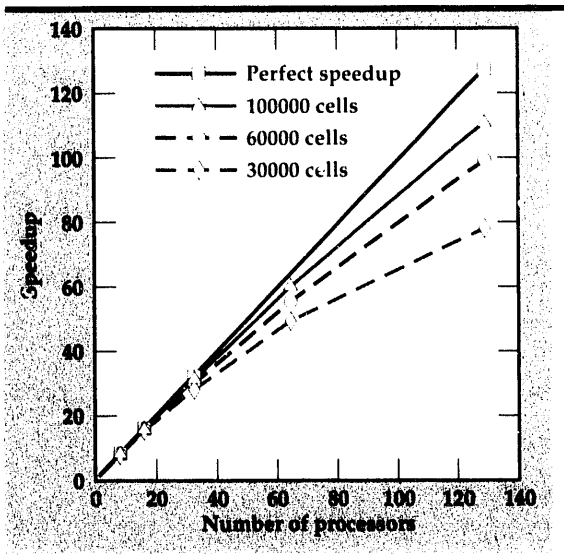


Figure 3. Parallel performance of DSI3D for waveguide problems of three different sizes, using various numbers of processors on the Intel IPSC/860.

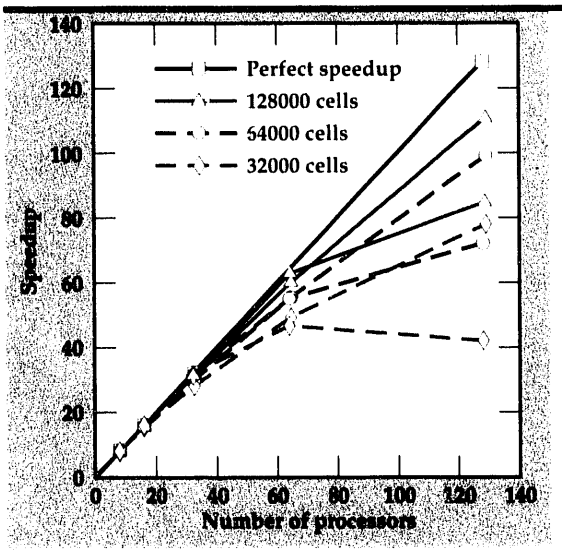


Figure 4. Parallel performance of PREDSI3D for waveguide problems of three different sizes, using various numbers of processors on the Intel IPSC/860.

which repeatedly uses the dependency graph and coefficients to update the field components in a time-marching manner.

In FY-91, we completed the implementation of message-passing versions of DSI3D for the Intel iPSC/860 parallel computer. In FY-92, our primary emphasis has been to produce a message-passing version of the preprocessor, PREDSI3D. In addition, the execution module DSI3D has been ported to other parallel machines: the BBN Butterfly, the Thinking Machines CM-5, and the Kendall Square KSR-1 machine. Generally, we have found that if the problem is sufficiently large, there is considerable benefit to using parallel computers. For smaller problems, it is more efficient to use conventional serial-processing computers. **Figure 3** shows the performance of DSI3D for waveguide propagation problems of three different sizes, using varying numbers of processors on the Intel iPSC/860 parallel computer. **Figure 4** shows the

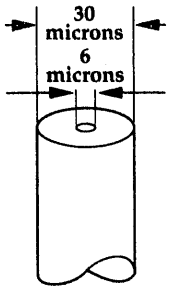


Figure 5. Schematic for a weakly guided fiber optical cable with an offset end-cleave.

performance of PREDSI3D for a similar set of problems. It is clear from these figures that larger problems run uniformly more efficiently than do smaller problems. The preprocessor, PREDSI3D, in general runs less efficiently than the execution module, because it requires considerably more interprocessor communication. However, it is run only once for a particular problem, whereas the execution module may be run repeatedly for the same problem.

Selected Applications

The overall purpose of our work using MPP computers is to be able to easily solve problems that heretofore have not been solvable on conventional serial-processing computers. While the existing MPP computers are not quite at that level of capability, the next generation will be, and we are now ready to address this type of problem.

One area of new interest to us has been the use of DSI3D to perform full-wave analyses of very-high-frequency optical signals propagating in weakly guided optical fiber cables. In splicing optical fibers, it is desired to cleave (or cut) them in a manner so that the cleave is tent-shaped and centered with respect to the fiber core (see **Fig. 5**). Due to their small size, it is not always easy to determine if the cleave is appropriately centered. One idea for determining if the proper centering exists is to launch optical signals in the fiber cable toward the cleave, and then to analyze the signal reflected from the cleaved end back down the cable. If the cleave is centered, most of the reflected energy should remain in the cable core in the fundamental mode. If the cleave is offset significantly, much of the reflected energy will be reflected out of the cable's core. DSI3D is well suited for studying this type of problem. The cleave intersection with the cylindrical fiber is easily handled using the unstructured and nonorthogonal grid features of DSI3D. **Figures 6** and **7** show the nature of the reflected pulses for a centered cleave and an offset cleave, respectively. The differences between the two reflected signals are obvious.

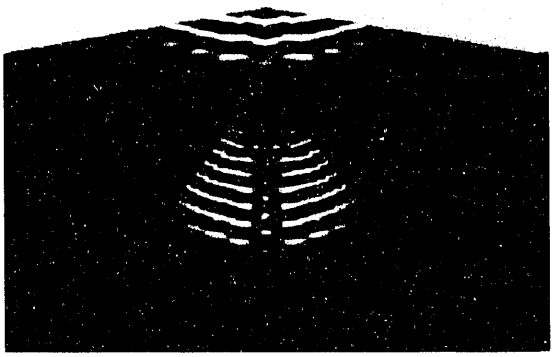


Figure 6. Two-dimensional planar cut in the center of a fiber optical cable with a centered end-cleave, showing the reflected pulse field fringes.

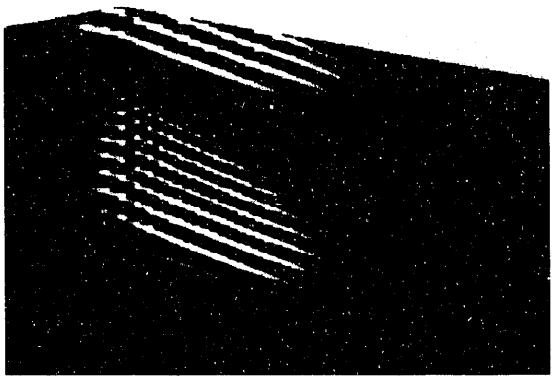


Figure 7. Two-dimensional planar cut in the center of a fiber optical cable with an offset end-cleave, showing the reflected pulse field fringes.

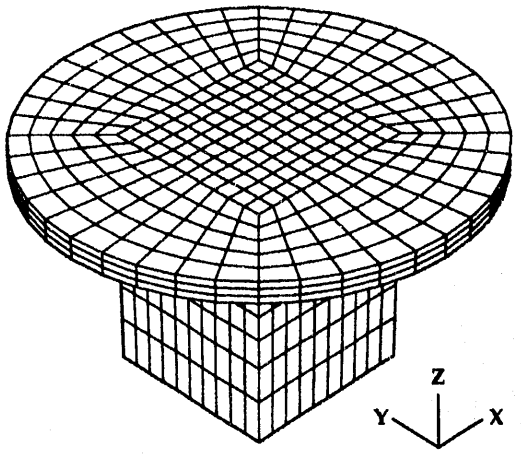


Figure 8. DSI3D grid used to model the behavior of the microdisk laser and pedestal.

Another new interesting application has been the use of DSI3D to study the behavior of whispering-gallery-mode microdisk lasers.[3] These novel devices have potential for the integrability and low-power operation required for large-scale photonic circuits. The disks are formed using selective etching techniques in a InP/InGaAsP system to achieve 3- to 10-μm-dia disks as thin as 500 Å suspended in air or SiO_2 on an InP pedestal. Optical confinement within the thin disk plane results in a microresonator with potential for single-mode, ultra-low threshold lasers. **Figure 8** shows the

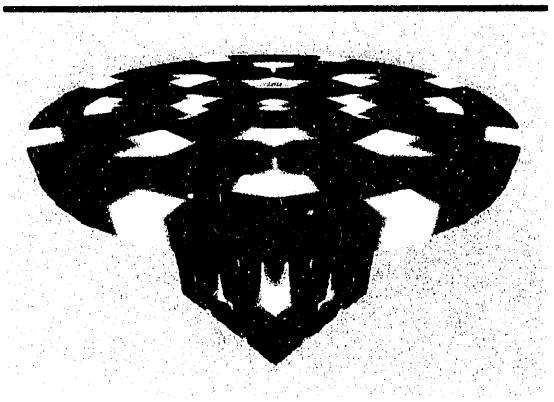


Figure 9. Field plot showing the M = 8 mode for the microdisk laser and pedestal.

DSI3D grid for the disk and pedestal. The grid for the surrounding medium is not shown. **Figure 9** shows the excited M = 8 mode for the disk and pedestal. **Figure 10** shows the structure of the radiated fields in the plane of the disk.

Future Work

Our DSI3D algorithm and code, together with the ever-more-capable MPP computers, give us a unique opportunity for significant new contributions to 3-D EM modeling. We now have the flexibility and capability to solve problems of a size and detail that were previously unimaginable. We intend to address to a much greater extent some of the areas of application mentioned above. In addition, we plan to add a charged-particle capability to DSI3D, so that these new capabilities will be available to the plasma physics community.

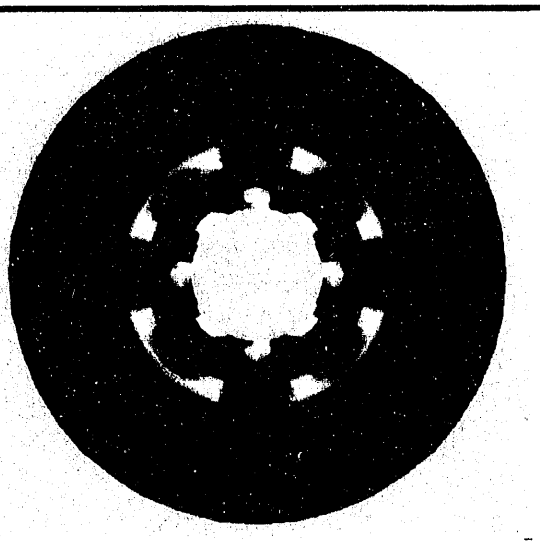


Figure 10. Field plot showing the radiated fields from the microdisk laser and pedestal structure. Plot shows fields in the center plane of the microdisk exterior to the disk.

We have attracted the interest of several industrial partners, and Cooperative Research and Development Agreement efforts are underway with these partners to develop specialized versions of DSI3D for use in RCS analysis and for gyrotron design.

1. N.K. Madsen, *Divergence Preserving Discrete Surface Integral Methods for Maxwell's Curl Equations Using Non-Orthogonal Unstructured Grids*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-109787 (1992).
2. H.D. Simon, *Computing Systems in Engineering* **2** (2/3), 135 (1991).
3. S. McCall, A. Levi, R. Slusher, S. Pearton, and R. Logan, *Appl. Phys. Lett.* **60** (3), 289 (1992).

Computational Integrated Photonics

Raymond J. Hawkins and Jeffrey S. Kallman
Engineering Research Division
Electronics Engineering

Richard W. Ziolkowski
Department of Electrical and Computer Engineering
University of Arizona
Tucson, Arizona

We have continued our innovative work in computational integrated optics, a field important both to programs at Lawrence Livermore National Laboratory (LLNL) and to industry. Integrated optical device design has been our primary research topic. The results of this project have been applied to device design at LLNL, at Bellcore in Red Bank, New Jersey, and at Hughes. A second leading project, device design code integration and graphical user interface development, has also proved to be of great significance, with our simulation results proving to be of interest to a number of companies.

Introduction

Computational integrated photonics (CIP) is the area of computational physics that studies the propagation of light in optical fibers and in integrated optical circuits (the photonics equivalent of electronic circuits). The purpose of integrated photonics simulation is to develop the computational tools that will support the design of photonic and optoelectronic integrated devices. These devices will form the basis of all future high-speed and high-bandwidth information-processing systems and are key to the future industrial competitiveness of the U.S. CIP has, in general, two thrusts: (1) to develop predictive models of photonic device behavior that can be used reliably to enhance significantly the speed with which designs are optimized for applications, and (2) to further our ability to describe the linear and nonlinear processes that occur and can be exploited in real photonic devices.

Our efforts in FY-92 have been focused in three general areas: (1) pseudospectral optical propagation codes; (2) linear finite-difference time-domain (FDTD) codes; and (3) nonlinear FDTD codes. This year we have focused on both the development of codes of interest to the integrated optics community, and on packaging these codes in a user-friendly manner, so that they can be used by other researchers in both academic and industrial laboratories. We have developed two new design codes, BEEMER and TSARLITE, with graphical user interface (GUI), and have made significant advances in nonlinear FDTD.

As FTDT becomes increasingly popular for the study of integrated optical systems, the need to include material dispersion and nonlinear effects has forced us to examine these issues. We found a particularly convenient way of including linear material dispersion in FDTD calculations, and have funded studies in the inclusion of material nonlinearities in FDTD calculations.

Progress

Our work in integrated optical device design continues to give us our leading role in the design of integrated optical components both for Lawrence Livermore National Laboratory programs and for U.S. industry. This research is of particular interest, since we have predictive codes that significantly reduce the time required to bring a device from concept to prototype.

Our work with the pseudospectral optical propagation code, called the beam propagation method (BPM), has addressed the issue of understanding optical field evolution in multilayer, integrated guided-wave detector structures. This work, which previously led to the development of extremely short integrated waveguide/photodiodes with high quantum efficiency, has now resulted in the development of the polarization diversity detector shown in **Fig. 1**[1] and the coherent receiver shown in **Fig. 2**.[2]

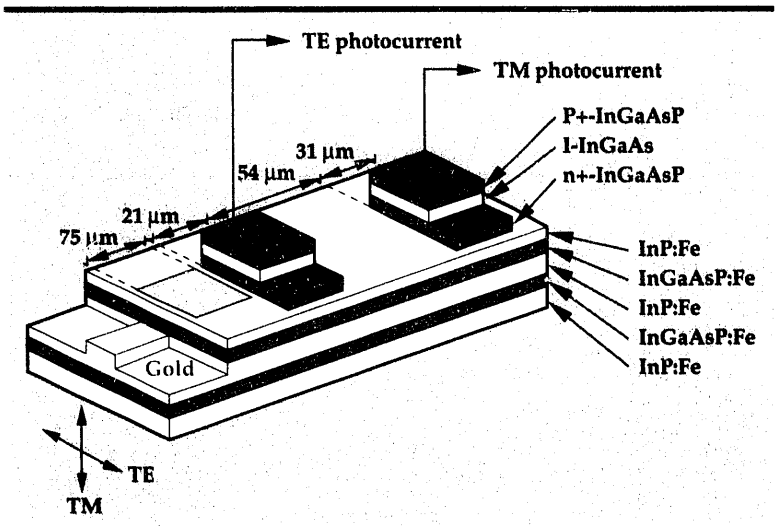


Figure 1. The Bellcore polarization-diversity photodetector that produces two photocurrent outputs proportional to guide-input intensities in each of two orthogonal polarization states. Our device is significantly (a factor of 5 to 20) smaller than previous monolithic realizations of this circuit.

We were also able to help researchers at Hughes understand the operation of their photodetectors, since they were based on a very similar design. Our work with Bellcore was selected as an example of leading work in optical interconnections.[3]

For several years, BPM has been the method of choice for computational physicists studying integrated optical waveguide/device behavior. Unfortunately, the special methods underlying the BPM that made it so efficient also made it difficult for many to code from scratch. Distribution of source code was found, empirically, to be an unsatisfactory alternative, since most codes are written by and for computational physicists. Consequently, the BPM was often admired from afar by those who would best benefit from a hands-on capability.

To fill the void, we wrote BEEMER, a BPM code with a GUI that allows construction of a device layout, simulation, and optimization, all within the same window structure. The designer can specify a variety of material parameters including gain, loss, and Kerr nonlinearity. Thus, this tool can easily handle design problems from linear photodetectors to all-optical soliton-based switches. BEEMER is written in C and has been compiled successfully on a number of workstations, including SUN, IBM, DEC, and SGI. An illustration of the type of problem that BEEMER can handle is shown in **Fig. 3**. The manual for BEEMER guides the user through a number of examples drawn from various areas in optics, to acquaint the user with the program. BEEMER and the manual have been released for distribution outside of LLNL, and we have installed BEEMER at both academic and industrial sites.

To meet the needs of a variety of photonics device designs, we have continued our development of FDTD as a tool for integrated optical device simulation, extending our previous expertise in pseudospectral-code-based device simulation. Our FDTD work has provided information on a variety of devices that could not be modeled by any existing codes. For example, we have demonstrated the ability to model diffraction gratings and facet reflections. The FDTD treatment of electromagnetic pulse propagation holds much prom-

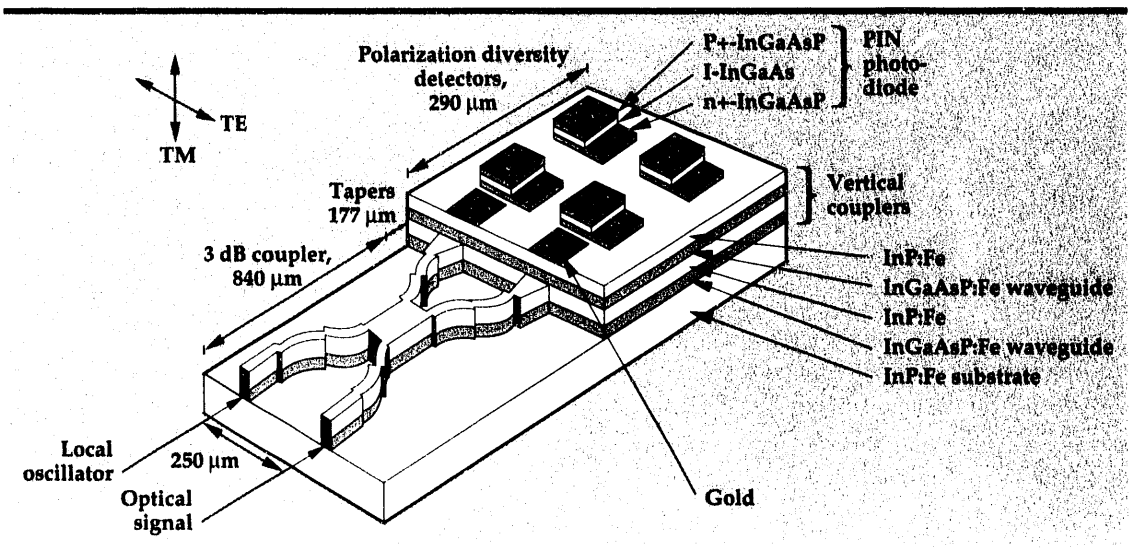


Figure 2. The Bellcore ultracompact, balanced, polarization diversity photodetector. The two detectors corresponding to each polarization state are interconnected for on-chip photocurrent subtraction, which is essential for broadband, balanced operation without microwave hybrids.

ise for the complete numerical description of integrated optical device behavior, where reflections and/or coherent effects are important. The recent application of FDTD to problems in integrated optics[4,5,6] has indicated that electronic dispersion must be included to treat realistically the broadband behavior of integrated optical devices. The inclusion of material dispersion (electronic or magnetic) in FDTD calculations has historically been quite limited. The first formulation of broadband dispersion in FDTD was presented in a pioneering paper[7] that demonstrated that if the electronic susceptibility was expanded as a series of exponentials, the treatment of dispersion could be reduced to a recursive update. The incorporation of this update, however, requires a substantial rewriting of the standard electric-field update equations.

More recently, others[8,9] demonstrated a different formulation of the linear problem, explicitly solving the equation of motion for the polarizability using finite-differencing. This alternative formulation has been extended[10] to nonlinear optical propagation. In our work, we have exploited a simple causality argument that enabled us to write dispersion as a simple, recursive, additive term in the common electric-field update equations. This is of particular interest, since it enabled the treatment of dispersion in a large number of existing FDTD design codes with minimal computational modification.

While there has been a great increase in integrated optical devices for which only a solution of Maxwell's curl equations will suffice, ease of use has not been the hallmark of these codes. To meet this need and to provide ease of user access, we have written TSARLITE[9]: a two-dimensional FDTD code with a fully integrated GUI. [TSAR is an established, three-dimensional (3-D), FDTD code with limited GUI application.] TSARLITE has been constructed with the integrated optics community in mind, and thus has desirable features such as the ability to launch spatially and temporally shaped pulses, and our latest dispersion model. An example of the type of problem that TSARLITE can handle is shown in **Fig. 4**. Unlike BEEMER, TSARLITE does not yet have a manual and has not yet been released for use outside of LLNL. We anticipate that this will happen in the coming year.

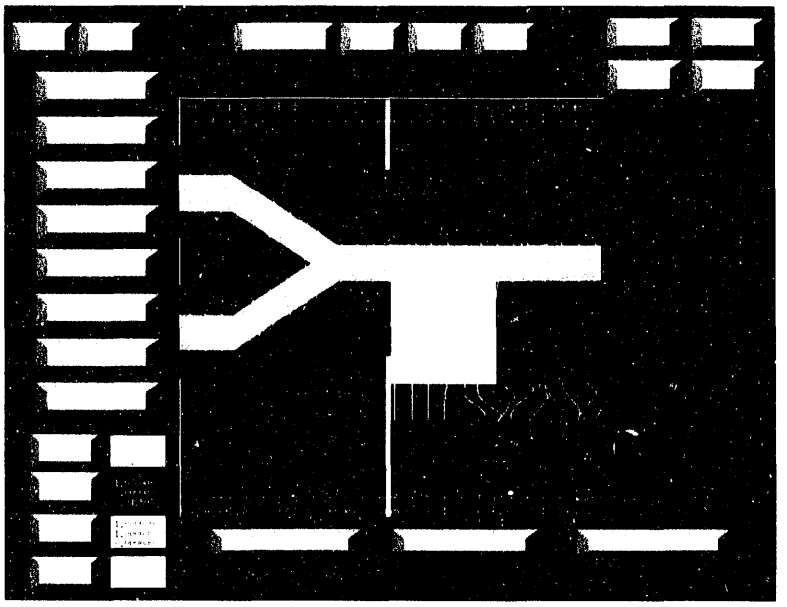


Figure 3. An all-optical switch based on spatial solitons. Light coming in from the left is combined into a waveguide that is placed next to a nonlinear medium. If the combined intensity is great enough (as shown), then the evanescent field in the nonlinear medium is strong enough to form a spatial soliton that splits off and is subsequently captured by the lower arm.

With the continuing and heightened interest in nonlinear semiconductor and optically integrated

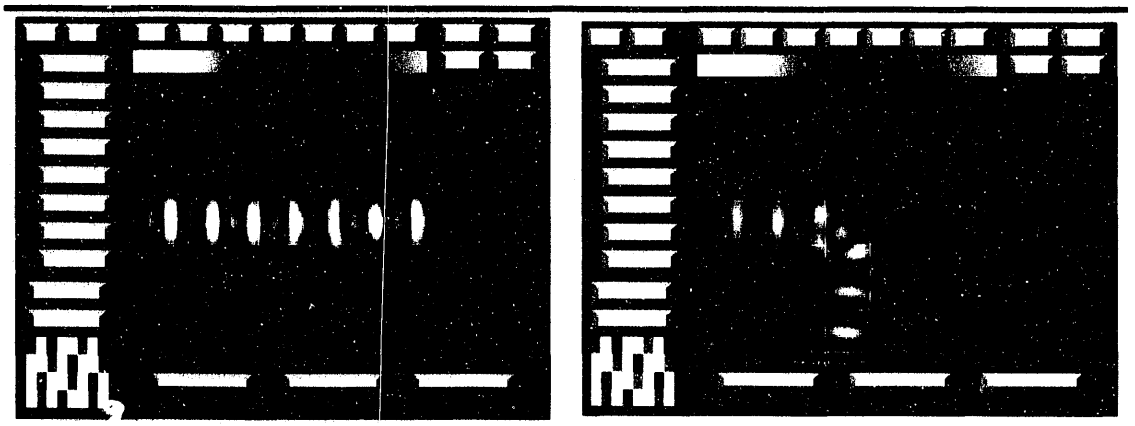


Figure 4. An optical crossbar element. On the left, the element is in transmit mode, but the degree of confinement of the light in the waveguide leads to significant loss in the cross region. With the mirror in place (right), the light is coupled into the waveguide, but the offset of the mirror from an ideal position results in some scattering losses.

devices, more accurate and realistic numerical simulations of these devices and systems are in demand. To date, most of the modeling of pulse propagation in and scattering from nonlinear media has been accomplished with one-dimensional, scalar models. These models have become quite sophisticated; they have predicted and explained many of the nonlinear as well as linear effects in present devices and systems. Unfortunately, they cannot be used to explain many observed phenomena, and are probably not adequately modeling linear and nonlinear phenomena that could lead to new effects and devices. Vector and higher dimensional properties of Maxwell's equations that are not currently included either in existing scalar models or in more detailed material models, may significantly impact the scientific and engineering results. Moreover, because they are limited to simpler geometries, current modeling capabilities are not adequate for linear/nonlinear optical-component engineering design studies. The successful development of general, linear, and nonlinear electromagnetic modeling capabilities will significantly impact the concept and design stages associated with novel linear and nonlinear phenomena and the resulting optical components.

We have developed the first multi-dimensional, full-wave, vector solutions to Maxwell's equations for problems describing the interaction of ultra-short, pulsed beams with a nonlinear Kerr material having a finite response time.[12] These solutions have been obtained with a nonlinear finite-difference time-domain (NL-FDTD) method developed by investigators at the University of Arizona. This NL-FDTD method combines a nonlinear generalization of a standard, FDTD, full-wave, vector, linear Maxwell's equation solver, with a currently used phenomenological time relaxation (Debye) model of a nonlinear Kerr material. In contrast to a number of recently reported numerical solutions of the full-wave, vector, time-independent Maxwell's equations and of vector paraxial equations, the FDTD approach is a time-dependent analysis that accounts for the complete time evolution of the system, with no envelope approximations. Nonlinear, self-focusing numerical solutions in two space dimensions and time that are obtained with this NL-FDTD method, as well as related NL-FDTD results for normal and oblique incidence, nonlinear interface problems, have been investigated. Although these basic geometries are straightforward, the NL-FDTD approach can readily handle very complex, realistic structures.

The chosen sample TE and TM nonlinear optics problems highlight the differences between the scalar and the vector approaches, and the effects of the finite response time of the medium. The NL-FDTD method is beginning to resolve several very basic physics and engineering issues concerning the behavior of the full electromagnetic field during its interaction with a self-focusing medium. In particular, using the NL-FDTD approach we have (1) shown the first back reflections from the nonlinear self-focus; (2) discovered optical vortices formed in the trailing wakefield behind the nonlinear self-focus; (3) identified that the longitudinal field component plays a significant role in limiting the self-focusing process; (4) performed the first complete full-wave, vector treatment of both the TM and TE models of an optical diode (linear/nonlinear interface switch); (5) characterized the performance of an optical diode to single-cycle pulsed Gaussian beams, including the appearance of a nonlinear Goos–Hänchen effect, the stimulation of stable surface modes, and the effects of a finite response time of the Kerr material; (6) shown definitively that the linear/nonlinear interface does not act like an optical diode for a tightly focused, single-cycle pulsed Gaussian beam; and (7) characterized the performance of some basic linear/nonlinear slab waveguides as optical threshold devices.

In all of these analyses, we have identified the role of the longitudinal field component and the resulting transverse power flows in the associated scattering/coupling processes.

Future Work

We will continue our efforts in the design of novel integrated optical devices, both for LLNL programs and for industry. It is our intention to transfer BEEMER and TSARLITE to industry. Our development of linear FDTD applications to integrated optics will be extended to 3-D structures, and our studies of NL-FDTD will continue in the area of nonlinear waveguides and couplers.

1. R.J. Deri, R.J. Hawkins, E.C.M. Pennings, C. Caneau, and N.C. Andreadakis, *Appl. Phys. Lett.* **59** (15), 1823 (1991).

2. R.J. Deri, E.C.M. Pennings, A. Scherer, A.S. Gozdz, C. Caneau, N.C. Andreadakis, V. Shah, L. Curtis, R.J. Hawkins, J.B.D. Soole, and J.-I. Song, *Photonics Tech. Lett.* **22**, 1238 (1992).

3. R.J. Deri, E.C.M. Pennings, and R.J. Hawkins, *Optics and Photonics News* (December, 1991).

4. S.T. Chu and S.K. Chaudhuri, *J. Lightwave Technol.* **LT-7**, 2033 (1989).

5. S.T. Chu, *Modelling of Guided-Wave Optical Structures by the FDTD Method*, Ph.D. Thesis, University of Waterloo (1990).

6. S.T. Chu and S. Chaudhuri, *IEEE Trans. Microwave Theory Tech.* **38**, 1755 (1990).

7. R. Luebbers, F.P. Hunsberger, K.S. Kunz, R.B. Standler, and M. Schneider, *IEEE Trans. Electromagn. Compat.* **EMC-32**, 222 (1990).

8. C.F. Lee, R.T Shin, and J.A. Kong, *PIER4 Progress in Electromagnetics Research*, J.A. Kong (Ed.), Elsevier Science Publishing Company, Inc. (New York), 415, 1991.

9. R.M. Joseph, S.C. Hagness, and A. Taflove, *Opt. Lett.* **16**, 1412 (1991).

10. P.M. Goorjian and A. Taflove, *Opt. Lett.* **17**, 1412 (1992).

11. R.W. Ziolkowski and J.B. Judkins, "Propagation Characteristics of Ultra-Wide Bandwidth Pulsed Gaussian Beams," accepted for publication in *JOSA A* (November 1992).

12. R.W. Ziolkowski and J.B. Judkins, "Full-Wave Vector Maxwell Equation Modeling of the Self-Focusing of Ultrashort Optical Pulses in a Nonlinear Kerr Medium Exhibiting a Finite Response Time," *JOSA B* **10**, 186 (1993).

Analysis of High-Average-Power, Millimeter-Wave Microwave Components and Induction Linear Accelerator Modules

Clifford C. Shang and John F. DeFord
Engineering Research Division
Electronics Engineering

Malcolm Caplan
Magnetic Fusion Energy Program

In FY-92, we analyzed high-average-power, millimeter-wave microwave components in rf systems for heating fusion plasmas and induction linear accelerator modules for heavy ion fusion. The electrical properties of these structures were studied using time-dependent electromagnetic field codes and detailed material models.

We modeled gyrotron windows and gyrotron amplifier sever structures for transverse electric modes in the 100- to 150-GHz range, and computed the reflection and transmission characteristics from the field data. Good agreement between frequency domain codes and analytic results has been obtained for some simple geometries. We describe results for realistic structures with lossy dielectrics and the implementation of microwave diagnostics.

For the modeling of induction accelerators (electron machines), understanding the coupling of the beam to the cavity is of fundamental importance in estimating the effects of transverse beam instabilities. Our accelerator modeling work focused on examining the beam-cavity interaction impedances (impulse response of cavity) for subrelativistic beams in drivers for heavy ion fusion, to better understand longitudinal (n = 0, monopole) and transverse (n = 1, quadrupole) beam instabilities. Results for simple segmented cell configurations show that the pulse power system and induction cores are largely decoupled from wakefields.

Introduction

Robust algorithms for the solution of Maxwell's equations in the time domain have been known for some time.[1,2] Since 1966, specializations of these algorithms to include more sophisticated boundary conditions[3,4] and detailed material models[5,6] have allowed the application of the basic numerical techniques to interesting problems. Further, recent algorithm developments[7,8] for Maxwell solvers on conforming meshes now allow high geometrical fidelity that may be required for a certain class of problems.

Progress

In FY-92, we examined two sets of problems. The first set involves high-average-power millimeter-wave (mmw) structures; the second involves induction linear accelerator cells. The principle features in modeling the mmw structures are the launching of modes, the modeling of lossy dielectrics, and the development of microwave diagnostics. The fundamental aspects of modeling the heavy-ion induction cells include implementing realistic, magnetically dispersive material models and computing subrelativistic wake potentials.[9]

Modeling mmw Components

The use of high-power microwaves to heat the plasma in a magnetic fusion energy (MFE) reactor at the electron-cyclotron resonance can yield a number of benefits, such as bulk-heating and preionization of the plasma; reaction startup; and instability suppression. The use of electron-cyclo-

tron heating (ECH) in tokamak and stellerator reactors has been studied in many significant MFE experiments, including C-mod at the Massachusetts Institute of Technology and DIII-D at General Atomics in the U.S.; Compass at Culham, England; T-10 at the Kurchatov Institute, Russia; and the Heliotron at Nagoya, Japan.

Operating parameters of interest for ECH applications include frequencies in the 140- to 250-GHz range and output power in the vicinity of 1 MW per bottle.[10] Currently, the fixed-frequency mmw source available for use in the 1-MW range is the gyrotron. Understanding the microwave properties of high-average-power rf structures is crucial to the design of gyrotron tubes and amplifier devices. Dissipation of the rf (ohmic loss) and excessive mode conversion are often limiting factors in the performance and robustness of the overall device. These issues and others pertaining to mmw devices can be investigated using time-domain electromagnetic (EM) field codes.[11] An advantage of simulation in the time domain is that EM characteristics can be obtained over a wide bandwidth from a single calculation. Excitation of the frequencies of interest can be obtained by launching modulated pulses driven by magnetic currents. A general field code such as AMOS[12] can be used to launch the prescribed modes at the frequency or frequencies of interest to examine mmw component performance by numerical integration of Maxwell's equations.

Mode Launching. Gyrotron oscillators operate with whispering gallery (WG) modes, for which the radial mode number greatly exceeds the axial mode number. Thus, most of the rf is distributed near the beam-pipe wall. As the mode propagates near the window, the modes couple into gaps in the window assembly, leading to mode conversion. To model gyrotron components requires the launching of transverse electric modes ($TE_{n,m}$). This is accomplished by driving magnetic currents over the beam-pipe aperture.

To describe the location of the TE drive-nodes, we rewrite the EM time-dependent curl equations:

$$\nabla \times \mathbf{H} = \sigma \mathbf{E} + \varepsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J}_s \tag{1}$$

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} - \mathbf{K}_s \tag{2}$$

in the integral form

$$\oint \mathbf{H} \bullet d\ell = \iint \left(\sigma \mathbf{E} + \varepsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J}_s \right) \bullet d\bar{A} \tag{3}$$

$$\oint \mathbf{E} \bullet d\ell = -\iint \left(\mu \frac{\partial \mathbf{H}}{\partial t} + \mathbf{K}_s \right) \bullet d\bar{A}. \tag{4}$$

K source components are co-located with H field components on the Yee lattice.[1] Referring to **Fig. 1** and **Eq. 4**, one can see that driving the K_r component of the magnetic current will excite the proper H_r, H_z and E_ϕ fields. Similarly, K_ϕ currents excite E_r and H_ϕ field components.

The proper spatial variation of magnetic currents required to obtain propagating WG $TE_{22,2}$ modes are the Bessel function $J_{22}(x)$ out to the second zero, and its derivative $J'_{22}(x)$, which directly drive K_ϕ and K_r, respectively. The amplitude distribution in time can be a modulated pulse to obtain the required frequency content (**Fig. 2**).

Field diagnostics for computing the voltage standing wave ratio (VSWR) were incorporated into AMOS by sampling electric fields at 'numerical' probes and computing the VSWR directly from the field values. If F [f(t)] denotes the forward Fourier transform, then the VSWR can be computed from the field data by first computing the reflection coefficient (no mode conversion),

$$\Gamma = \sqrt{1.0 - \left(\left| \frac{\Im[e_{smp}(t)]}{\Im[p_{mod}(t)]} \right| \right)^2}, \tag{5}$$

where $e_{smp}(t)$ is the sampled electric field on the 'downstream' side of the window, and $p_{mod}(t)$ is the modulated pulse in time. The VSWR is computed according to the definition VSWR = (1.0 + Γ)/(1.0 − Γ).

Results of mmw: High-Power rf Window Analysis and Gyrotron Amplifier Sever. Presently, gyrotrons operate in the 100- to 140-GHz and ~ 1-MW regime. Future performance requirements

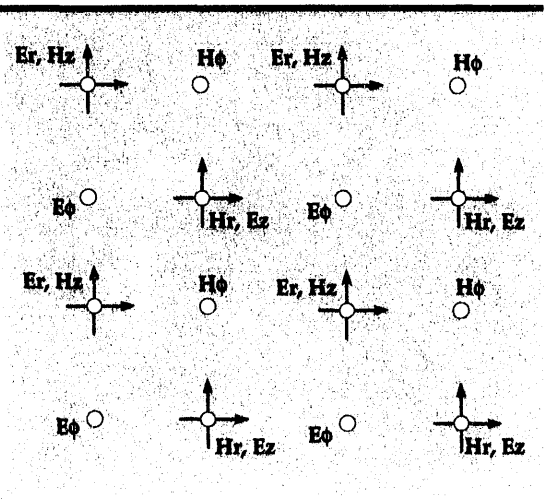


Figure 1. TE drive-node location on the Yee lattice.

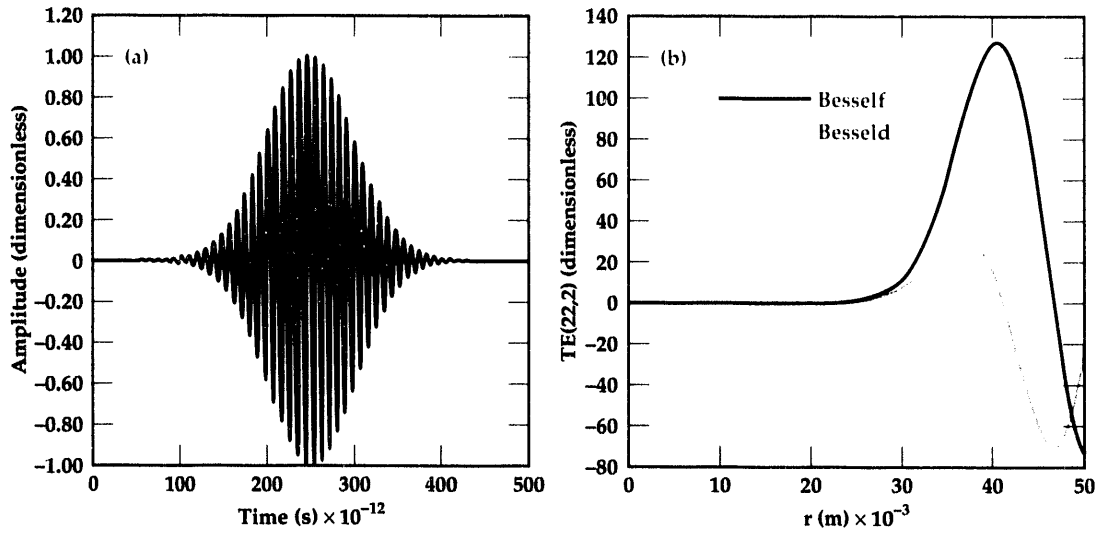


Figure 2. Launching $TE_{22,2}$ WG mode—spatial and temporal magnetic drive function.

will increase power levels to the multi-megawatt range with frequencies approaching 250 GHz. In this scenario, mmw components will be placed under severe mechanical and thermal stress. Until now, less demanding performance constraints have rendered non-ideal component effects less important. However, understanding these effects is now critical to the operation of the device.

We now examine high-order mode scattering caused by various rf window geometries at the exit of the gyrotron. The VSWR associated with the window can be determined over a broad spectrum of frequencies, using data from a single time-domain run with the technique described in the previous section.

In **Fig. 3**, we find good agreement between AMOS and analytic values[13] for the VSWR of a three-layer rf window. The gyrotron window geometry includes a beam-pipe radius of 5.08 cm with the longitudinal extent of the window at 0.443 cm. The window material has $\varepsilon_r = 9.387$, and the dielectric cooling fluid has $\varepsilon_r = 1.797$. A small difference between the AMOS and frequency code results is evident, caused by a minor variation in window element thicknesses resulting from the use of a regular grid in AMOS.

The gyrotron window structure is grown from a sapphire crystal. The window assembly is expensive and difficult to fabricate, but more realistic window geometries cannot be easily treated analytically. In **Fig. 4**, a realistic window structure with the 'coolant reservoir' is modeled. Compared

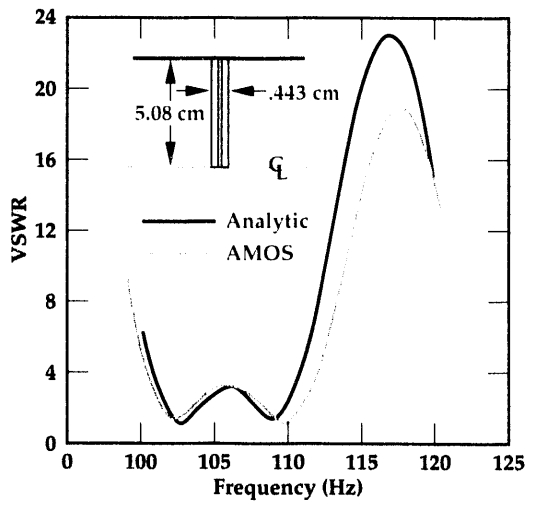


Figure 3. VSWR for idealized 110-GHz bandpass, from analytic calculations and from AMOS. The inset shows the gyrotron window geometry.

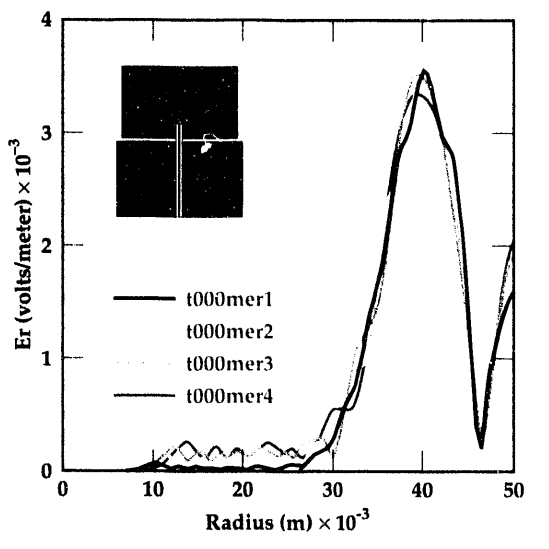


Figure 4. Radial field profile at varying longitudinal locations for realistic gyrotron rf window structures. The inset shows the window geometry.

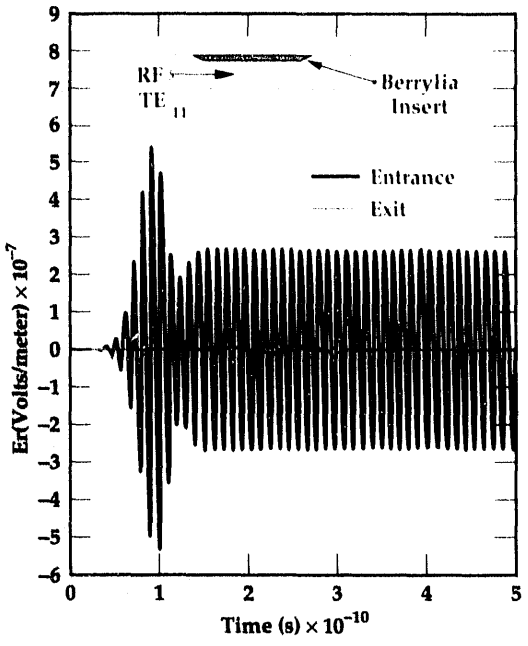


Figure 5. Sever performance for berrylia 60/40.

to the idealized window, the electric fields near axis highlight coupling to modes through the rf window near the beam-pipe wall. At the multimegawatt range, this amount of rf may be significant. However, the exact level of power per mode awaits further analysis.

We performed a second set of calculations in which we examined wave propagation through a microwave sever, a device for stopping or absorbing microwave energy while allowing a charged particle beam to pass undisturbed. Of interest is the understanding of the absorption in the lossy dielectric insert (see **Fig. 5**) for a variety of lossy rf mixtures. The beam-pipe radius is 0.95 mm, which is near the cutoff radius. As before, TE_{11} modes were launched by driving magnetic currents at the sever aperture. The material conductivity characteristics for five berrylia mixtures were obtained from the available experimental data at 12 GHz. Two materials, berrylia 80/20 and berrylia 60/40, are representative: the dielectric constants K' for berrylia 60/40 and 80/20 are 49.54 and 17.81, respectively, and the loss tangents are 0.72 and 0.22, respectively.[14]

AMOS predicted ~ 40 dB attenuation for a 95-GHz TE_{11} mode propagating toward the sever for the berrylia 60/40 mixture. In comparison, unacceptably low rf absorption characteristics for the other berrylia mixtures (**Fig. 5**) were evident. In the limit, when the conductivity is large (berrylia 60/40), the relevant diameter is not that of the beam-pipe, but instead it is the diameter inside the sever section. With cutoff given by $\lambda_c = 2\pi a/1.84$, a TE_{11} mode at 95 GHz is well below cutoff, and the fields will be attenuated. This set of calculations can be repeated when updated berrylia measurements in the ~ 100-GHz range are available. For the previous class of modeling problem, we plan to examine simulation issues such as the launching of waveguide modes near cutoff. Further, the taper of the lossy section was initially limited to a minimum of 5° because of numerical limitations of a shallow-angle staircasing of the mesh. The conforming mesh algorithm in CG-AMOS[15] will allow exact boundary representation, and thus any shallow taper.

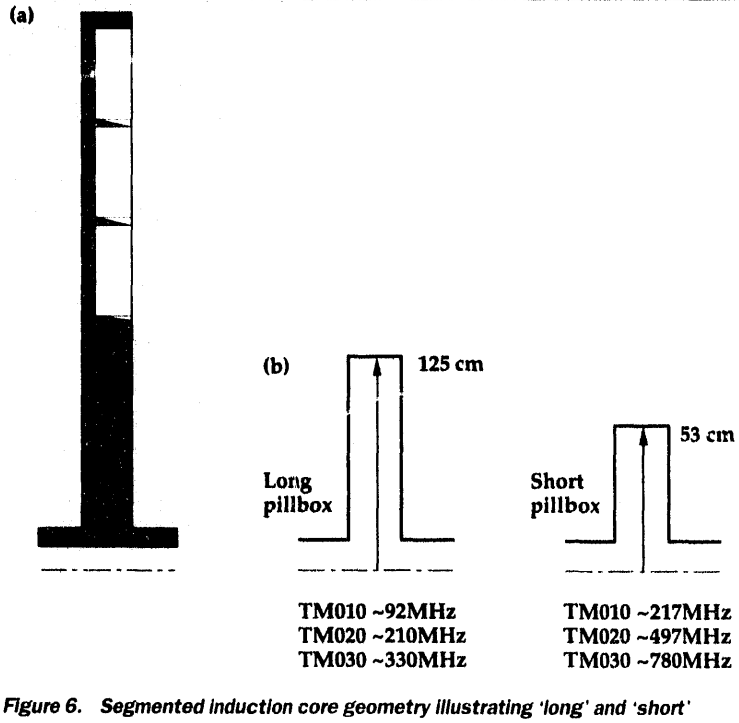


Figure 6. Segmented induction core geometry illustrating 'long' and 'short' pillbox regions.

Modeling Induction Linear Accelerator Modules

We have modeled the beam-cavity interaction impedances for induction linear accelerator cells for heavy ion fusion. The induction cell works conceptually much like a 1:1 transformer with the beam as secondary. The primary one-turn loop has a pulsed voltage V applied to it. The secondary loop around the core will have a voltage induced across its terminals that is the same as the primary voltage, i.e., from Faraday's law, V = A dB/dt. The core consists of wound metallic glass (Metglas), which has good dB/dt characteristics (1 to 5 T/μs). In the three-segment core configuration proposed by Lawrence Berkeley Laboratory (LBL), each core is fed in parallel. The secondary loop encloses all three cores, providing

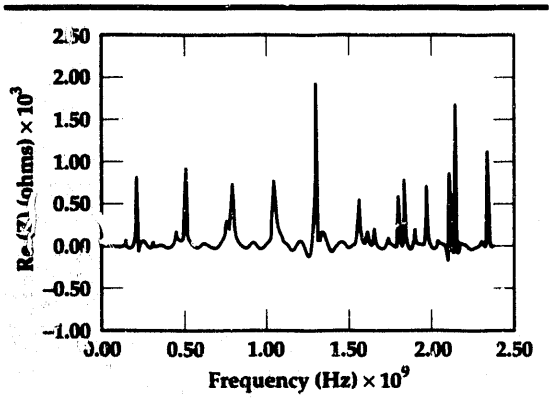


Figure 7. Impedance spectrum for the base induction module.

3 A dB/dt at the accelerating gap. The equivalent single core configuration would require either three 1-V accelerating gaps or a 3-V pulse power system. Another advantage of the segmented configuration is that one part of the core will not saturate before any other part.

In FY 92, we concentrated on understanding the segmented cell geometry and multicell accelerating modules for subrelativistic heavy ions from a beam-cavity coupling point of view.

Results from Accelerator Modeling. For the base case (**Fig. 6**), the gap width is 1.5 cm, the radial length is 1.25 m, and the overall cell width is 10 cm. In **Fig. 2**, the Fourier transform of the impulse response (wake potential)[9] of the cavity due to charge bunch transiting the accelerating gap shows the beam-cavity coupling (interaction impedances) for the monopole fields.

The dependence of the impedance as a function of v, the charge velocity, goes as $\sin^2(\omega d/2v)$ (transit time factor),[16] where ω is the angular frequency and d is the gap width. We can see, in fact, that the modes at 217 MHz and 497 MHz correspond closely to the TM_{010} and TM_{020} modes of the short pillbox geometry, and the weaker coupling of the 92-MHz mode corresponds to the TM_{010} mode of the 1.25-m radial line (**Fig. 7**).

For this simplified model, we can see that the dominant feature is the gap width. To determine if the segmented core has good damping features, we opened the gap width to 4.5 cm. The field calculations (**Fig. 8**) show low Q resonances (Q = 3.91) corresponding to the 1.25-m (long-pillbox modes') radial line.

The final set of results to be discussed involves the stacked accelerating modules (**Fig. 9a**). In the absence of inter-cell interactions, it is expected that the impedance of a single cell will add in series.

The result (**Fig. 9b**) shows that this is indeed the case. Similar simulations were performed for the dipole modes (TM_{1nm}) to study the possible impact of beam break-up instability[16] in heavy ion drivers. These impedance calculations, coupled with analytic and calculational results from the BREAKUP beam dynamics code, indicate controllable beam break-up modes.[17]

Future Work

Field calculations show that for the current, heavy-ion, linear accelerator cell configuration, the pulse power system and accelerating cores are largely decoupled from wakefields. Although this idealized cell has a high impedance-geometry figure of merit, schemes to lower the Q of the lower-order modes can be developed. We will continue our work to model the fully three-dimensional, multi-beam-pipe cell as proposed by LBL for the Induction Linac Systems Experiments. We intend to develop detailed anisotropic, dispersive media models of Metglas in the coming year. We also will be involved in research on designs for the next-generation induction accelerator for radiography. In the latter work, all cavity modeling results will be incorporated into beam dynamics codes, with the goal being end-to-end simulations leading to dosage estimates.

For high-average-power mmw components, we have shown how application of field codes can be used to analyze complex geometrical aspects that are not amenable to analytical techniques. TE modes may be launched in a beam-pipe by use of the dual K term (magnetic current) in the Faraday–Maxwell equation. TM modes may be launched using a similar dual technique. Since the rf impinging on the window (110-GHz tube) is in the WG mode, it remains to be seen if the approach for extracting usable modes will involve either (1) converting WG to usable modes external to the

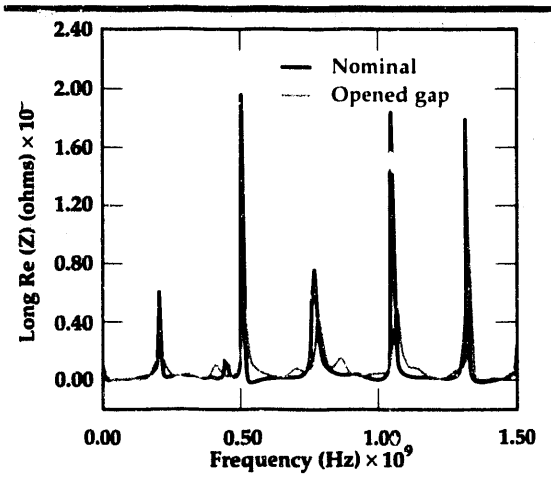


Figure 8. Impedance spectrum for opened gap width.

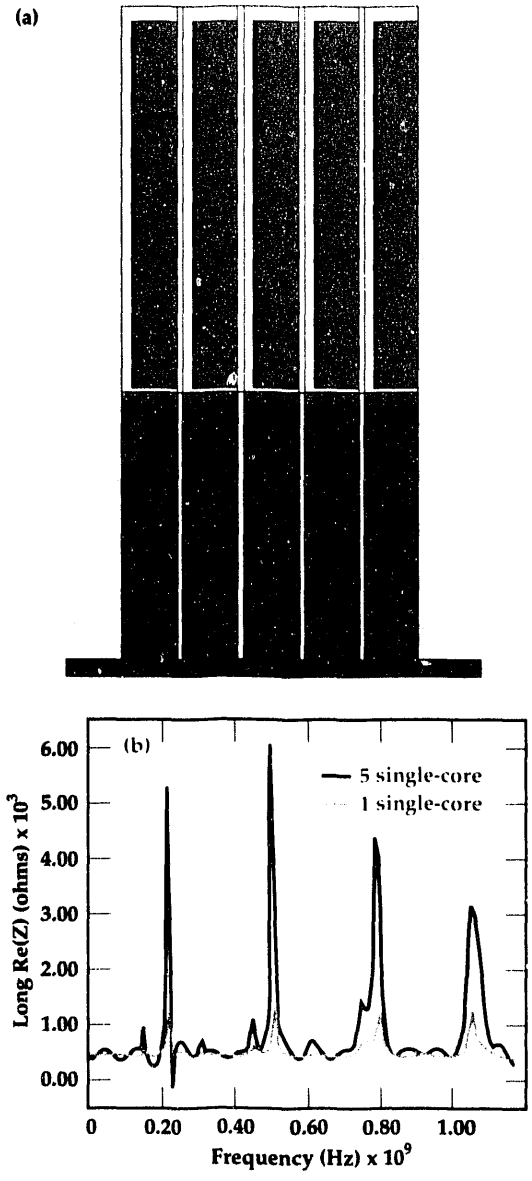


Figure 9. Longitudinal impedance for multi-cell accelerating module.

window or (2) designing a window that can allow a Gaussian power distribution for long pulse operation.

Our future work will also involve the use of unstructured meshes to simulate the important aspects of high-bandwidth rf windows by detailed modeling of window surfaces. We will also implement the diagnostics that will derive mode information directly from the field data.

Our work in mmw has been formalized in a two-year, $2.8 M Cooperative Research and Development Agreement with Varian Associates, Inc., Microwave Products Tube Division, of Palo Alto, California and is funded by the Defense Systems and Engineering Directorates at Lawrence Livermore National Laboratory (LLNL) for FY93 and FY94. We will be examining new mmw circuit concepts, input/output couplers, water loads, and lossy drift tube sections, as well as continuing to model high-average-power windows (fixed frequency and wideband), using field codes.

Acknowledgements

Stimulating discussions with Drs. J.J. Barnard and G.J. Caporaso (LLNL), and W. DeHope (Varian) are gratefully acknowledged.

1. K.S. Yee, *IEEE Trans. Antennas Propag.* **AP-14**, 302 (1966).
2. A. Taflove and G. Brodwin, *IEEE Trans. Microwave Theory and Techniques* **MTT-23**, 623 (1975).
3. G. Mur, *IEEE Trans. Electromagnetic Compat.* **EMC-23**, 1073 (1981).
4. K.K. Mei and J. Fang, *IEEE Trans. Antennas Propag.* **AP-40**, 1001 (1992).
5. J.F. DeFord, and G. Kamin, *Application of Linear Magnetic Loss Model of Ferrite to Induction Cavity Simulation. Proc of 1990 Linear Accelerator Conf.* (Albuquerque, New Mexico), (September 14, 1990).
6. R. Luebbers et al., *IEEE Trans. Antennas Propag.* **AP-32**, 222 (1990).
7. N.K. Madsen, *A Divergence Preserving Discrete Surface Integration Algorithm for Time-Domain Maxwell Equations*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC 109787 (1993).
8. N.K. Madsen and R.W. Ziolokowski, *Electromagnetics* **10** (1,2), 127 (1990).
9. J.F. DeFord and G.D. Craig, *Part. Accel.*, **37-38**, 111, (1992).
10. W.H. Urbanus, R.W.B. Best, A.G. Verhoeven, and M.J. van der Wiel (FOM-Instituut voor Plasmafysica, Nederland); M. Caplan (Lawrence Livermore National Laboratory, Livermore, California); V. Bratman and G. Denisov (Institute of Applied Physics, Russia); and A.A. Varfolomeev (Kurchatov Institute, Russia), "A 1-MW Free Electron Maser for Fusion Application," *Proc. Third European Particle Accelerator Conf.* (Berlin, Federal Republic of Germany), (March 24–28, 1992).
11. M. Caplan, *Application of AMOS Code to Design of Comp' x Structures for Selective Microwave Absorption and Transmission*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-ID-109254 (1991).
12. J.F. DeFord, G.D. Craig, and R.R. McLeod, "The AMOS Wakefield Code," *Proc. Conf. Computer Codes and the Linear Accelerator Community* (Los Alamos, New Mexico), 265 (January 22–25, 1990).
13. J.M. Neilson, P.E. Latham, M. Caplan, and W.G.

Lawson, *IEEE Trans. Microwave Theory and Techniques* **37**, 1165 (1989).

14. W. DeHope, Private communication (April 1992).

15. C.C. Shang and J.F. DeFord, "Modified-Yee Field Solutions in the AMOS Wakefield Code," *Proc. 1990 Linear Accelerator Conf.* (Albuquerque, New Mexico), (September 14, 1990).

16. R.J. Briggs, D.L. Birx, G.J. Caporaso, V.K. Neil, and T.C. Genoni, *Part. Accel.* **18**, 41 (1985).

17. G.J. Caporaso, "Transverse Instability in a Heavy Ion Fusion Induction Linac," *Proc. Longitudinal Instability Workshop* (Berkeley, California), (February 1992).

Electromagnetic Modeling and Experiments for Dispersive Media

Scott D. Nelson and
Carlos A. Avalle
Defense Sciences Engineering Division
Electronics Engineering

The Ground Penetrating Imaging Radar Project was established to investigate the feasibility of designing an electromagnetic (EM) radar system to examine the internal structure of concrete structures typically found in the highway industry. The central project involved the coordination of the EM modeling, imaging, code design, and experimentation efforts at Lawrence Livermore National Laboratory. The modeling effort generated data for EM imaging and enabled the precise control of individual parameters in the model.

Introduction

The modeling effort consisted of three phases: (1) complex permittivity analysis of cement using a coaxial line; (2) model construction and experimental verification in one dimension, which was represented by a coaxial line in the time domain; and (3) model construction and experimental verification in two dimensions, which was represented in the model by a slice through a concrete block and experimentally by an antenna with a fan beam.

Progress

One-dimensional (1-D) and two-dimensional (2-D) models were constructed and compared with experimental data. The 1-D data served as a preliminary test of the dispersion algorithms added to the AMOS[1] 2-1/2 D FDTD (Finite Difference Time Domain) electromagnetic (EM) modeling code. The 2-D data served as a verification for the concrete model and as a method to create waveforms for the image reconstruction algorithms.

The following parametric studies were performed:

Pulse width:	100 ps to 1000 ps
Change depth:	10-mm to 150-mm voids
Change size:	5-mm to 75-mm voids
Aggregate %:	10% to 50% aggregate probability
Multi-Receiver:	5-mm to 45-mm spacing (simulates highway speed or prf changes)
Multi-Target:	no targets, 1 void, 2 rebars, 2 rebars + 1 void, grate, shadowing
Air/Concrete:	1 case
Bistatic data:	1 set

The parametric studies gave some results that were already hypothesized[2,3]: (1) the desired frequency is close to 2 GHz; (2) the correction filters have an image 'gain' of a factor of two; (3) large targets reradiate in addition to reflecting; (4) monostatic data spatially averages out the aggregate effects for small- and medium-sized particles; (5) bistatic data is more susceptible to aggregate effects near the transmitter than monostatic data, which is why commercial systems do not see a lot of aggregate effects; and (6) shadowing is not a significant problem as long as the spacing between the rebars is greater than three times the pulse width, and the rebars are not appreciably larger than the pulse width[4] (to give the diffracted field time to repair the wave front).

1-D and 2-D Verification Effort

A coaxial line was used for the 1-D simulation, with the cement sample embedded in a removable section of 2-in. coaxial line. The Lorentzian param-

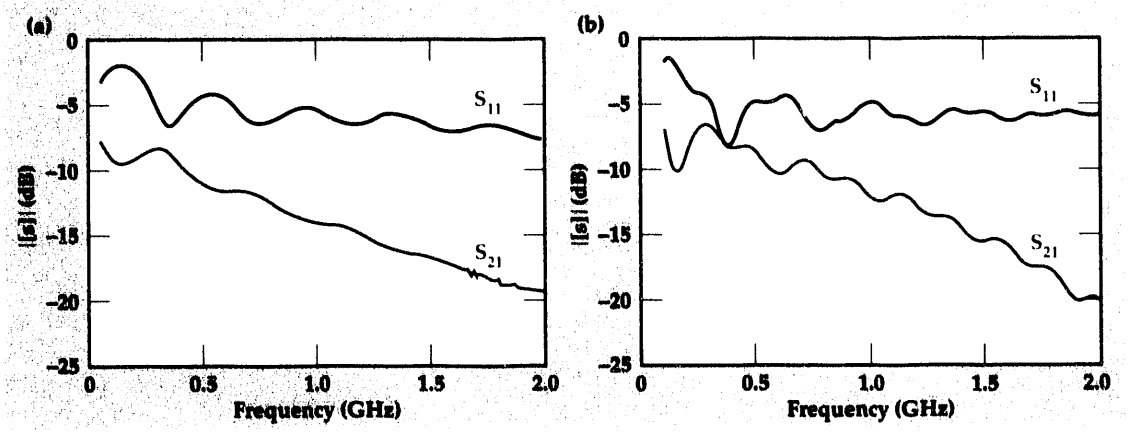


Figure 1. The s parameters for the 1-D (a) experimental and (b) modeled results. Note the 400-MHz ripple due to the sample size.

eters used for the initial 1-D and 2-D dispersion cases are as follows:

$$F\left\{\sum_{i=1}^{i=\infty}\alpha_i e^{-\beta_i t}\sinh(\gamma_i t)\right\}$$

$$=\frac{1}{2}\left[\sum_{i=1}^{i=\infty}\frac{1}{\beta_i-\gamma_i+j\omega}-\frac{1}{\beta_i+\gamma_i+j\omega}\right]$$

$$\alpha_1 = 1.55\cdot 10^{10},$$
$$\beta_1-\gamma_1 = 3.29\cdot 10^{8},$$
$$\beta_1+\gamma_1 = 5.58\cdot 10^{10},$$

$$\alpha_2 = 1.63\cdot 10^{12},$$
$$\beta_2-\gamma_2 = 1.16\cdot 10^{11},$$
$$\beta_2+\gamma_2 = 3.62\cdot 10^{12}.$$

The results for the 1-D experiment, performed in the Lawrence Livermore National Laboratory (LLNL) EM lab,[5] are shown in **Fig. 1**. The 400-MHz ripple seen in **Fig. 1** represents the resonance in the material sample due to its length. The 2-D verification experiment was performed in the LLNL Anechoic Chamber using a broadband antenna (with a fan beam pattern), a concrete block, a broadband field probe (Prodyne Ddot probe), and a transient digitizer. The time domain waveforms for the experimental and modeled cases are shown in **Fig. 2**. The antenna beam pattern, pulse shape, aggregate, and dispersion effects of the concrete block were included in the model. The finite size of the block and the diffraction around the block were also included in the model.

2-D Concrete Simulation

Figure 3 shows the received waveform from one of the receiver antennas (1 of 15), with the individual reflection identified for a geometry typical of the project.[6] In this, there were two rebars and one void at different depths and cross-range distances. The effects of the aggregate in the problem are clearly visible. The aggregate is modeled as discrete scattering bodies of finite size. The lighter

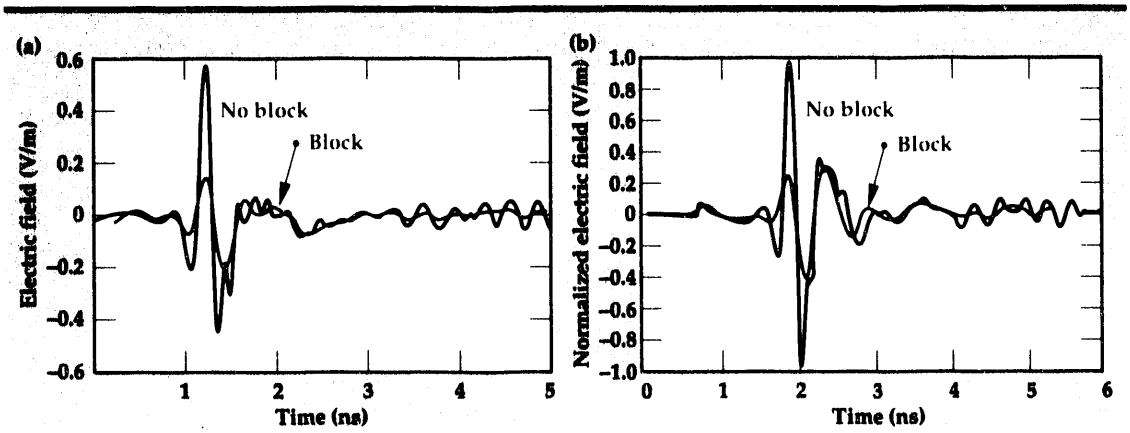


Figure 2. The time domain waveforms for the 2-D concrete block experiment compared to the modeled results. The modeled results are normalized. The negative-going double peak in the experimental results is combir into a single negative-going peak in the modeled results.

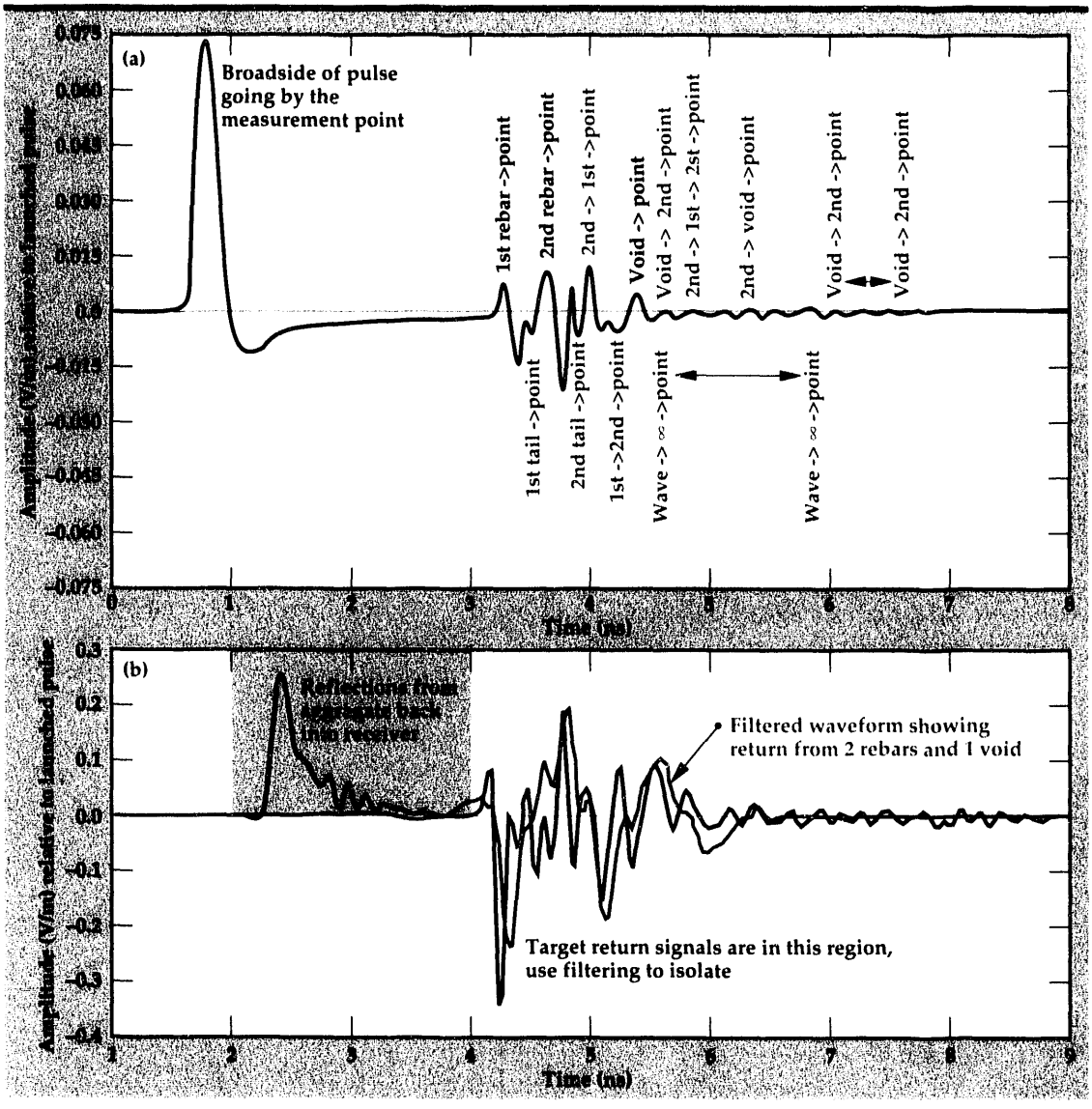


Figure 3. Received waveforms from one of 15 receiver antennas. The first waveform shows the received time domain waveform from one of the receiver elements in the modeled case with no aggregate and no dispersion. The individual target reflections are identified. The second waveform shows the effects of the aggregate in the problem. The receiver in this case is closer to the transmitter than in the previous case. The lighter curve shows the resulting waveform after the application of the imaging team's adaptive filter.

curve in **Fig. 3** shows the results from the imaging team's adaptive filter. The reflections from the two rebars and from the void are clearly visible. The parametric study listed above was performed; **Fig. 4** shows a typical wave propagation scenario. The 90° beamwidth of the antenna, the two rebars, the void, and the aggregate are all visible. The aggregate radius is one half that of the rebars and represents a 30% probability duty cycle. The early-time backward-propagating waves are due to the aggregate.

Conclusions

Aggregate sizes less than one third of the pulse width did not create significant reflections at the receivers. Aggregate sizes on the order of the pulse width created discrete waveforms in the received signals in the areas around the transmitter where the power density was the strongest. Due to the dispersion effects of the concrete, the aggregate was most reflective (in a relative sense) near the transmitter when the pulse was still short. Since

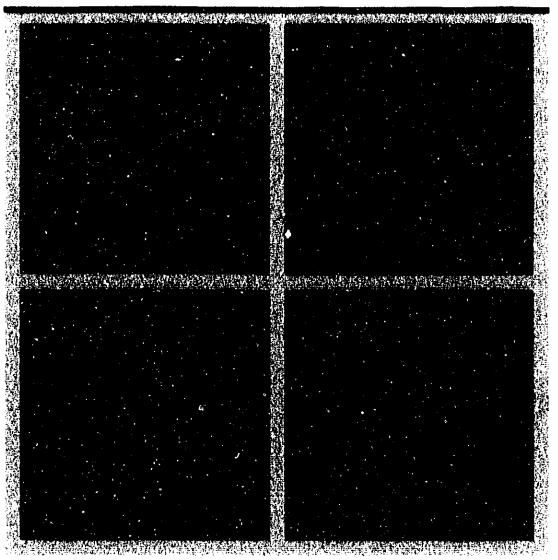


Figure 4. Four frames from a propagating EM wave sequence showing the reflections from the various targets. The transmit horn is on the upper surface of the concrete, and the wave propagates down into the material. The two rebars are on the right; the single void is on the left. Note the polarity difference between the rebars and the void. Also observe the multiple reflections between the outer targets and the center target.

the area around the transmitter also had the greatest power density, then the maximum returned waveform (in an absolute sense) from the aggregate was also seen in this region.

This modeling effort demonstrated that the original complex permittivity data obtained in the 1-D case does support large-scale material modeling, as was expected. More important, this effort confirmed the assumptions that were made about the aggregate and its modelability.[7] The size of the individual rocks constituting the aggregate and their probability distribution were more important than some exact spatial placement for each rock in the model. This result provided direct support for the usability of an adaptive filter as part of the imaging effort, to remove the aggregate effects even when individual aggregate particles generate discrete reflection waveforms in the received waveform. The aggregate specifications are known, or at least are specified, for concrete structures.

Future Work

A realistic dispersive concrete model will be introduced in three dimensions, using the TSAR code, and a realistic synthetic aperture radar (SAR) will be modeled using a section of a typical bridge deck as the target of interest. The transmitter/receiver designs will be optimized to efficiently use the spectral information in the waveforms. Comparisons will be made to the experimental effort being performed on the 6 ft-x-6 ft-x-1 ft concrete test slab. This effort is already under way.

Issues that still need to be addressed are: (1) single vs multiple transmitters; (2) different antenna beamwidths based on distance from the transmitter(s); (3) a single linear array sweeping a synthetic aperture vs a real SAR; (4) optimum receiver antenna size vs receiver density; and (5) the temporal holographic use of the time domain waveforms.

Acknowledgements

Thanks go to John DeFord (LLNL) for his advice and timely modifications to the AMOS code and to Robert McLeod (LLNL) for his efforts in adding dispersion to the TSAR code. The imaging team consisted of Jose M. Hernandez (LLNL) and Joe Arellano[8] (Sandia National Laboratory) with assistance from James Brase (LLNL).

1. J.F. DeFord, G. Kamin, L. Walling, and G.D. Craig, *Development and Applications of Dispersive Soft Ferrite Models for Time-Domain Simulation*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-109495 (1992).
2. K. Olp, G. Otto, W.C. Chew, and J.F. Young, *J. Mater. Sci.* **26**, 2978 (1991).
3. K.S. Cole and R.H. Cole, *J. Chem. Phys.*, 341 (April 1941).
4. M. Kanda, *IEEE Trans. Antennas Propag.*, **26**, 439.
5. C. Avalle, *Broadband Complex Permittivity Measurements of Cement*, Lawrence Livermore National Laboratory, Livermore, California (in preparation).
6. R. Zoughi, G.L. Cone, and P.S. Nowak, "Microwave Nondestructive Detection of Rebars in Concrete Slabs," *Materials Evaluation—American Society for Nondestructive Testing*, 1385 (November 1991).
7. K.R. Maser, "Detection of Progressive Deterioration in Bridge Decks Using Ground Penetrating Radar," *Proc. ASCE Convention* (Boston, Massachusetts), (October 27, 1986).
8. J. Arellano, "Adaptive Filter for LLNL's Impulse Radar Inspection of Roadways and Bridges," *EE373A, Adaptive Signal Processing* (Winter 1991–92).

Band Gap Engineering for Infrared Detectors

J. Brian Grant
Engineering Research Division
Electronics Engineering

We have extended and improved modeling codes for strained layer superlattices. Significant improvements include capabilities for reliable subband tracing and multilayer modeling; better validation of eigenstates; and the calculation of physical quantities such as wave function and optical absorption profiles and effective masses.

Introduction

Applications for infrared (IR) detectors include military, civilian, and medical devices with current interest focused on the far IR spectrum (wavelengths greater than 10 μm). Small band gaps corresponding to this range push the engineering parameters in commercial IR detector technology, which is alloying $Hg_{1-x}Cd_xTe$. This process is not only very sensitive to composition, but toxic and volatile too. While a switch from alloy to superlattice technology would relax the conditions on exact composition,[1] an even greater benefit is obtained from a switch to III-V semiconductors,[2] where internal strain can be used in designing band gaps. Recent advances in $Ge_{1-x}Sb_x$ alloy fabrication provide another alternative,[3] which is less sensitive to alloy composition. Severe lattice mismatches prevent the formation of Ge/Sb superlattices.

Of particular interest are superlattices of GaSb/$In_{1-x}Al_xAs$ in which the alloy composition controls lattice mismatch, the source of internal strain. The major tradeoff is reduced optical absorption because the superlattice is type II, i.e., the conduction and valence states are principally confined to different material layers. Thinner layering and strain help to overcome this by increasing tunneling and consequently increasing wave function overlaps.

The modeling of strained layer superlattices is perhaps best achieved through the use of interface transfer matrices. These matrices identify how wave functions in one bulk-like material are transformed into those of the next. While this idealizes interfaces, it eliminates the excessive basis size of traditional approaches, which must model each atom in the periodic structure. The matrix approach also allows modeling of non-periodic, finite-sized structures.

Computational speed is further enhanced by use of the **k·p** theory, which expands bulk wave functions in terms of those at the Brillouin zone center. Because interest is in the lowest conduction and highest valence, i.e, in bands near the Brillouin zone center, only the eight spin split s- and p-wave functions of the bulk are kept. Löwden perturbation theory is used to extend the range of accuracy by including effects of other wave functions to first-order.

Progress

A significant portion of our project has focused on extending and improving existing modeling codes.[4,5] By combining several codes into a single package and by reducing and simplifying required

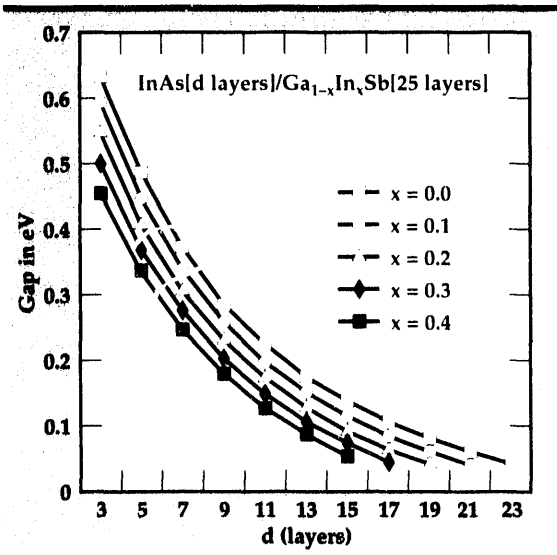


Figure 1. Strain and composition effects on band gap.

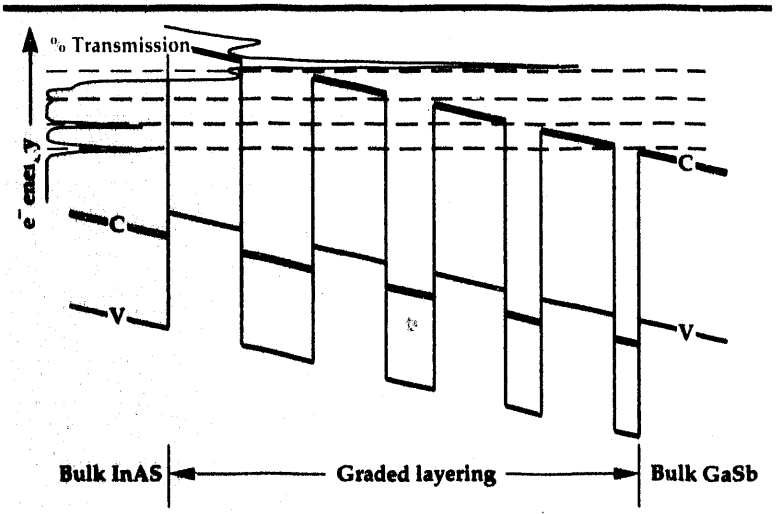


Figure 2. Percent transmission as a function of electron energy relative to the bulk conduction (C) and valence (V) band edges. Slope of band edges indicates an applied voltage.

input, the ease of use was greatly improved. Further, the resulting **sls** code was made truly 'user friendly' by building an **X** windows[TM] interface that wraps around it. The incorporation of capabilities for automatically locating subband-edges, as well as for reliable energy subband tracing, has also enhanced the code's versatility.

Other significant improvements include better validation of eigenstates and allowance for calculation of a variable number of bands. Both help to alleviate the numerical instabilities introduced by the inclusion of extremely weak, but finite, tunneling wave functions sometimes required for accuracy.

Significant capabilities were added to the **sls** code to allow multilayer superlattices. Previously, only two layers could be modeled. Further advantage was taken of the interface-matching approach by allowing the non-periodic boundary conditions of a finite-sized superlattice sandwiched between two semi-infinite bulk materials. Calculation of various physical quantities were also added. Among the more important are wave function and optical absorption profiles, and effective masses.

Another significant component of this project was technology transfer. Working with Bill Ahlgren of Santa Barbara Research Center (SBRC), a subsidiary of Hughes, we engaged in numerous modeling activities. The **sls** code has been transferred to SBRC for further use and continued collaborations.

Figure 1 shows a sample calculation for SBRC that demonstrates the effects of strain (alloy) and layer thickness on band gaps in InAs / $Ga_{1-x}In_xSb$ superlattices. Note that a target wavelength of 12 μm would require very thick layers of InAs without alloying (x = 0.0), but that only 11 or 12 atomic layers would be required for a composition where x = 0.4. In addition to the band gap, SBRC is interested in the actual position of the conduction band with respect to the chemical potential (Fermi level) and the optical absorption as functions of the same parameters. Similar plots can be made of those values. Values of effective masses calculated by the **sls** code can also be used for additional calculations by SBRC.

Modeling for programs at Lawrence Livermore National Laboratory (LLNL) has been based on graded-layer superlattices that slowly accommodate large band offsets. Externally applied voltages then provide a rather constant-energy conduction band throughout the structure. The **sls** code easily identifies acceptable-transmission energy ranges, as shown in **Fig. 2,** as well as resonant tunneling states, which build up much larger electron concentrations between barrier layers. While this code is unable to provide self-consistent estimates of current densities, the information available can provide a guide to superlattice grading, and quantitative estimates of the necessary external voltages.

Future Work

This project has included software development and improvement, modeling, and technology transfer. The result is a rather complete and usable strained-layer superlattice code. Relations with SBRC are continuing, supported in part by the Physics Department at LLNL.

1. D.L. Smith, T.C. McGill, and J.N. Schulman, *Appl. Phys. Lett.* **43**, 180 (1983).

2. D.K. Arch, G. Wicks, T. Tonae, and J.-L. Staudenmann, *J. Appl. Phys.* **58**, 3933 (1985).

3. S.M. Lee and W. Paul, "Electronic Band Gap Measurements of *Bulk* Metastable Crystalline $Ge_{1-x}Sn_x$ Alloys, $0 \leq x \leq 0.311$," preprint.

4. C. Mailhiot and D.L. Smith, *Crit. Rev. Solid State Mater. Sci.* **16**, 131 (1990).

5. D.L. Smith and C. Mailhiot, *Rev. of Mod. Phys.* **62**, 173 (1990).

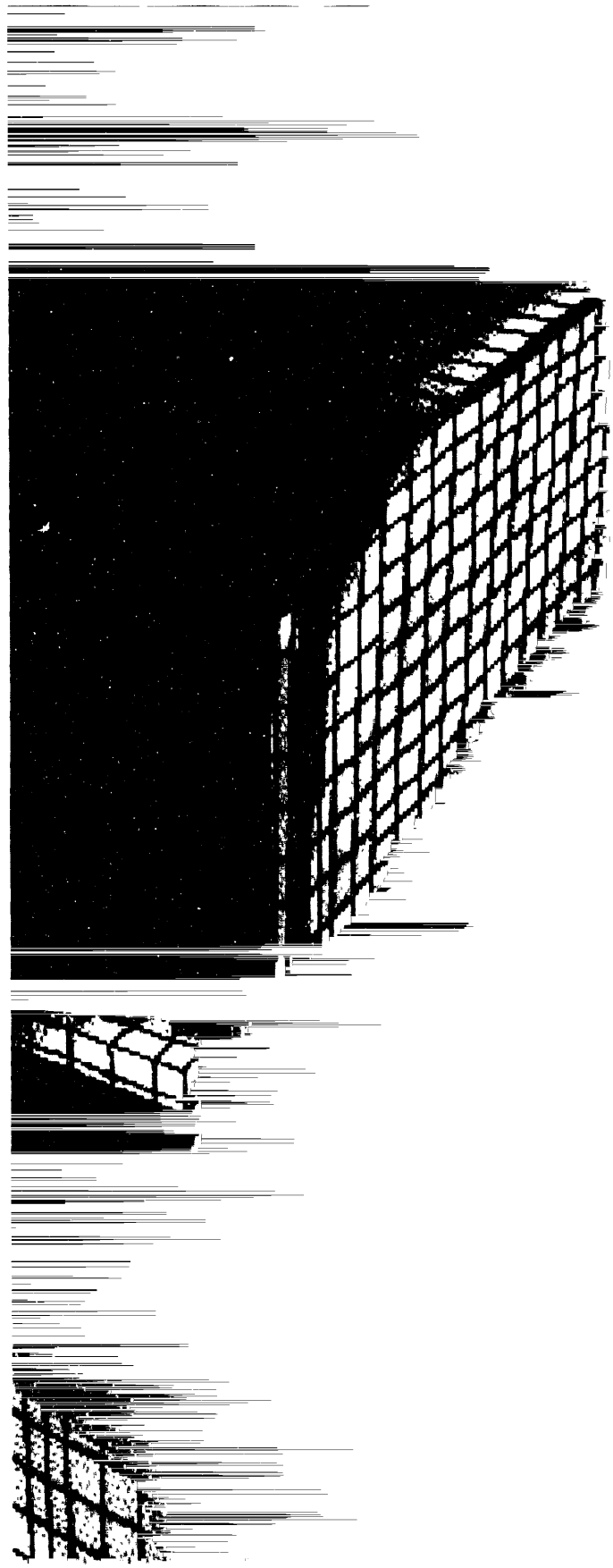
Computational Mechanics

The Computational Mechanics thrust area sponsors research into the underlying solid, structural, and fluid mechanics and heat transfer necessary for the development of state-of-the-art general purpose computational software. The scale of computational capability spans office workstations, departmental computer servers, and Cray-class supercomputers. The DYNA, NIKE, and TOPAZ codes have achieved world fame through our broad collaborators program, in addition to their strong support of on-going Lawrence Livermore National Laboratory (LLNL) programs. Several technology transfer initiatives have been based on these established codes, teaming LLNL analysts and researchers with counterparts in industry, extending code capability to specific industrial interests of casting, metalforming, and automobile crash dynamics. The next-generation solid/structural mechanics code, ParaDyn, is targeted toward massively parallel computers, which will extend performance from gigaflop to teraflop power.

Our work for FY-92 is described in the following eight articles: (1) Solution Strategies: New Approaches for Strongly Nonlinear Quasistatic Problems Using DYNA3D; (2) Enhanced Enforcement of Mechanical Contact: The Method of Augmented Lagrangians; (3) ParaDyn: New Generation Solid/Structural Mechanics Codes for Massively Parallel Processors; (4) Composite Damage Modeling; (5) HYDRA: A Parallel/Vector Flow Solver for Three-Dimensional, Transient, Incompressible Viscous Flow; (6) Development and Testing of the TRIM3D Radiation Heat Transfer Code; (7) A Methodology for Calculating the Seismic Response of Critical Structures; and (8) Reinforced Concrete Damage Modeling.

Gerald L. Goudreau
Thrust Area Leader

Section 2



2. Computational Mechanics

Overview

Gerald L. Goudreau, Thrust Area Leader

Solution Strategies: New Approaches for Strongly Nonlinear Quasistatic Problems Using DYNA3D

Robert G. Whirley and Bruce E. Engelmann 2-1

Enhanced Enforcement of Mechanical Contact: The Method of Augmented Lagrangians

Bradley N. Maker and Tod A. Laursen 2-7

ParaDyn: New Generation Solid/Structural Mechanics Codes for Massively Parallel Processors

Carol G. Hoover, Anthony J. De Groot, James D. Maltby, and Robert G. Whirley 2-11

Composite Damage Modeling

Edward Zywicz 2-15

HYDRA: A Flow Solver for Three-Dimensional, Transient, Incompressible Viscous Fluid

Mark A. Christon 2-19

Development and Testing of the TRIM3D Radiation Heat Transfer Code

James D. Maltby 2-23

A Methodology for Calculating the Seismic Response of Critical Structures

David B. McCallen, Francois E. Heuze, Lawrence J. Hutchings, and Stephen P. Jarpe 2-27

Reinforced Concrete Damage Modeling

Sanjay Govindjee and Gregory J. Kay 2-35

Solution Strategies: New Approaches for Strongly Nonlinear Quasistatic Problems Using DYNA3D

Robert G. Whirley and
Bruce E. Engelmann
Nuclear Explosives Engineering
Mechanical Engineering

The analysis of large, three-dimensional, strongly nonlinear structures under quasistatic loading is an important component of many programs at Lawrence Livermore National Laboratory (LLNL). The most widely used formulation for this type of problem is an implicit solution process with a linearization and iteration approach to solving the coupled nonlinear equations that arise. Our research investigates an alternative approach, in which an iterative solution method is applied directly to the nonlinear equations without the use of a linearization. This approach alleviates some of the difficulties encountered when linearizing nonsmooth phenomena such as mechanical contact. The first iterative method explored is the dynamic relaxation method, which has been implemented into the LLNL DYNA3D code, and combined with software architecture and computational mechanics technology developed for explicit transient finite element analysis. Preliminary analysis results are presented here for two strongly nonlinear quasistatic problems to demonstrate the promise of a linearization-free approach.

Introduction

Many programs at Lawrence Livermore National Laboratory (LLNL) use nonlinear finite element structural analysis to guide engineering projects. Applications include the determination of weapon component response to a variety of structural and thermal environments; the study of stresses in nuclear fuel transportation casks; and the simulation of the forming of sheet metal parts to optimize processing parameters and minimize waste. These applications share the common features of being three-dimensional (3-D), quasistatic, and strongly nonlinear, and illustrate wide use of this type of computer analysis.

Nonlinear finite element structural analysis methods may be divided into two categories: implicit methods and explicit methods. Implicit methods are typically used for quasistatic and low-frequency dynamic problems. The LLNL NIKE2D[1] and NIKE3D[2] codes are based on an implicit formulation. This approach uses a small number of large increments to step through the simulation, with the increment size chosen by the analyst to satisfy accuracy and convergence requirements. An implicit analysis code must solve a coupled system of nonlinear algebraic equations at each step, usually by a linearization and iteration procedure. This linearization leads to a coupled system of linear algebraic equations that must be solved at each iteration of each step in the analysis. Typically, the iteration process is continued within a step until some convergence measure is satisfied, then the solution is advanced to the next step.

Explicit methods are typically used for high-frequency dynamics, wave propagation, and impact problems. The LLNL DYNA2D[3] and DYNA3D[4] codes are based on an explicit formulation. In contrast to implicit methods, explicit methods use a large number of small increments to step through a problem, with the increment size chosen automatically to satisfy stability requirements. This stability requirement essentially dictates that

the time increment size must be smaller than the time it would take a stress wave to propagate across the smallest dimension of the smallest element in the mesh. An explicit code does not solve coupled equations at each step, and therefore the update from step to step is much faster than in an implicit code.

In practice, implicit methods have worked well for strongly nonlinear quasistatic problems in two dimensions, but have encountered difficulties on 3-D problems. These difficulties can be attributed to three primary factors. First, large 3-D contact problems, especially sheet-forming problems, have large matrix bandwidths due to the large contact area between the sheet and tool surface. This large matrix bandwidth translates into high computer memory requirements and expensive linear solutions at each iteration of the nonlinear solution process. Second, strongly nonlinear problems often contain discontinuous phenomena that are difficult to linearize. For example, in a contact problem, the interface pressure abruptly changes from zero when two bodies are separated to a finite value when the bodies come into contact. Obtaining an accurate linearization of such abrupt changes is difficult, and this is manifest in the code as slow convergence or nonconvergence of the linearization and iteration procedure within a step. Finally, when solution difficulties are encountered in a large 3-D problem, it is much more difficult to troubleshoot the model than it would be in a similar two-dimensional model. Often there are few clues to suggest why the iteration procedure is having difficulty converging to a solution. The development of a more robust solution strategy for strongly nonlinear quasistatic problems is the primary objective of this effort.

One approach to improving the performance of implicit methods for large, 3-D, strongly nonlinear quasistatic problems focuses on the solution of the large linear system that arises from the linearization and iteration approach. An iterative method, such as the use of a preconditioned conjugate gradient, is one approach to solving the linear system. This approach was investigated in the LLNL NIKE3D code,[5] and culminated in the development of an iterative solver now used in the production code version. Although this approach reduces memory requirements and may reduce CPU costs for the linear equation solution, it does nothing to improve the convergence of the nonlinear iteration.

An alternative approach for difficult quasistatic problems is to use an explicit transient dynamics code, and apply the loads so slowly that the dynamic effects are negligible, and therefore a quasistatic solution is obtained. Although this approach is often used by engineering analysts, it does have several disadvantages. First, the best rate of load application to minimize dynamic effects while keeping the analysis cost tolerable is not known *a priori*, and often requires some experimentation. Also, it is important to minimize artificial oscillations in the solution when history-dependent material models such as plasticity are included, and this further complicates the choice of analysis parameters. Finally, this approach obtains only an approximate quasistatic solution, and the amount of error due to dynamic effects requires some effort to quantify.

These observations suggest the alternate approach followed in our work. The basic linearization and iteration paradigm is abandoned, and an iterative solution method is applied directly to the nonlinear equations. This method is combined with much of the computational mechanics technology and software architecture developed for explicit transient dynamic analysis to produce a code that solves the nonlinear problem directly by using a large number of rather inexpensive iterations, and without solving a coupled linear system. The essential elements of this approach and its development in the LLNL DYNA3D code are described below.

Progress

In FY-92, we developed an iterative quasistatic solution capability in DYNA3D, based on the dynamic relaxation (DR) method. In addition to the implementation of the basic DR procedure, a load incrementation framework has been incorporated into DYNA3D that allows a true quasistatic solution to be obtained at a load level before the load is increased for the next increment. In addition, a spectrum contraction algorithm has been implemented that greatly improves the efficiency of the method. Also, extensions have been developed for the rigid-body mechanics formulation and the treatment of boundary conditions to accommodate nonlinear quasistatic problems within the DYNA3D framework. The resulting code is now being used as a testbed to evaluate the overall robustness and efficiency of the DR method, and to study improvements in the formulation, contact algorithms, and adaptive damping procedures.

Overview of the DR Approach

In the DR method, the equations governing a quasistatic analysis are first transformed into those governing a dynamic system. The nonlinear cou-

pled equations for a quasistatic problem may be written as

$$\mathbf{p}(\mathbf{x}_0) = \mathbf{f}, \qquad (1)$$

where **p** is the internal nodal force vector resulting from stress states in the finite elements; $\mathbf{x}_0$ is the nodal displacement solution; and **f** is a vector of externally applied loads. An associated dynamic problem may be written as

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{p}(\mathbf{x}) = \mathbf{f}, \qquad (2)$$

where dots denote differentiation with respect to time. With the appropriate choice of mass and damping matrices, **M** and **C**, the solution of the dynamic problem as time gets large approaches the solution of the quasistatic problem, i.e.,

$$\lim_{t\to\infty} \mathbf{x}(t) = \mathbf{x}_0. \qquad (3)$$

The iterative scheme is defined by applying the explicit central difference method to integrate the dynamic equations in time.

The success of the dynamic relaxation iterative method to solve highly nonlinear quasistatic problems depends on many factors including the specification of mass and damping, as well as the development of an incremental loading strategy.

Spectrum Contraction

The efficiency of the DR method may be improved by contracting the spectrum of the global equations. This is easily accomplished by proper choice of the mass matrix **M** in **Eq. 2**. The construction of the mass matrix should not dominate the computation, and thus it should be based on conveniently available quantities. In our algorithm, the mass matrix is constructed from an assemblage of element contributions. The mass matrix of each element is scaled so that all elements have a uniform critical time step, and thus information flows throughout the mesh at an optimal rate during the iteration process. This technique has proven quite useful in accelerating the convergence of the DR method.

Damping

The type and amount of damping can also significantly affect convergence. For linear systems, optimal damping depends on the both the highest and lowest eigenvalues of the system. Although bounds on the highest eigenvalue are readily available from the element eigenvalue inequality,

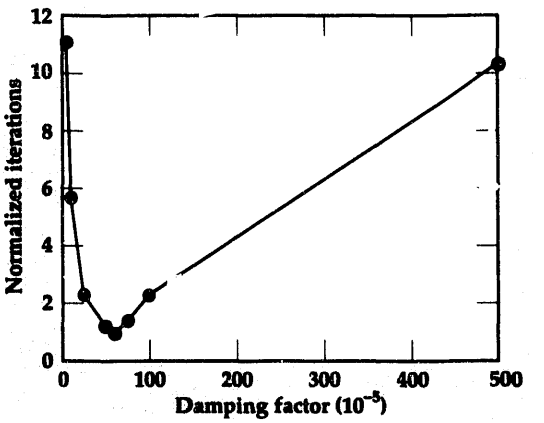


Figure 1. Variation in the number of normalized iterations required for convergence, relative to the magnitude of the damping factor in the dynamic relaxation method.

estimates must be used for the lowest eigenvalue, which can vary greatly throughout a nonlinear simulation. When insufficient damping is used, iterates will oscillate around a solution and reach it very slowly. Too much damping will dramatically retard convergence, especially for problems that include large rigid-body motions such as the motion of the sheet in sheet-forming simulations. Results thus far indicate that adaptive damping approaches, based on the evolving physics of the problem, may prove most effective for highly nonlinear problems. **Figure 1** shows the variation in the number of iterations required to converge relative to the magnitude of damping in the DR algorithm. The graph depicts the strong influence of damping value on the number of iterations required by the DR method to converge to a solution. The automated determination of the optimal damping value is a subject of ongoing investigation.

Cantilever Elastic Plate Example

To demonstrate some essential features of the new quasistatic solution capability in DYNA3D, an elastic cantilever plate was subjected to an applied moment on the free end. The problem was solved with two magnitudes of applied load: one

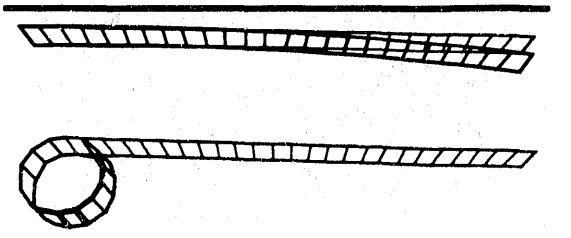


Figure 2. The initial and final deformed shapes of a cantilever plate subjected to an end moment. The upper figure corresponds to a small load, and the lower figure to a larger load. The solution to each of these problems was obtained using only one load increment with the DR method in DYNA3D.

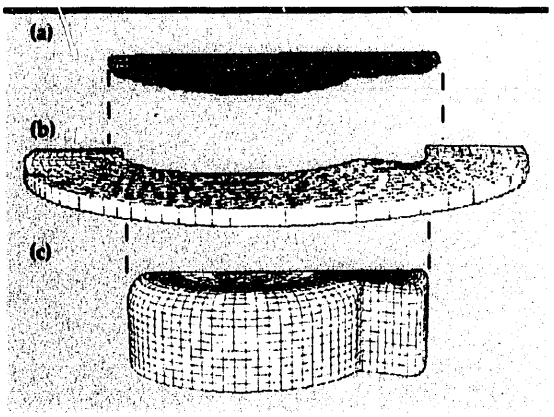


Figure 3. Initial geometry for hydroforming simulation, showing the punch, blankholder, and sheet. The punch and blankholder are geometrically represented by 8-node continuum elements and are treated as rigid bodies. The sheet is represented by 4-node thin shell elements and is modeled as an elastic-plastic material.

that causes a small deformation of the plate, and one that causes an extensive 'roll-up' deformation. The initial geometry and the two final deformed shapes are shown in **Fig. 2**. An interesting observation is that the DR algorithm required approximately the same number of iterations to converge for both load cases. This is in contrast to conventional implicit solution techniques, where the number of iterations required to converge increases quickly with the degree of nonlinearity. This insensitivity of DR to the degree of nonlinearity is a powerful advantage of the DR method.

Sheet Metal-Forming Example

One major application of the quasistatic solution capability developed in this research is the numerical simulation of sheet metal-forming processes. These problems pose a real challenge since they involve large strains, material nonlinearities such as plasticity, and extensive sliding contact with friction. In addition, the thin sheets have a wide spectrum due to the large difference between in-plane and bending stiffnesses, thus making them even more difficult for an iterative solver.

Figure 3 shows the finite element model for the numerical simulation of an aluminum hydroforming process. Pressure is applied to the upper surface of the sheet to hold it against the blankholder, and the punch is then advanced to form the sheet into the final shape shown in **Fig. 4**. The good comparison between the computed results and the shape of the actual part, including the failure locations, is illustrative of the power of a versatile quasistatic analysis tool.

This problem was first solved at LLNL in 1988 by running DYNA3D in an explicit dynamic analysis mode and applying the loads slowly to minimize dynamic effects, an approach requiring approximately two hours of CPU time. More recently, this problem was solved using the LLNL implicit code NIKE3D, but it required somewhat more computation time. Using the newly developed iterative methods in DYNA3D, this solution has been obtained in approximately 20 minutes of CPU time. Further improvements in contact algorithms, adaptive damping algorithms, and code optimization should enable solution of problems such as this in even less CPU time and without trial and error. Although much remains to be done, these initial results indicate the promise of the iterative quasistatic solution method in DYNA3D.

Future Work

Our research in FY-92 has led to the development and implementation of a DR iterative strategy for quasistatic problems in the LLNL DYNA3D code. Four general conclusions can be made from our experience thus far: (1) overdamping in the

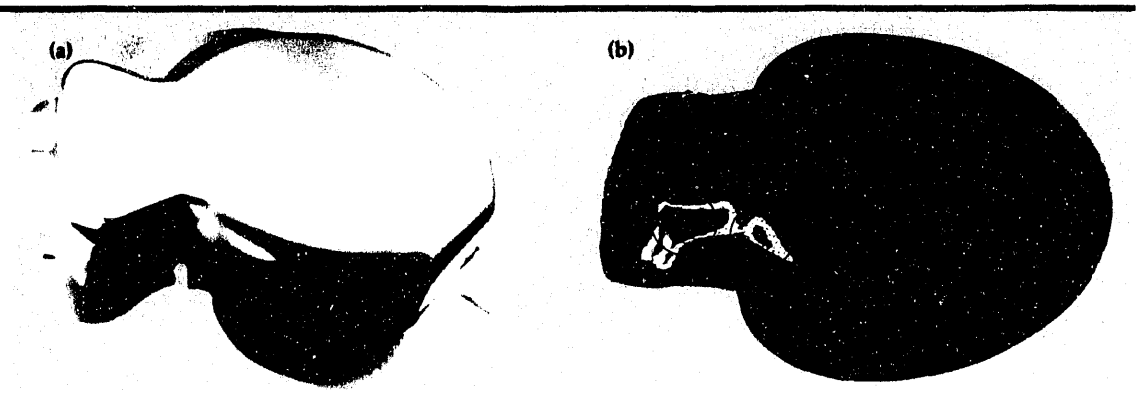


Figure 4. Comparison of actual deformed shape with that predicted by numerical simulation. The circled areas in the numerical results indicate regions of large strains, and these correspond closely with the tears observed in the real part.

DR method significantly slows the convergence rate, especially for problems with large rigid body motions; (2) the convergence rate of DR appears insensitive to the degree of nonlinearity in many problems; (3) the rate of load application within an increment is important, and a step function is probably not optimal; and (4) adaptive damping algorithms work extremely well for some problems, and are clearly desirable. More study and development will be required, however, before these algorithms can be used for general production analysis.

Our research efforts in FY-93 will explore the promising directions discussed above. We will refine adaptive damping DR algorithms and develop optimal load application schemes for a range of nonlinear quasistatic problems. We will also investigate new contact formulations to eliminate the solution noise introduced by the current penalty-based procedures. In addition, we will evaluate the utility of the nonlinear conjugate gradient algorithm for the problem classes of interest. Finally, the results of this effort will be optimized for vector computers and implemented into a future production version of the LLNL DYNA3D code for general use. In addition, the algorithms developed in this project will be implemented into the ParaDyn project to allow the solution of large quasistatic problems on massively parallel computers.

Acknowledgements

The authors wish to acknowledge Dr. Brad Maker of the LLNL Methods Development Group for sharing his early experiences on sheet forming with DYNA3D and for providing the finite element model and photographs for the sheet forming example.

1. B.E. Engelmann and J.O. Hallquist, *NIKE2D: A Nonlinear, Implicit, Two-Dimensional Finite Element Code for Solid Mechanics—User Manual*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-MA-105413 (1991).

2. B.N. Maker, R.M. Ferencz, and J.O. Hallquist, *NIKE3D: A Nonlinear, Implicit, Three-Dimensional Finite Element Code for Solid and Structural Mechanics—User Manual*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-MA-105268 (1991).

3. R.G. Whirley, B.E. Engelmann, and J.O. Hallquist, *DYNA2D: A Nonlinear, Explicit, Two-Dimensional Finite Element Code for Solid Mechanics—User Manual*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-MA-110630 (1992).

4. R.G. Whirley and J.O. Hallquist, *DYNA3D: A Nonlinear, Explicit, Three-Dimensional Finite Element Code for Solid and Structural Mechanics—User Manual*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-MA-107254 (1991).

5. R.M. Ferenez, *Element-By-Element Preconditioing Techniques for Large-Scale, Vectorized Finite Element Anaysis in Nonlinear Solid and Structural Mechanics*, Ph.D. Thesis, Stanford University, Palo Alto, California (1989).

Enhanced Enforcement of Mechanical Contact: The Method of Augmented Lagrangians

Bradley N. Maker
Nuclear Explosives Engineering
Mechanical Engineering

Tod A. Laursen
Duke University
North Carolina

We have introduced the method of augmented Lagrangians into our stress analysis codes, NIKE2D and NIKE3D. This approach provides a simple and effective enhancement to the penalty method for enforcing contact constraints. Also, by using augmented Lagrangians, accuracy is determined by physically motivated convergence criteria, independent of the penalty parameter.

Introduction

Contact between deformable bodies occurs commonly in mechanical systems. Stress analysis codes that are applied to multi-body systems and assemblies must accommodate this contact to be useful to design engineers. Our NIKE and DYNA finite element codes have a widely recognized capability to capture the mechanics of contact in complex systems, as the models in **Fig. 1** demonstrate. The results of this research effort have further enhanced our contact algorithms by introducing the method of augmented Lagrangians into NIKE2D and NIKE3D.

In the finite element method, bodies are discretized into assemblies of elements whose boundaries are described by a set of node points. In this context, mechanical contact conditions act to constrain the node points of one body from penetrating the boundary surface of another. **Figure 2** represents the discrete contact problem in two dimensions. Driven by the action of externally applied loads, a single node point from the 'slave' body penetrates the boundary of the 'master' body. This penetration is identified by a search algorithm, and activates the constraint enforcement algorithm. As the contact constraint is enforced, penetration is minimized, and stress and deformation are induced in each body. This deformation may cause other slave nodes to penetrate the master body, which in turn activates additional constraints. As this iterative process reaches equilibrium, the proper contact area and pressure distributions that balance the applied loads are obtained.

This simple example highlights the nonlinear nature of the contact problem. Indeed, the deformation of each body may be large, generating both geometric and material nonlinearities. But the more fundamental nonlinearity in the contact problem arises from the discontinuous manner in which the contact area evolves. Since the surfaces are faceted, the contact area grows or shrinks in discrete increments. These abrupt changes in contact area are sharp nonlinearities, which complicate the equilibrium search process.

Progress

The constraint algorithm used to minimize penetration in most finite element codes, including our own, is the penalty method. This simple but effective approach introduces penalty springs between the two bodies wherever penetration occurs. As the penetration increases, the springs are stretched, generating forces that oppose further penetration. The springs act unilaterally, i.e., when the bodies separate, the penalty springs are removed, allowing gaps to open.

One obvious drawback of the penalty method is that penetration must occur before any constraint forces are generated. Thus, in the equilibrium state, where each penalty spring is properly stretched to balance the applied loads, the two bodies are interpenetrated, and the exact contact condition is violated.

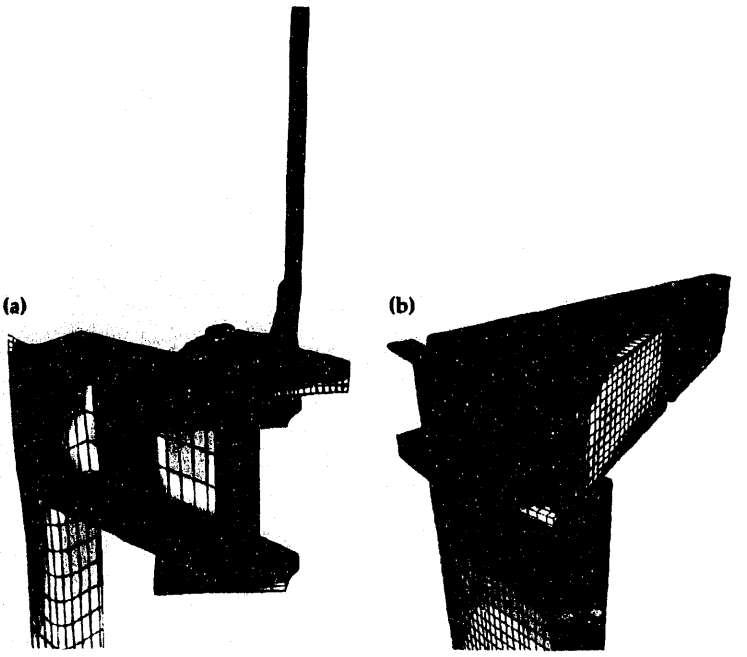


Figure 1. Examples of NIKE3D contact algorithms applied to engineering problems: (a) the bolted superflange and (b) the Kestrel bulkhead.

To minimize this penetration, the penalty spring stiffness may be increased, generating a large contact force through a very small penetration. This approach works well in theory, but in practice introduces poor numerical conditioning, and inevitably numerical errors. But a more fundamental deficiency of the penalty method is that the results are dependent upon the value chosen for the penalty stiffness. This effect is demonstrated in **Fig. 3**. Clearly, as a larger stiffness is chosen, the bodies are driven further apart, and the contact area and/or pressure changes. This arbitrariness motivated our work toward an enhanced constraint algorithm.

The augmented Lagrangian method is an effective and intuitively obvious enhancement to the penalty method, and proceeds as follows. Using the penalty method as a kernel, equilibrium is obtained in the usual manner. With known penetration depth and penalty stiffness, the contact force may be computed. This force is taken as the initial value for the Lagrange multiplier. The Lagrange multiplier defines a static load that is applied to the slave node, and the equilibrium search is then repeated. In the presence of the Lagrange multiplier load, penetration is reduced. The new penetration distance is then used to compute a new increment in contact force, the Lagrange multiplier is augmented by this increment, and the iteration process is repeated.

This equilibrium search and Lagrange multiplier augmentation loop proceeds until convergence is obtained. But now convergence may be defined in physically meaningful terms. For example, the augmentation loop can proceed until the contact force (Lagrange multiplier) stabilizes to within 1%, or until the largest penetration is less than a user-specified distance.

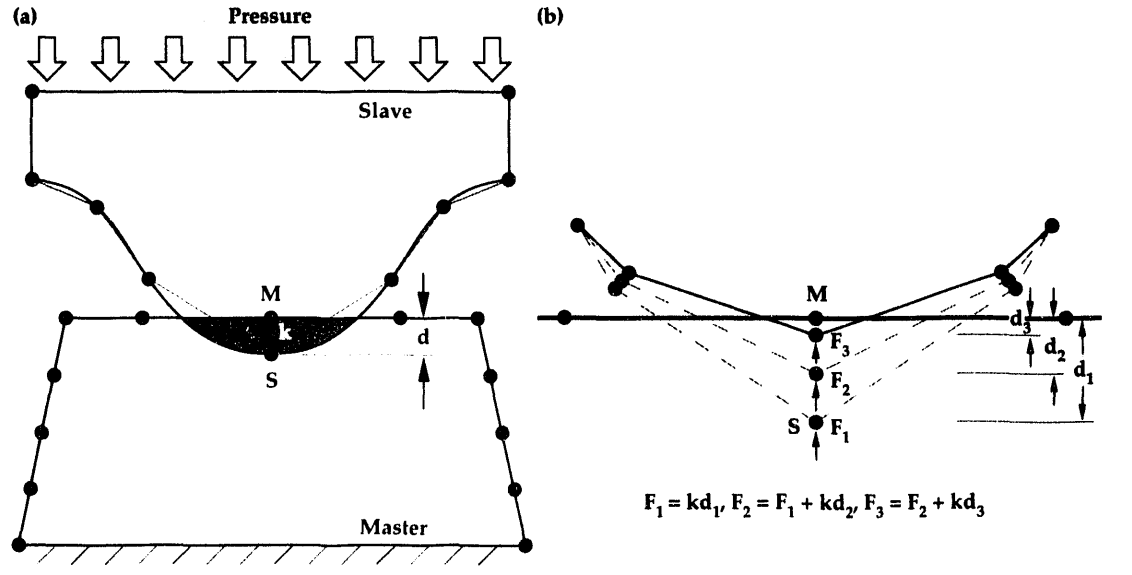


Figure 2. Enhanced enforcement of mechanical contact. (a) A search algorithm detects penetration of the master body by the slave node point. The penalty method introduces a spring of stiffness k between nodes S and M. When stretched, the spring generates interface force F = kd. (b) The augmented Lagrangian method applies F as a static force on nodes S and M, and iteratively augments this force, i.e., $F_{n+1} = F_n + kd_n$ until a convergence criterion is satisfied.

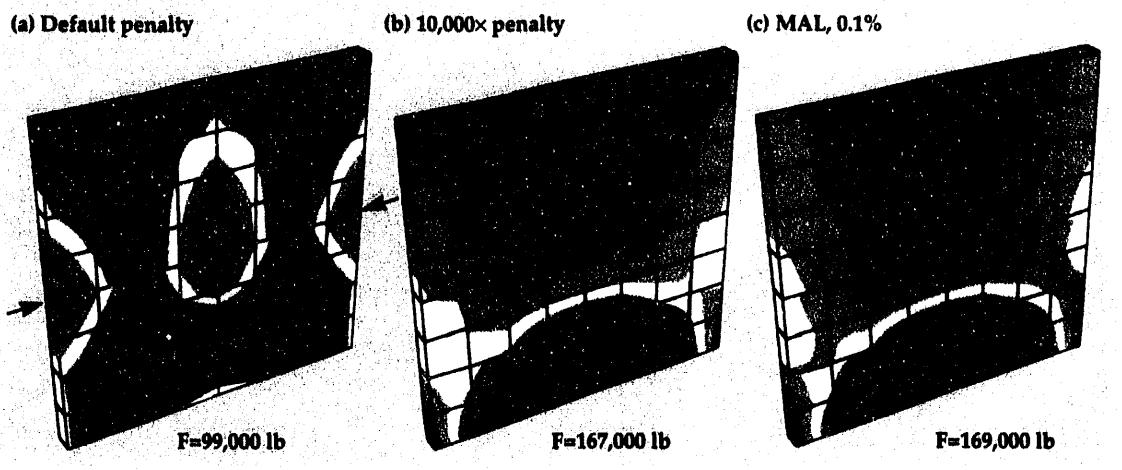


Figure 3. Using the penalty method, results vary dramatically with penalty stiffness. In this example, a contact interface is defined between two flat plates (arrow). The lower plate is fixed at its lower edge. A downward motion is prescribed to the upper edge of the upper plate. Erratic stress distributions result using NIKE3D's default penalty stiffness (a). Increasing the penalty stiffness by 10^4 produces a more uniform stress distribution (b). The augmented Lagrangian method gives the most accurate solution (c) using a convergence tolerance of 0.1% on the interface force. This same answer was obtained for penalty stiffnesses ranging from 10^{-2} to 10^4. The new method therefore provides an insurance policy against errors from a poorly chosen penalty parameter.

The new method has several advantages. In the limit of a large number of augmentations, equilibrium contact force is obtained without penetration. Further, the solution is independent of penalty parameter, since augmentations proceed until the (physically based) convergence criterion are satisfied. The exception to this independence is the case where the penalty stiffness is chosen so large that the original penalty method (the kernel of the new method) will not converge due to numerical conditioning. This case is obviously moot, since both methods fail.

The obvious drawback to the augmented Lagrangian method is that an additional iterative loop is introduced into the solution process. For a very soft choice of penalty parameter, this iteration loop can be slow to converge. However, our implementation allows for immediate convergence with no iteration if the penalty stiffness is cleverly (or luckily) chosen to satisfy convergence criteria in the first step. The method is therefore an insurance policy against a poor choice of penalty parameter, which before would have yielded an inaccurate result.

The final and perhaps most dramatic advantage to the new method is that the Lagrange multipliers are preserved for use in the next loading step. Thus, for a problem in which load is applied in several steps, the initial guess at contact pressure is the converged value from the previous step. This history information often speeds convergence of the equilibrium search in the second and later steps in the problem, and can result in an overall reduction in CPU run time for a complex problem.

The augmented Lagrangian method provides a simple and effective enhancement to the penalty method for enforcing contact constraints in NIKE2D and NIKE3D. Accuracy is determined by physically motivated convergence criteria, and is independent of the penalty parameter.

Future Work

The method of augmented Lagrangian also offers a new mathematical framework for considering the frictional contact problem, which will be pursued in future work.

Acknowledgement

We gratefully acknowledge the extensive collaboration of Dr. Bruce Engelmann in the algorithm development and NIKE2D implementation, and Messrs. M.A. Gerhard, D.J. Trummer, and E.A. Platt for the superflange and Kestrel examples.

ParaDyn: New Generation Solid/ Structural Mechanics Codes for Massively Parallel Processors

Carol G. Hoover
National Energy Research Supercomputer Center
Computation Directorate

James D. Maltby
Nuclear Test Engineering
Mechanical Engineering

Anthony J. De Groot
Engineering Research Division
Electronics Engineering

Robert G. Whirley
Nuclear Explosives Engineering
Mechanical Engineering

The objective of this work is to develop DYNA3D for massively parallel computers. In this last year, we have worked with the DYNA2D program on a Thinking Machines CM-5 computer to develop strategies for distributing the data and parallelizing the finite element algorithms. We are using the experiences gained with DYNA2D to guide the parallelization of the algorithms for the much larger and more complex DYNA3D. We have measured performances comparable to Cray Y-MP speeds for a DYNA2D test problem on systems with as many as 512 processors. The performance results show moderately large communication times relative to computing times, particularly for the global force assembly (scatter). We attribute this performance to the early developmental releases of the CM-5 software.

Introduction

Recent advances in microprocessor chip technology and parallel computer architectures are revolutionizing the concept of supercomputing. Vector supercomputer architectures have reached technology limits that preclude the orders-of-magnitude performance improvements expected for the massively parallel architectures.[1] A massively parallel computer is an arrangement of hundreds to thousands of microprocessors interconnected with a high-speed internal network (currently up to 250 megabytes/s). Typical microprocessor peak speeds range from a low of 10 MFLOPS per processor to a high of 100 MFLOPS per processor for pipelined (vector-like) processors. Performances between 10 and 100 GFLOPS are possible today on systems of 1000 processors. By comparison, the latest vector supercomputer is a 16-processor system with a peak performance of 1 GFLOP per processor. The motivation for developing a parallelized version of the solid mechanics programs (DYNA and NIKE) is the potential in the next three to five years for running applications that are larger by two or three orders of magnitude than are possible on vector supercomputers. This would allow simulations of hundreds of millions of elements rather than a few hundred thousand elements with DYNA3D.

Figure 1 illustrates the speed and storage requirements for typical advanced applications in metal forming, materials science, earthquake simulations, and crash dynamics. Notice in **Fig. 1** the increased complexity of the models for points in the upper right portion. These applications are of high value in government research and for their impact on industrial competitiveness.

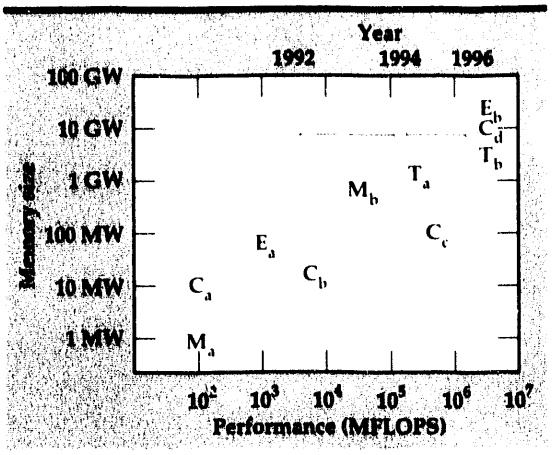


Figure 1. Advanced applications for massively parallel processors. The data represents systems as follows: E = earthquake simulations: (a) bridges, buildings and other structures, and (b) full Bay Area earthquake simulations. C = crash dynamics simulations: (a) automobile component simulations, (b) automobile/barrier simulations, (c) multiple automobile crash simulations, and (d) aircraft crash simulations. M = metal forming applications: (a) two-dimensional simulations, and (b) three-dimensional simulations. T = tribology and nanometer-scale machining simulations: (a) large-scale (10 to 100 million atoms) molecular dynamics simulations with no electronic structure (ab initio) calculations, and (b) hybrid molecular dynamics and continuum mechanics models with billions of particles/zones and electronic structure calculations with million particle molecular dynamics. Projected times for the next generations of systems are given along the top of the plot.

Progress

The DYNA3D program is nearly twice the size of DYNA2D, and the three-dimensional algorithms (e.g., for contact between slide surfaces) are more elaborate. Our strategy is to experiment with conversion techniques, parallel language paradigms, and algorithm parallelization with DYNA2D rather than with DYNA3D.

The development of a parallelized version of a large vectorized program necessarily proceeds in steps. The first and most tedious step is the conversion of array storage. The storage allocation for a distributed-memory massively parallel computer is dramatically different than for a common-memory serial computer. Careful analysis of reused storage, detailed conversion of array layout, and parameterization of the element vector block length absorbed well over one third of our effort. A benefit of this work and of the following timing analysis has been the insight we have gained into techniques for greatly reducing this same effort for the DYNA3D conversion.

The computationally expensive step in the DYNA algorithm is the element-by-element force evaluation. The vectorized version of this element processing translates readily into the data parallel paradigm on the CM-5. We have completed a data parallel version of the force update and time integration for

Table 1. Timing for the 7 cycles of an elastic/plastic bar impacting a rigid wall. There are 32,768 elements in the 64-×-512 mesh. Results are for a 512-processor CM-5.† The gather time is associated with the block processing for multiple material and element formulations. The scatter time is associated with the global force assembly step. The parallel reduction time accrues for calls to an intrinsic parallel library routine.

	June 1992	November 1992
Processor CPU time	29.3%	20.8%
Gather/scatter time	33.8%	24.5%
Parallel Reductions	0.8%	0.7%
Front-end to processor time	36.1%	54.0%
Totals	183.7 s	0.962 s
CM-5 elapsed time for 7 cycles:		2.03 s
Time per element-cycle (November 1992):		9 μs for the 512 processor CM-5 6 μs for the Cray Y-MP

The above results do not include the use of vector software. At the time of this printing our timing results for the CM-5 have improved to .7(2 μs per element time step.

† *Disclaimer by Thinking Machines Corporation. These results are based upon a test version of the software where the emphasis was on providing functionality and the tools necessary to begin testing the CM-5 with vector units. This software release has not had the benefit of optimization or performance tuning and, consequently, is not necessarily representative of the performance of the full version of this software.*

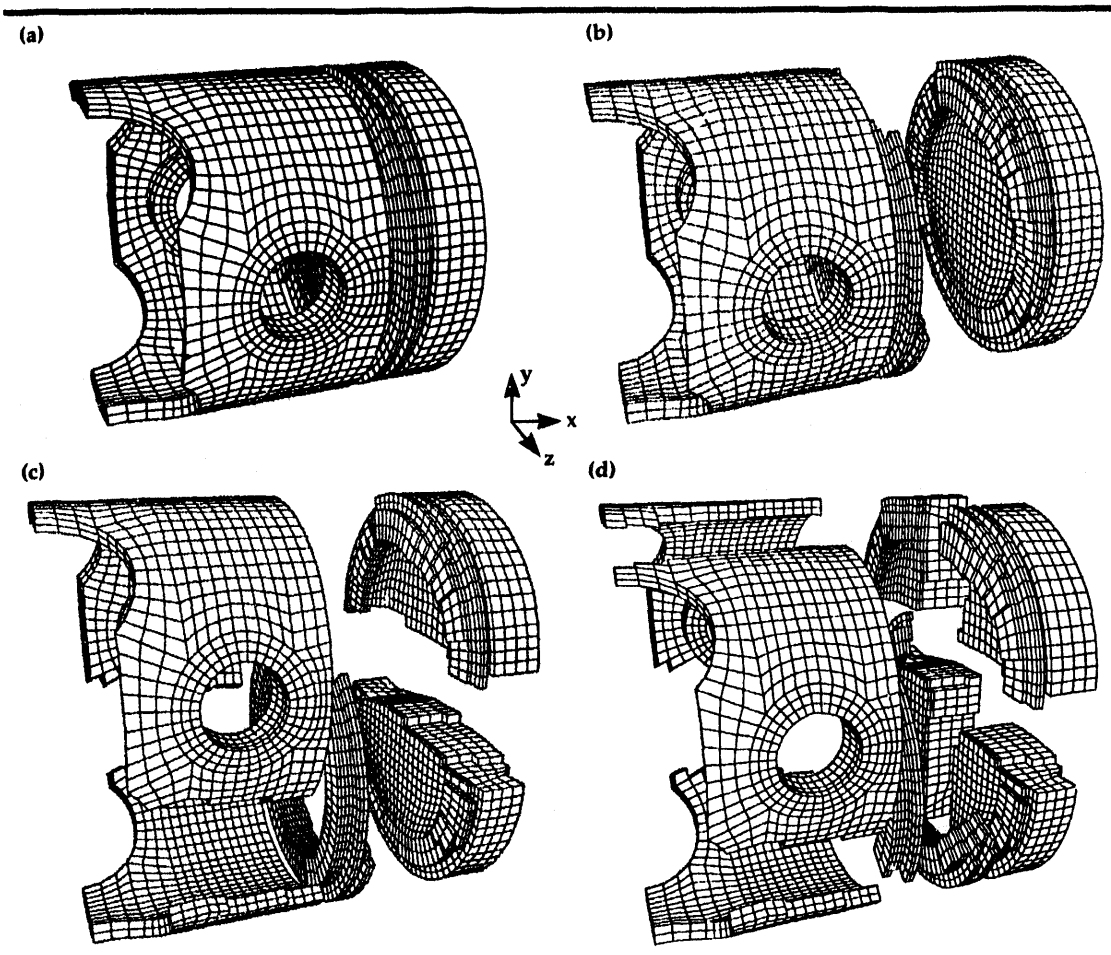


Figure 2. Partitioning of a standard three-dimensional mesh for an automobile piston; (a) the unstructured, three-dimensional finite element mesh, (b) the partitioning of the piston mesh for two processors, (c) the partitioning of the piston mesh for four processors, (d) the partitioning of the piston mesh for eight processors.

both elastic and elastic/plastic material models. We chose a standard test case, a bar impacting a rigid wall, for timing and performance analysis.

Balancing the parallel and scalar calculation time with network time is essential for efficient use of a massively parallel computer. An unbalanced problem with communication times exceeding calculation times prevents the desirable linear speedups predicted by Amdahl's law. On the CM-5, the performance analysis tool, PRISM, has been effective in providing the breakdown of hardware times. The most valuable feature of PRISM is the availability of timing data throughout the program, from upper level subroutines down to individual FORTRAN statements. Two timing analyses are shown in **Table 1**. The performance difference in the two runs is a combined effect of hardware/software changes at the Army High Performance Computing Research Center and several programming changes inspired by the PRISM statistics. The speed achieved for our most recent run is 9 microseconds per element-cycle, which is comparable to the one processor Y-MP performance of 6 microseconds per element-cycle.

The communication times we have measured are still excessively high for a balanced calculation. We have collaborated with computational analysts at Thinking Machines to analyze the imbalance in the timings for this test problem. The development of system software such as compilers and communication libraries for massively parallel systems is in its infancy. We find that the newer alpha-test versions of the software, used now by the company analysts, will change these results up to an order of magnitude. This software may be available to us within the next three to six months. With the new versions of the software, we expect to exceed single processor Cray C-90 performances for a single material problem with a regular topology. At the time of this printing our timing results for the CM-5 have improved to .732 μs per element time step.

Several techniques have been developed for balancing the computational work among processors while minimizing the communication time. We are testing a recursive spectral bisection technique[2] with three-dimensional meshes. We have developed a method for visualizing the results, as shown in **Fig. 2**.

Future Work

We will continue to use DYNA2D to experiment with algorithm parallelization. In this next year, we plan to investigate: (1) message-passing and data-parallel versions of selected contact algorithms, (2) a data-parallel and message-passing hybrid system software model available in the next year from Thinking Machines, and (3) parallel table lookup and sort algorithms, which are appropriate for contact algorithms. We will begin the conversion and parallelization of DYNA3D and develop kernel algorithms for DYNA3D for further evaluation of parallel programming paradigms and architectures.

DYNA3D is an 80,000 line analysis program including ten finite element formulations (solid elements, shells, and beams), 35 material models, equations of state for hydrodynamic models, several algorithms for contact at arbitrary interfaces, and a list of additional boundary conditions and mechanics algorithms, all of which make the program one of the most widely used tools for nonlinear structural response simulations. Our plan over the next three years for demonstrating a prototype massively parallel version of DYNA3D includes implementing an eight-node solid (continuum) element, at least one shell element, one contact algorithm, and solid/structural mechanics capabilities needed for three large-scale demonstration problems. The demonstration problems include the simulation of a nanoindentation problem, an automobile/barrier simulation, and a weapons penetration application.

Acknowledgement

We gratefully acknowledge the Army High Performance Computing Research Center for providing CM-5 computer time for this work as part of their General Plan for Developing Structural Analysis Programs for Advanced Massively Parallel Computers. Funding for computer time was supported by, or in part by the Army Research Office contract number DAAL03-89-C-0038 with the University of Minnesota Army High Performance Computing Research Center. We thank Earl Renaud from Thinking Machines Corporation for his advice and cooperation.

1. B. Bhoghosian, *Comput. Phys.* **4**, 1 (1990).
2. H.D. Simon, *Computing Systems in Engineering* **2** (2/3), 135 (1991).

Composite Damage Modeling

Edward Zywicz
Nuclear Explosives Engineering
Mechanical Engineering

A progress damage model for continuously reinforced, polymeric-matrix composites is being developed and implemented in the implicit finite element code NIKE3D. The constitutive model replicates the discrete laminae with an equivalent homogenized material prior to the onset of damage. Failure criteria eventually trigger damage evolution laws that track individual failure mechanisms within each lamina and degrade the stiffness and strength of the laminated composite. Failure criteria and damage evolution laws are currently being developed, as well as numerical procedures, to efficiently address the multilayer nature of laminates. This work will allow analysts to simulate the redistribution of load as the composite materials degrade and, therefore, to design minimal mass composite structures.

Introduction

Continuously reinforced, polymeric-matrix composites offer substantial weight savings over conventional materials, such as steels and aluminums, and at the same time provide equal or superior mechanical properties. For example, at Lawrence Livermore National Laboratory, continuously reinforced graphite/epoxy (Gr/Ep) composites are used in lightweight earth-penetrator weapons, advanced conventional munitions, and enhanced nuclear safety systems. Commercial applications of Gr/Ep composites include high-speed aircraft, automobile drive shafts, bicycles, and tennis rackets. Currently, components manufactured with continuously reinforced, polymer-based composites are designed very conservatively or must be tested extensively, because the failure response of the material is not fully understood. To overcome this barrier, a composite damage model is currently being developed and implemented in the implicit finite element code NIKE3D.[1]

A progress composite damage model permits analysts to simulate the complex three-dimensional (3-D) response of composite components in both subcritical (e.g., dings in an aircraft wing) and catastrophic (e.g., car crashes) loading environments. To be useful, the damage model must accommodate a wide range of fiber layups, track damage evolution based upon individual failure mechanisms, and predict residual life and strength. At the same time, the material model must be numerically efficient and resolve the complex lamina behavior within each laminate region modeled.

Progress

A continuum-based framework has been assembled to represent composite behavior. The approach uses conventional 8-node, solid isoparametric, 3-D elements with conventional $2 \times 2 \times 2$ Gaussian quadrature. Element stresses and stiffness are calculated in the usual way at each Gaussian point; however, the constitutive evaluations use homogenized material properties that are calculated uniquely for each element. During initialization, the 'virgin' elastic properties of all laminae present within each element, which can vary between one and two hundred, are homogenized and stored along with element-level, strain-based, failure criterion coefficients. Throughout the analysis, the small-strain, finite-deformation (total-Lagrangian)-based constitutive relation continually updates the element stresses, using effective stiffnesses and monitors for failure initiation, prior to the onset of damage.

Element-level failure triggers an in-depth lamina level or microanalysis. The microanalysis checks for failure, and tracks and evolves individual damage mechanisms for each lamina present in the element. Furthermore, it degrades

the individual lamina stiffnesses and calculates a material tangent matrix. The updated stiffnesses and tangent matrices are then homogenized for use at the element level.

The two-tier homogenized approach provides a rational and precise mechanism for tracking and integrating the complex response of damaged laminated composites. For undamaged material points, the homogenization technique, which incorporates bending and coupling effects, yields accurate solutions at a substantial computational savings, since laminate integration is performed only in initialization. Traditional methods use single elements or integration points for each lamina or homogenized material property and neglect bending and coupling effects. Efficiency in the undamaged region is very important since, in general, only small regions of typical composite components reach critical load levels. To date, the element homogenization technique, a conservative, element-level, strain-based failure criteria for fiber-direction strain to failure, and a micro-lamina-level subintegration scheme have been developed, implemented, and verified.

Element Homogenization and Representation

Homogenized stiffness functions[2] are used in the element to represent the total sub-laminate response. Within an element, the homogenized local stiffness $C^{II}(z)$ is given by

$$C^{II}(z) = C_0^{II} + C_1^{II} z + C_2^{II} \cos(\alpha z), \qquad (1)$$

where C_0^{II}, C_1^{II}, and C_2^{II} are element-based stiffness matrices, z is the normalized distance from the element's central plane, and α is a constant used to minimize element integration error. To determine C_0^{II}, C_1^{II}, and C_2^{II}, all laminae present in an element are identified. Next, using the closed-form long wavelength solution of Pagano,[3] the current lamina stiffnesses are 'integrated' through the element thickness, yielding the effective extensional (**A**), coupling (**B**), and bending (**D**) matrices of the element. This approach treats each element as a unique sublaminate. Using the same long wavelength procedure, **Eq. 1** is integrated. The resulting extensional, coupling, and bending matrices are equated with the previous ones and manipulated to yield C_0^{II}, C_1^{II}, and C_2^{II} directly in terms of the actual lamina properties and local geometry.

This approach, as we have noted,[2] ensures identical net mid-surface forces and moments between any two systems for a specified displacement field. The kinematics assumed in the effective long wavelength solution are, with the exception of the through thickness shear strains, identical to an 8-node, rectangular, isoparametric brick element for small strains. Therefore, the finite element solution reflects the same behavior assumed in the homogenization. Thus, the effective properties represent precisely the varying lamina orientations, and the stacking sequence relates behavior, i.e., the bending and coupling responses.

Equation 1 allows approximate, but very accurate, element-level integration with conventional Gaussian quadrature. With $\alpha = 0.25$, the maximum normalized error in any single stiffness or force term is less than 8.4×10^{-3}. Although smaller values of α reduce the integration error, they introduce other undesirable numerical problems. By restricting individual laminae to be orthotropic, only 19 coefficients per pair of Gaussian points are necessary to describe $C^{II}(z)$, independent of the number of laminae present.

Failure Criteria

Accurate, strain-based, element-level failure criteria minimize computational costs by postponing, as long as possible, the use of expensive microlevel analysis. Criteria must be conservative to ensure that failure initiation is not missed within any of the sublaminae, and thus, a criterion is needed for each failure mechanism.

Laminate strengths are inherently limited by the extreme stresses and strains that the individual fibers and matrix can sustain. Since fibers are typically brittle, one convenient and commonly used criterion bases damage initiation or failure upon the minimum and maximum fiber direction strains.[4] A conservative, element-level failure criterion based upon fiber direction strains was formulated. Tensile and compressive failure initiates when

$$a\left\{\max\left(0, \frac{\varepsilon_{11}}{\varepsilon_f^t}\right)\right\}^2 + b\left\{\max\left(0, \frac{\varepsilon_{22}}{\varepsilon_f^t}\right)\right\}^2 + \left\{\max\left(c_1 \frac{\varepsilon_{12}}{\varepsilon_f^t}, c_2 \frac{\varepsilon_{12}}{\varepsilon_f^t}\right)\right\} = \frac{4\beta - 1}{4\beta^2} \qquad (2)$$

and

$$a\left\{\min\left(0, \frac{\varepsilon_{11}}{\varepsilon_f^c}\right)\right\}^2 + b\left\{\min\left(0, \frac{\varepsilon_{22}}{\varepsilon_f^c}\right)\right\}^2 + \left\{\min\left(c_1 \frac{\varepsilon_{12}}{\varepsilon_f^c}, c_2 \frac{\varepsilon_{12}}{\varepsilon_f^c}\right)\right\} = \frac{4\beta - 1}{4\beta^2}, \qquad (3)$$

respectively, where

$$a = \max\left\{\cos^2\theta_i\right\} \quad (4)$$

$$b = \max\left\{\sin^2\theta_i\right\} \quad (5)$$

$$c_1 = \max\left\{\left[\max\left(0, \frac{2\cos\theta_i \sin\theta_i}{\alpha}\right)\right]_i\right\} \quad (6)$$

$$c_2 = \min\left\{\left[\min\left(0, \frac{2\cos\theta_i \sin\theta_i}{\alpha}\right)\right]_i\right\}. \quad (7)$$

In these expressions, ε_{11}, ε_{22}, and ε_{12} are the in-plane strains, expressed in the element's natural coordinate system; ε_f^t and ε_f^c represent the composite's tensile and compressive fiber-direction strains to failure, respectively; and θ_i is the angle between the fiber direction (in the i-th lamina) and the 1-axis of the elements' natural coordinate. The value of β positions the failure surface and is bounded by $1/2 \le \beta \le 3/2$. For arbitrary layups, the optimal value is $\beta = 0.785$. **Equations 4** through **7** are evaluated using all lamina present within the element. In 8-node brick elements, the failure criteria, **Eqs. 2** and **3**, need be checked on only the 'upper' and 'lower' element surfaces.

Internally Pressured Thick-Walled Cylinder

To demonstrate the new model's ability to predict the elastic response of a laminated composite material, a thick-walled composite cylinder subjected to an internal pressure was analyzed. The cylinder was axially constrained and had an inside-to-outside ratio of $r_i/r_0 = 3/4$. There were 72 Gr/Ep plies randomly oriented with their fibers in either the axial (0°) or hoop (90°) direction. The 3-D quarter model used, shown in **Fig. 1**, contained only 60 elements, i.e., 12 circumferentially and 5 radially oriented. Radial displacements on the inner [$\delta r(r = r_i)$] and outer [$\delta r(r = r_0)$] surfaces were compared to a baseline solution and are listed in **Table 1**. The baseline solution was obtained with an axisymmetric, two-dimensional model that contained 1152 elements in the radial direction. It modeled each lamina with 16 elements. Calculations were performed with and without incompatible modes. Overall, there is excellent agreement between the axisymmetric baseline and the 3-D homogenized solutions.

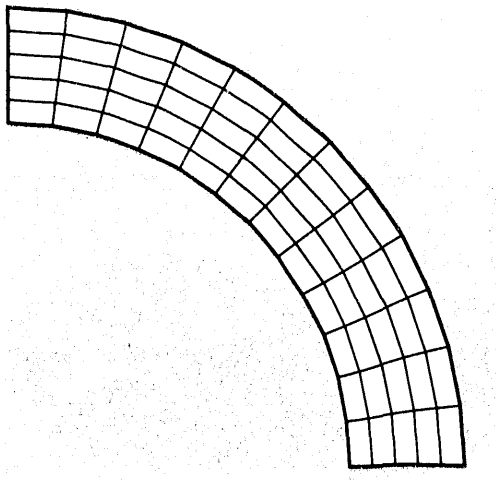


Figure 1. 3-D finite element discretization of the internally pressurized thick-walled cylinder. A quarter section is represented here.

	$\delta r(r = r_i)$	$\delta r(r = r_0)$
Baseline	1.112×10^{-4}	8.294×10^{-4}
3-D	1.112×10^{-4}	8.297×10^{-4}
3-D*	1.110×10^{-4}	8.260×10^{-4}

Table 1. Baseline and 3-D calculated radial displacements. *With 'incompatible modes' turned on.

Future Work

Additional failure mechanisms are necessary before component responses can be realistically tracked beyond initial failure. This requires that an element-level failure criterion as well as damage evolution relationships be developed for each mechanism. The lamina-level constitutive laws must ensure solution convergence with mesh refinement, include all mechanisms, and permit interaction between the various modes. In the immediate future, development efforts will focus on the predominant failure modes, namely, tensile and compressive failure, delamination, and in-plane shear failure. Development of the evolution laws will use both micromechanical models and non-traditional experimental results.

1. B.N. Maker, *NIKE3D User's Manual*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-MA-105268 (1991).
2. E. Zywicz, *Int. J. Num. Meth. Eng.* **35**, 1031 (1992).
3. N.J. Pagano, "Exact Moduli of Anisotropic Laminates," *Mechanics of Composite Materials* **2**, Academic Press (New York, New York), 23 (1974).
4. R.M. Christensen, *J. Compos. Mater.* **22**, 948 (1990)

HYDRA: A Flow Solver for Three-Dimensional, Transient, Incompressible Viscous Flow

Mark A. Christon
Nuclear Explosives Engineering
Mechanical Engineering

This article describes the current effort to develop a high-performance flow solver for addressing the incompressible class of complex-geometry transient flow problems that require very-high-resolution meshes. The code development effort is described in terms of the algorithm-to-architecture mapping issues involved in both vector and parallel supercomputers. An example problem showing the application of the current code to a streamline submarine hull is presented to demonstrate the class of problems being considered.

Introduction

This work is part of a collaborative effort involving the Mechanical Engineering and Physics Departments, and Military Applications at Lawrence Livermore National Laboratory (LLNL). The development of a high-performance, three-dimensional, transient, incompressible, viscous flow code is being undertaken primarily to study submarine performance in a fluid dynamics sense. The effects of flow separation and vorticity upon vehicle lift, drag, and ultimately steering, are of primary interest. The final goal is to provide a design simulation tool that will help to reduce the costly submarine design cycle.

While this effort addresses one of the National Grand Challenges of Computing, simulating flow fields about vehicles and in turbomachinery, this computational fluid dynamics (CFD) capability is unique because it also finds application within multiple divisions at LLNL, the Department of Energy, and in U.S. industry. Applications include the study of casting processes, heavy gas dispersion, and flow in the planetary boundary layer. There is also immediate application in industries critical to U.S. competitiveness in the world economy, such as the automotive industry where CFD is being used to augment engineering design in the areas of vehicle aerodynamics, heating and air conditioning, and engine and underhood cooling.

For the full-body, transient flow simulation of a submarine, it is anticipated that upwards of one million elements will be required to resolve important flow-field features such as vortices and regions of separated flow, which directly influence vehicle lift and drag. In addition to the high degree of spatial discretization, the temporal resolution is also demanding, requiring the efficient mapping of the flow-solution algorithm to current vector and parallel supercomputer architectures to make such simulations practicable.

For the solution of the time-dependent Navier-Stokes equations with complex geometry, it is estimated that computers with memory sizes of 1000 to 10,000 million words and performance rates of 10 to 1000 GFLOP's (1 GFLOP = 1 billion floating point operations per second) will be required. Today, a fully configured CRAY C-90 vector supercomputer provides a peak performance rate of 16 GFLOP's with a memory size of 256 million words. In contrast, a fully configured parallel computer such as the Thinking Machines CM-5 provides a peak performance rate of 120 GFLOP's with 4096 million 64-bit words of memory. By focusing on the rapidly evolving parallel computing platforms and making use of advanced numerical algorithms, the goal of rapid simulation of complex geometry flow simulations for design may be achievable in the near future.

Progress

The current finite element code for solving the Navier-Stokes equations is based primarily upon the work of Gresho *et al.*,[1-4] making use of ad-

vanced solution algorithms for both implicit and explicit time integration. In the case of the second-order fractional step algorithm[3,4] (implicit), a consistent-mass predictor in conjunction with a lumped mass corrector legitimately decouples the velocity and pressure fields, thereby reducing both memory and CPU requirements relative to traditional, fully coupled solution strategies. The consistent-mass predictor retains phase speed accuracy, while the lumped mass corrector (projection) maintains a divergence-free velocity field. Both the predictor and the corrector steps are amenable to solution via direct or preconditioned iterative techniques, making it possible to tune the algorithm to the computing platform, i.e., parallel, vector, or superscalar. The second-order projection algorithm can accurately track shed vortices, and is amenable to the incorporation of either simple or complex (multi-equation) turbulence submodels appropriate for the driving applications.

The explicit solution algorithm[1,2] sacrifices some of the phase-speed accuracy of the fractional-step algorithm for the sake of minimizing memory and CPU requirements. However, the momentum equations are still decoupled in the explicit algorithm. While both the diffusive and Courant stability limits must be respected in the explicit algorithm, balancing tensor diffusivity[1,2] is used to lessen the restrictive diffusive stability limit in the explicit algorithm. This, in combination with single-point integration and hourglass stabilization, makes the explicit algorithm very efficient computationally, and because of this, the explicit algorithm was chosen as the initial focus of the parallelization effort.

The fractional step and the explicit algorithms both rely upon the implicit solution of a linear system arising from an elliptic operator. In the case of the fractional step algorithm, this solution is used to project an intermediate velocity field to a divergence-free space. In the explicit solution strategy, the elliptic operator appears in the pressure Poisson equation, which is used to advance the pressure field in time. Because the linear system solver is a key component of the algorithm-to-architecture mapping for both algorithms, it has been necessary to develop modified, conjugate-gradient iterative solvers[5,6] that minimize the impact on memory requirements and allow the natural data parallelism of element-level processing to be exploited. For both the explicit and the projection algorithms, no additional storage is required for the elliptic operator itself, making the current conjugate gradient solver essentially matrix-free.

During the past year, our efforts have been directed primarily towards the vector and data-parallel or SIMD (Single Instruction Multiple Data) implementation of the code suitable for the laminar flow regime. In the case of the vectorized version of the code, element operations are blocked into groups of contiguous, data-independent operations by using a simplified domain-decomposition strategy to group the elements. This approach results in a code that is completely vectorized, yielding performance comparable to DYNA3D for the time integration of the momentum equations. However, the solution to the pressure Poisson operator limits the overall performance of the code, taking up to 95% of the CPU time per simulation time-step in problems with strong pressure–velocity coupling.

Because the element data structures for the vectorized version of HYDRA are adjustable, they are also used for the SIMD (CM-2/CM-5) implementation, where element-level operations are performed in a lock-step parallel fashion. For the CRAY architecture, the vector block size is configured as 128 (twice the length of the vector pipeline). In the case of the CM-2/CM-200, the element block size is configured as a multiple of the minimum virtual processor ratio (4) and the number of available processing elements. For the CM-5, the block size is configured as an integral multiple of the processor vector pipeline length and the number of available processors enabling processor pipelined operations in conjunction with SIMD parallelism.

In the SIMD (CM-2/CM-5) version of the code, data dependence in the element blocks may be resolved using hardware-specific communication/combining operations for the parallelized assembly of element data to nodal data. Instead of data dependency, the constraint on domain decomposition in the SIMD implementation requires that the elements be grouped in a spatially contiguous manner to minimize the deleterious effects on performance of off-processor communication. However, because the same data structures are used for the vector and SIMD versions of HYDRA, it is possible to appropriately reconfigure the block size for each architecture for the sake of performance. In effect, the element grouping strategy provides a mechanism to account for variations in granularity across supercomputer architectures ranging from vector to SIMD to Multiple Instruction Multiple Data (MIMD).

Many alternative domain-decomposition algorithms are available, including methods that consider the graph of the finite element mesh[7] when subdividing the physical domain, and are not restricted to logically regular meshes. By matching the domain decomposition strategy to the supercomputer architecture, it will be possible to maintain optimal performance on vector, SIMD, and MIMD machines.

HYDRA has been written using standard, UNIX

software-development tools, enabling the code to be simultaneously developed in FORTRAN 77 and FORTRAN 90 by making use of compile time configuration of the software. This approach has made it possible to provide HYDRA in a form suitable for computing platforms including workstations, and CRAY vector and Thinking Machines SIMD supercomputers. The top-down design and bottom-up implementation have required the design of a memory management package that makes it possible to perform dynamic memory allocation on a single processor workstation, multi-processor CRAY, or on the processors of the CM-2/CM-5 with a single interface definition.

Application

Currently, initial calculations are being performed on a range of simplified submarine-hull configurations. The top frame in **Fig. 1** illustrates the mesh used in the computation of the flow field around a streamline submarine hull at a Reynolds number of 830, based on the hull diameter. A 1/4 symmetry model has been used, resulting in a mesh with 18,000 nodes (16,000 elements) or 72,000 degrees of freedom. Tow-tank conditions were imposed on the computational domain to simulate the case when the vehicle is moving straight ahead.

Isosurfaces of pressure are shown in the middle frame of **Fig. 1** for the initially divergence-free and potential flow field. At the leading edge of the vehicle, a stagnation point is apparent, with the pressure decreasing in the streamwise direction along the hull of the vehicle. Near the trailing edge of the submarine, a low pressure 'bubble' is present due to the acceleration of the fluid as it tries to turn and follow the streamline surface of the hull. In the bottom frame of **Fig. 1**, isosurfaces of the x-velocity are shown. At the inlet to the domain and the far-field boundary, a uniform x-velocity has been imposed to simulate tow-tank conditions. The bubbles at the front and back of the straight section of the hull correspond to locations where the fluid has been accelerated to track the contours of the vehicle hull. At the surface of the hull, no-slip boundary conditions have been imposed.

Future Work

Future efforts will address two key issues: the acceleration of the solution to the linear system, arising from the pressure Poisson operator; and the inclusion of the recursive, spectral domain-decomposition strategy for SIMD architectures. The pressure computation currently relies upon a data-parallel, element-by-element diagonally scaled, matrix-free conjugate gradient solver. While this approach offers scalability in terms of the mesh resolution, multigrid acceleration can provide enhanced convergence rates by effectively damping the low-mode error components via coarse grid corrections. It also fits well in the current parallel-code architecture.

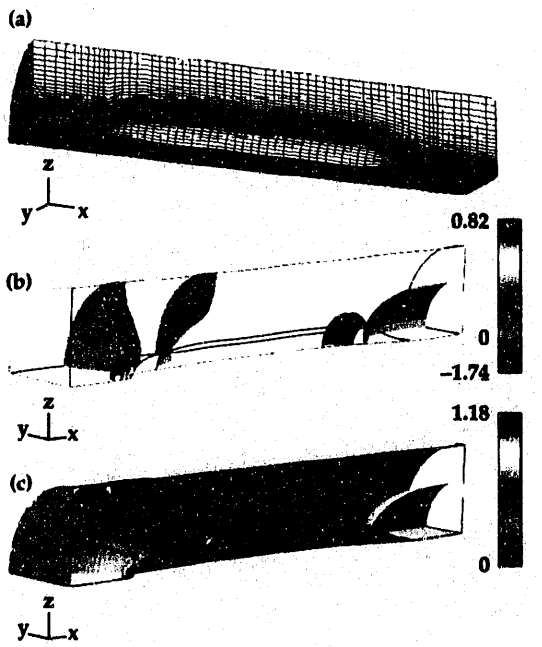


Figure 1. Results of calculations performed on simplified submarine-hull configurations, showing (a) the mesh used in the computation of the flow field; (b) isosurfaces of pressure; and (c) isosurfaces of the x-velocity.

While the current, vector-blocking, domain-decomposition algorithm is adequate for vector supercomputers, robust decomposition techniques yielding element-to-processor assignments that minimize communications overhead are necessary to achieve peak performance rates on both SIMD and MIMD parallel architectures. The recursive spectral bisection[7] algorithm, which uses the second eigenvector of the mesh connectivity graph, is currently being investigated as a candidate for performing domain decomposition.

1. P.M. Gresho, S.T. Chan, R.L. Lee, and C.D. Upson, *Int. J. Numer. Methods Fluids* **4**, 557 (1984).
2. P.M. Gresho, S.T. Chan, R.L. Lee, and C.D. Upson, *Int. J. Numer. Methods Fluids* **4**, 619 (1984).
3. P.M Gresho, *Int. J. Numer. Methods Fluids* **11**, 587 (1990).
4. P.M Gresho and S.T. Chan, *Int. J. Numer. Methods Fluids* **11**, 587 (1990).
5. F.S. Beckman, "The Solution of Linear Equations by the Conjugate Gradient Method," *Mathematical Methods for Digital Computers* **1**, 62 (1965).
6. R.M. Firencz, *Element-by-Element Preconditioning Techniques for Large-Scale, Vectorized Finite Element Analysis in Nonlinear Solid and Structural Mechanics*, Ph.D. Thesis, Stanford University, Palo Alto, California (1989).
7. H.D. Simon, "Partitioning of Unstructured Problems for Parallel Processing," *Computing Systems in Engineering* **2** (2), 135 (1991).

Development and Testing of the TRIM3D Radiation Heat Transfer Code

James D. Maltby
Nuclear Test Engineering
Mechanical Engineering

We have developed a new code, TRIM3D, to solve radiative heat transfer problems involving a participating medium. The code uses a Monte Carlo formulation to solve problems with absorption, anisotropic scattering, and specular boundaries. It is designed to work with other codes to solve coupled radiation/conduction/thermal stress problems, and has been verified against known analytic solutions.

Introduction

Radiation heat transfer problems involving a semi-transparent medium that participates in the radiative exchange occur often in areas such as high-power optics, crystal growth and glass manufacture, coal furnaces, annealing ovens, and analysis of fuel fires. Unfortunately, these problems are often difficult to solve due to the complex nature of the radiative transfer equation. A Monte Carlo approach to radiation heat transfer problems without a participating medium has proved very successful, resulting in the computer code MONT3D.[1,2] The objective of this research was to develop a Monte Carlo code to analyze radiation heat transfer in the presence of a participating medium.

The resulting code, TRIM3D, represents the state of the art for radiation heat transfer analysis, and is also the first production code with detailed participating medium capability. The addition of TRIM3D to our code suite allows the solution of coupled radiation/conduction/thermal stress problems with a level of detail not previously attainable.

Progress

During FY-92, a working version of the computer code TRIM3D was developed and given preliminary testing and verification. The TRIM3D code is formulated in a similar manner to the successful MONT3D non-participating medium heat transfer code used by programs at Lawrence Livermore National Laboratory (LLNL). The current working version of TRIM3D solves three-dimensional (3-D) radiation heat transfer problems in absorbing, emitting, and anisotropically scattering media. Problems may be solved that are non-homogenous and non-isothermal, and material properties may vary with wavelength. Boundaries may be diffuse, specular, or mixed, with directional reflectivity and transmissivity.

The code has been verified against a series of analytic problems with absorbing or scattering media and specular boundaries, with agreement within the statistical accuracy of the simulation. Currently, no other code exists that can handle participating media problems of this complexity.

Theoretical Formulation

TRIM3D generates a matrix of direct exchange areas (DEA's) that describe the radiative interaction among all surfaces and volumes in an enclo-

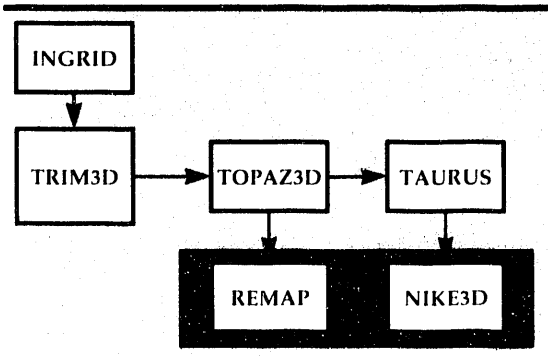


Figure 1. TRIM3D code flow. Temperature output from TOPAZ3D is passed through REMAP to NIKE3D for solution of radiation/conduction/thermal stress problems.

sure. The net exchange between any two surfaces or volumes may be described:

$$Q_{ij} = \sigma\, s_i\, s_j \left(T_i^4 - T_j^4\right) \quad \text{(surface to surface)}$$
$$= \sigma\, s_i\, g_j \left(T_i^4 - T_j^4\right) \quad \text{(surface to volume)}$$
$$= \sigma\, g_i\, g_j \left(T_i^4 - T_j^4\right) \quad \text{(volume to volume).}$$

This matrix is then passed to TOPAZ3D for solution of the coupled radiation/conduction problem. Since the matrix of DEA's is temperature-independent, boundary conditions and temperatures in an analysis may be changed without re-running TRIM3D. This approach has been very successful with MONT3D and TOPAZ3D, resulting in a large savings in computer time. The temperature output from TOPAZ3D may then be passed through REMAP to NIKE3D for solution of radiation/conduction/thermal stress problems. The code flow during the solution of such a problem is shown in **Fig. 1**. The mesh generator INGRID and post-processor TAURUS are also used.

TRIM3D simulates thermal radiation by emitting a large number of monoenergetic photons from each surface and volume. These photons are traced through multiple reflection and/or scattering events until they are absorbed in another surface or volume. The DEA's are then calculated from these photon tallies. For a given row of the DEA matrix,

$$s_i\, s_j = A_i\, \varepsilon_i\, N_{ij} / N_i$$
$$g_i\, s_j = 4\, V_i\, a_i\, N_{ij} / N_i\ ,$$

where N_{ij} is the number of photons emitted by element i and absorbed by element j, and N_i is the number emitted by element i.

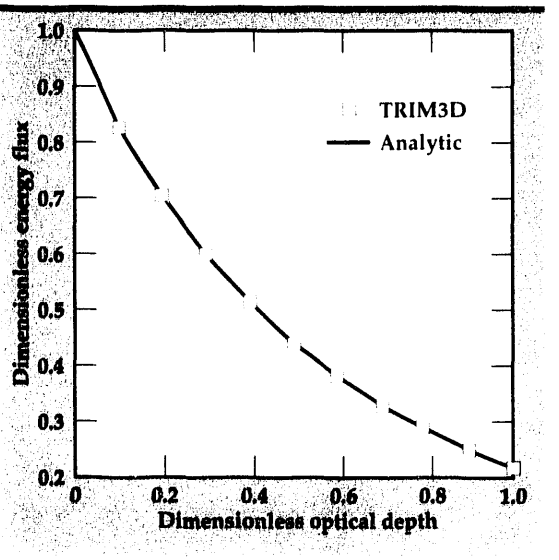


Figure 2. Analytic verification of TRIM3D for pure absorption.

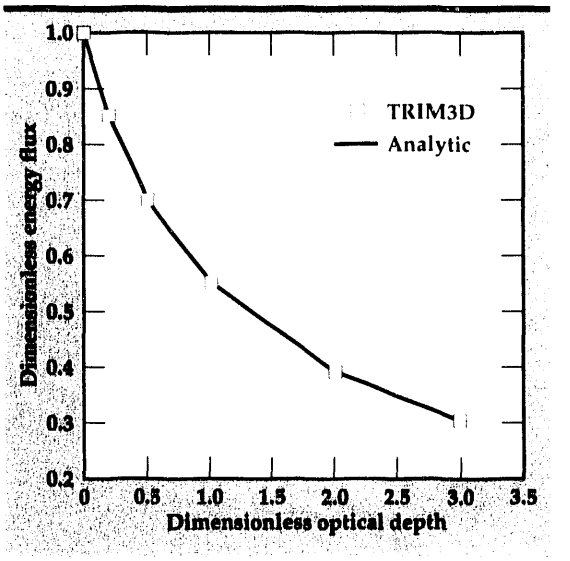


Figure 3. Analytic verification of TRIM3D for isotropic scattering.

If the material properties change significantly with wavelength, as is typical with gas-radiation problems, a band-wavelength model is available. This model splits the wavelength range into separate gray bands, with a separate simulation per band.

Surfaces consist of 4-node shell elements, degenerating to triangles. Volumes are represented as 8-node bricks, with triangular prisms and tetrahedra as subsets. Both surfaces and volumes are designed for mesh compatibility with INGRID and TOPAZ3D. Material properties are assumed constant within a single element, but any number of materials may be defined. In this manner, non-homogenous problems may be solved.

Analytic Verification

TRIM3D has been verified against a series of participating medium heat transfer problems with known analytic solutions. Though the analytic solutions are one-dimensional, they were simulated with a 3-D geometry with specular mirrors on four sides. Optical depths from 0.1 to 10 were tested, with good agreement throughout the entire range. Some of the results of this verification are shown in **Figs. 2** and **3** for pure absorption and isotropic scattering, respectively. Agreement with the analytic solution in both cases is very good.

An additional result from this verification was that the speed of the code appears to decrease only linearly with optical depth, and that even at an 'optically thick' depth of 10, the speed is practical on a SUN workstation.

Future Work

One of the difficulties in verifying a participating medium code is the small number of problems with analytic solutions that exercise all the code features. To address this problem, a symposium was held at the 1992 American Society of Mechanical Engineers Heat Transfer Conference to assess the current capability for solving non-gray, anisotropically scattering, multidimensional radiation problems. Thirty-four benchmark problems ranging from one to three dimensions at optical depths from 0.1 to 10 were specified.[3] These problems will be solved using TRIM3D and should provide a good platform for verification of the code features.

Additional features are planned for the production version of the code to simplify the solution of large problems and make the code more 'user friendly.' A complete user's manual, including test problems, will be produced for TRIM3D. In addition, all the solved analytic and benchmark problems will be organized into a quality assurance manual for code validation purposes.

Once the code is released, we intend to collaborate with groups inside or outside LLNL to test the utility and accuracy of TRIM3D on experimental problems. This will provide valuable feedback on code robustness and performance on large problems, as well as on which features are most useful to the analysis community.

Because TRIM3D uses a Monte Carlo formulation, it is very well suited to the new class of massively parallel computers. A test version of TRIM3D will be developed for whichever parallel computer becomes available at LLNL, and its performance will be assessed. If successful, it should provide a good example of a production-parallel application.

1. J.D. Maltby and P.J. Burns, *Numer. Heat Transfer* **9** (2), (1991).

2. R. Siegel and J.R. Howell, *Thermal Radiation Heat Transfer*, 4th ed., Hemisphere Publishing Corporation (Bristol, Pennsylvania), 1992.

3. T.W. Tong and R.D. Skyocypec, "Summary on Comparison of Radiative Heat Transfer Solutions for a Specified Problem," *Developments in Radiative Heat Transfer*, HTD **203** (New York), 1992.

A Methodology for Calculating the Seismic Response of Critical Structures

David B. McCallen
Nuclear Test Engineering
Mechanical Engineering

Francois E. Heuze,
Lawrence J. Hutchings, and
Stephen P. Jarpe
Earth Sciences Department

We are developing a methodology chain that will allow estimation of the seismic response of critical structures to large earthquakes. The methodology consists of three distinct steps: generation of synthetic bedrock motion at the structure site due to a postulated large earthquake; nonlinear finite element analysis of the soil profile at the site to transform the bedrock motion to surface motion; and linear/nonlinear finite element analysis of the structure based on the predicted surface motions. Progress in all steps is reported here. Our ultimate goal is to allow accurate, site-specific estimates of structural response for a specified earthquake on a specified fault.

Introduction

Our computational simulation of the seismic response of a critical structure is illustrated in **Fig. 1**. To envelope the motions that might be observed at the structure site, the seismological portion of the methodology develops a suite of possible earthquake rupture scenarios for each fault that can contribute significant ground motion at the site. Field instrumentation is placed on bedrock at the structure site, and over a period of time, bedrock motions due to micro-earthquakes emanating from the causative fault(s) are recorded. These recordings serve as empirical Green's functions, which characterize the motion at the structure site location due to slip of an elemental segment of the fault. By appropriate summation of the responses due to each element of the fault rupture zone for a given rupture scenario, the bedrock motion due to slip over a large area of the fault (corresponding to a large magnitude earthquake) can be estimated. By considering a standard suite of 25 possible fault rupture models, which characterize the different manners in which the fault rupture can propagate across the total fault rupture zone, a suite of 25 acceleration time histories are generated. The suite of time histories is representative of the maximum ground accelerations that could be expected at the site for a given size earthquake. Hutchings[1,2,3] and his coworkers have led the development of the empirical Green's function technique and demonstrated the utility of this method using Loma Prieta earthquake data.[2]

The transmission of earthquake motion from bedrock through the soil to the soil surface can result in significant modification of the bedrock motion. Traditionally, the nonlinear behavior of the soil under strong seismic motion has been modeled with 'equivalent linear' methods, which iterate with a linear model to approximate the nonlinear response of the soil deposits. The classical computer program SHAKE[4] has typically been used to perform site-response analysis. SHAKE is operational at Lawrence Livermore National Laboratory (LLNL), but such equivalent linear models cannot describe the evolution of pore pressure and predict liquifaction; i.e. they cannot perform "effective-stress" analysis which we deemed essential for this project. So, the effective stress nonlinear finite element program DYNAFLOW[5] has been obtained from Princeton University. As part of the methodology development and validation, the DYNAFLOW and SHAKE programs will be applied to the Loma Prieta earthquake data obtained at Treasure Island, California. The Treasure Island site consists of saturated soils that exhibited liquefaction during the Loma Prieta earthquake. Site-response calculations are being performed by

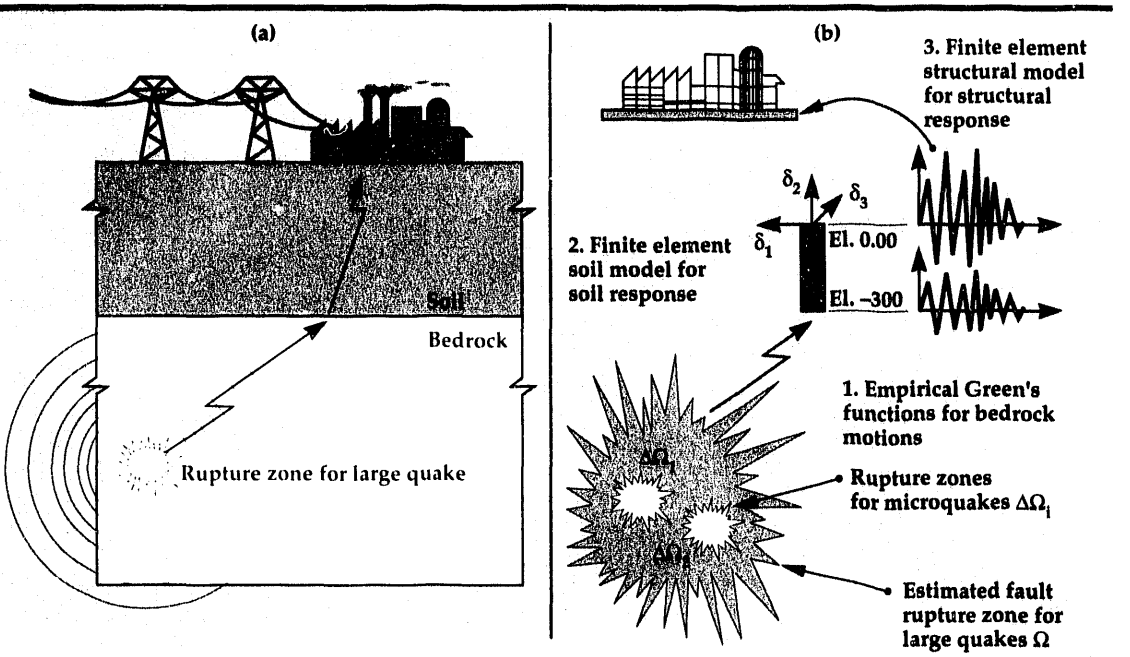


Figure 1. Computational simulation of the seismic response of a critical structure showing (a) the physical system and (b) the three-step computational model.

a number of researchers, and a portion of our model validation efforts will consist of a comparison of DYNAFLOW and SHAKE results with measured Treasure Island response data for the Loma Prieta earthquake.

Nonlinear structural-response computations are being performed with nonlinear finite element software developed at LLNL. The implicit, nonlinear, finite deformation program NIKE3D[6] is being used to model structures and the nonlinear near-field soil. NIKE3D has a number of nonlinear constitutive models and advanced contact-surface capabilities for modeling gap opening and closing.

The seismic analysis procedures and capabilities under development are being applied to two transportation structures in California. The first structure is the Dumbarton Bridge, which is the southern-most crossing of the San Francisco Bay. The Dumbarton Bridge study was initiated by LLNL at the request of the California Department of Transportation (CDOT). The second study is concerned with the seismic analysis of the Painter Street Bridge in Rio Dell, California. The Painter Street Bridge study is very important from the standpoint of validation of our methodology and procedures. This study is the focus of this report.

The Painter Street Bridge, which has been heavily instrumented by the California Department of Mines and Geology (CDMG), provides an excellent case study. The high rate of occurrence of

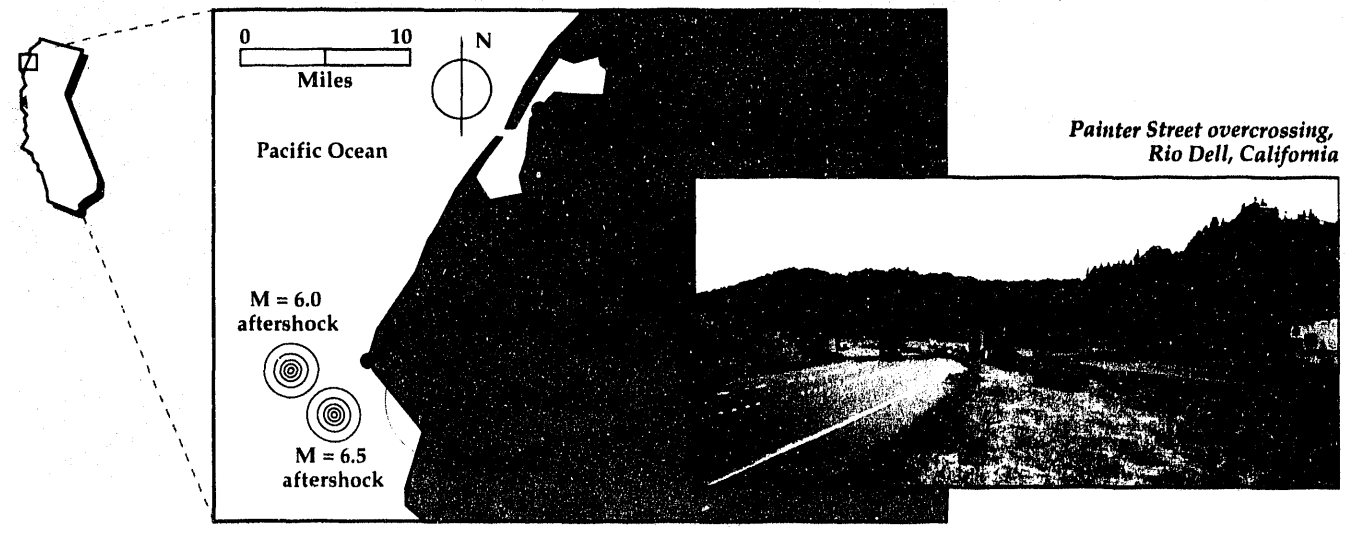


Figure 2. Location of April 1992 Petrolia earthquake epicenters and Painter Street Bridge site.

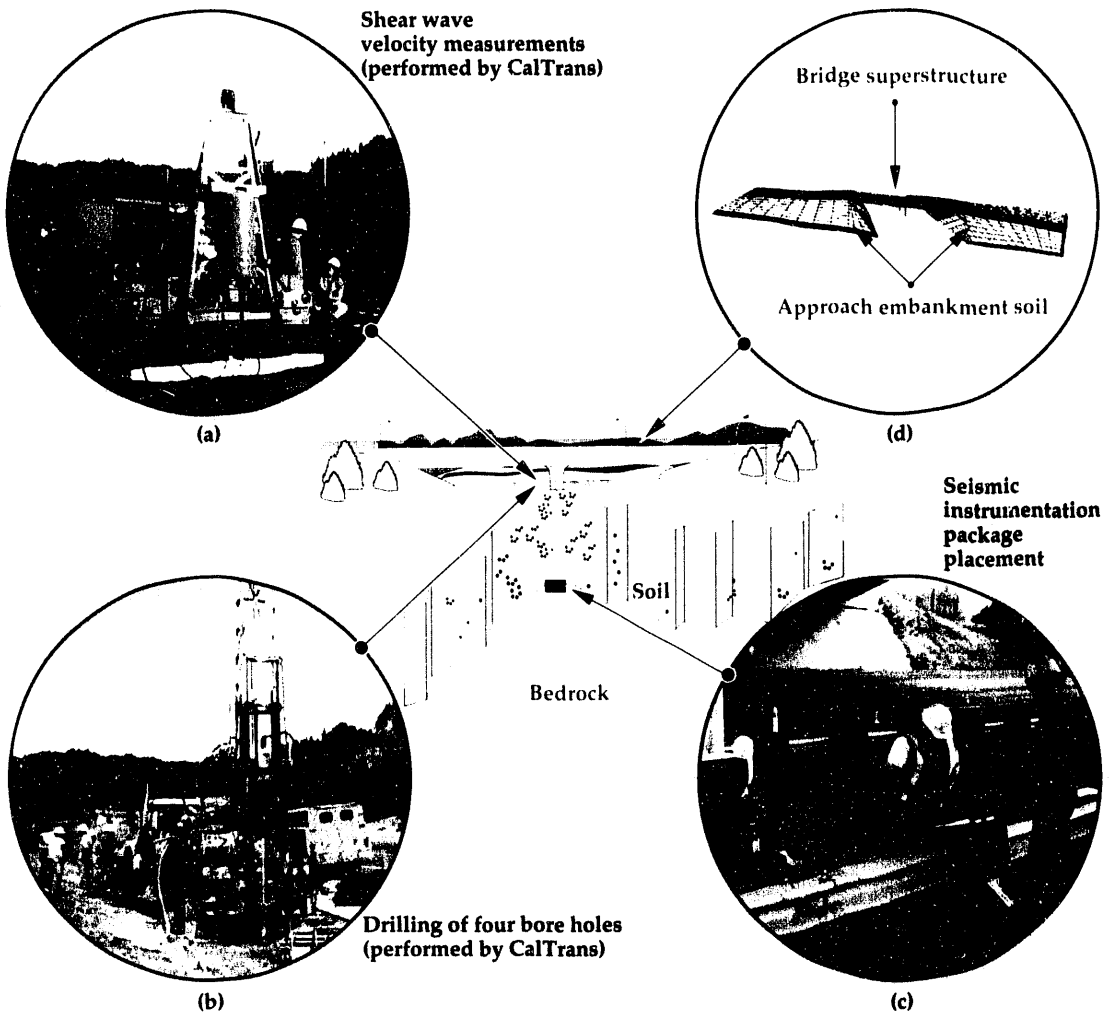


Figure 3. Photographs showing (a, b, and c) experimentation and field work for the Painter Street Bridge site; and (d) illustration of finite element model.

earthquakes in the seismically active region of northern California has allowed the measurement of response of this bridge to a number of significant earthquakes. In April 1992, three large earthquakes occurred within close proximity to Rio Dell and the Painter Street Bridge location (see **Fig. 2**). During the largest of these shocks, the Painter Street Bridge structure was shaken quite violently, with lateral deck accelerations on the order of 1.23 times the acceleration due to gravity. These measured accelerations represent the largest accelerations ever measured in any structure during an earthquake. Prior to the April earthquakes, McCallen, Romstad, and Goudreau[7] had constructed a detailed finite element model of the Painter Street bridge/abutment system (see **Fig. 3**) and had performed detailed parameter studies on the dynamic response of the system. Since an extensive modeling effort had already been initiated on this bridge, the latest set of quakes was a fortuitous event for our project.

Progress

As a result of the April quakes, the LLNL efforts at the Painter Street site have been scaled up signif-

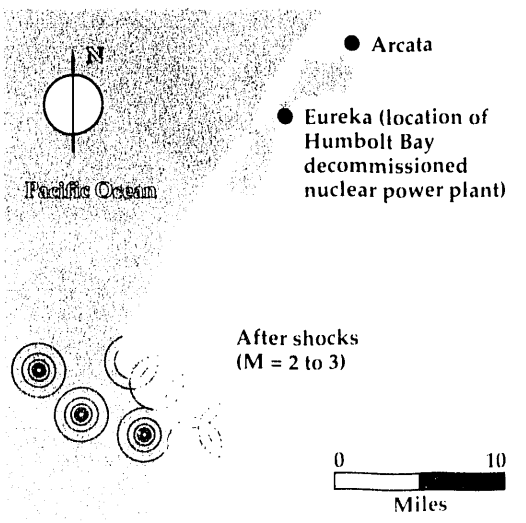


Figure 4. Region of aftershock locations used for measurement of empirical Green's functions.

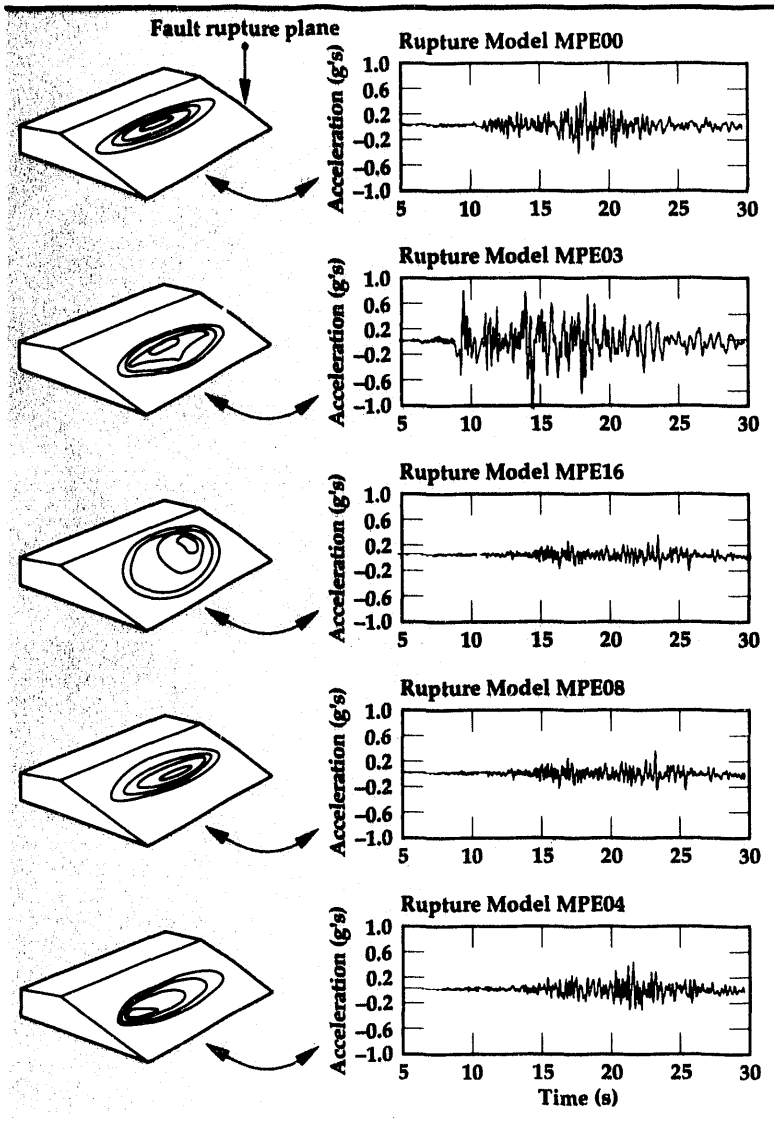


Figure 5. Five sample fault rupture scenarios with resulting Painter Street time histories.

icantly. At the request of Heuze,[8] the CDOT recently drilled four bore holes at the Painter Street site (see **Fig. 3**) and performed down-hole, shear-wave-velocity measurements. Soil samples were retrieved from the boreholes, and Heuze has contracted with the Department of Civil Engineering at (UCB) to perform laboratory tests on the soil samples. The field shear-wave-velocity measurements and the laboratory soil testing will provide quantitative soil properties for use in the site soil response calculations and the structural model calculations. Two of the boreholes were drilled to bedrock (a depth of approximately 80 ft), and two seismic instrumentation packages were placed, one at the surface and the second at the bedrock depth of 80 ft (**Fig. 3**). The package at bedrock depth is currently being used by seismologists to measure empirical Green's functions for micro-earthquakes emanating from nearby faults.

To date, Painter Street site bedrock responses have been measured for eight micro-earthquakes emanating from the fault locations indicated in **Fig. 4**. Based on the empirical Green's functions obtained from these measurements, synthetic bedrock-ground-motion time histories have recently been generated by Hutchings and Jarpe for a number of earthquakes. Samples of Painter Street site synthetic time histories, each based on a different fault rupture propagation model, are shown in **Fig. 5**.

In parallel to the seismological work, finite element modeling of the Painter Street bridge/abutment system has progressed into the nonlinear regime. For nonlinear time history analyses, the superstructure, pile foundation, and approach embankment soil masses have been modeled as shown in **Fig. 3**. In this type of bridge structure, nonlinear hysteretic behavior of the soil embankments has been experimentally identified as a very important factor in the dynamic response of the bridge system.[9,10] The primary objective of constructing a detailed, three-dimensional model of the bridge/soil system was to allow incorporation of the effects of nonlinear soil stiffness and soil mass. Traditional finite element models for this type of bridge, which are used in bridge design and analysis calculations, neglect the soil mass, and the soil stiffness is represented by linear elastic, amplitude-independent springs. We decided to truncate the detailed finite element model at approximately the original ground surface elevation, and apply the surface free field motion directly to the base of the model at this elevation (see **Fig. 3**). This approach neglects potential soil–structure interaction effects between the piles and soil below this level, and prevents radiation of energy vertically back into the soil. However, interaction between the soil and piles typically occurs in the top portion of the piles, and energy loss through radiation will be small relative to the energy dissipated by the nonlinear hysteretic behavior of the soil embankments.

Until the experimental tests are completed at UCB in January 1993, there is only sparse quantitative data on the soil properties for the Painter Street site. The small-amplitude shear moduli for the approach embankments and original grade soils have been estimated[11] based on P and S wave surface refraction measurements which were performed. To represent the nonlinearity of the soil in the bridge/abutment finite element model, the small-strain shear moduli obtained from these measurements were used with standard soil modulus degradation and damping curves.[12] To represent the standardized modulus degradation and damping curves in the NIKE3D finite element program,

a simple Ramberg–Osgood constitutive model was used to model the soil. The material parameters were set such that the Ramberg–Osgood hysteresis loop would yield modulus degradation and damping curves very similar to Seed's[12] standardized curves. The procedure for determining the Ramberg–Osgood parameters to approximate given modulus degradation and damping curves was developed by Ueng and Chen.[13] The modulus and damping curves obtained from the Ramberg–Osgood constitutive model fit with Ueng and Chen's technique are shown in **Fig. 6** along with the original curves of Seed. The shear stress-strain behavior generated with the fitted Ramberg–Osgood model in the NIKE3D finite element program is also shown in **Fig. 6**.

A number of time history analyses have been carried out with the detailed bridge/abutment model shown in **Fig. 3**, as well as with simple reduced-order stick models of the bridge.[4] The bridge instrumentation records for the April 1992 Petrolia earthquakes have not yet been completely processed by the CDMG; thus, the measured free field motions were not available to apply to our model prior to this report. However, free field and bridge-response data for a magnitude 5.5 earthquake of November 1986 were available and were used to examine the accuracy of the finite element models of the bridge system.

The 1986 free-field acceleration time histories were used as input motion to the base of the bridge system models. The model response predictions were compared to the actual bridge response data measured by the CDMG bridge instrumentation array. Since the details of all of the response predictions are given elsewhere,[14] only an illustrative example of the response predictions is provided here. The detailed model response predictions for the absolute displacement at channel 7 (transverse motion at mid-span) are shown in **Fig. 7**. **Figure 7a** shows the response of the detailed model when a linear elastic soil model is used, and mass- and stiffness-proportional Rayleigh damping is used to provide approximately 5% damping in the first transverse and longitudinal modes of the bridge system. For the linear analysis, soil properties were set equal to the small-strain soil properties estimated by Heuze and Swift from field measurements. Two observations can be made: (1) the frequency content of the bridge model is significantly too high when the small-strain soil properties are used; and (2) the amplitude of the response prediction is too large relative to the measured response. The bridge response prediction using the detailed model with the nonlinear Ramberg–Osgood soil model

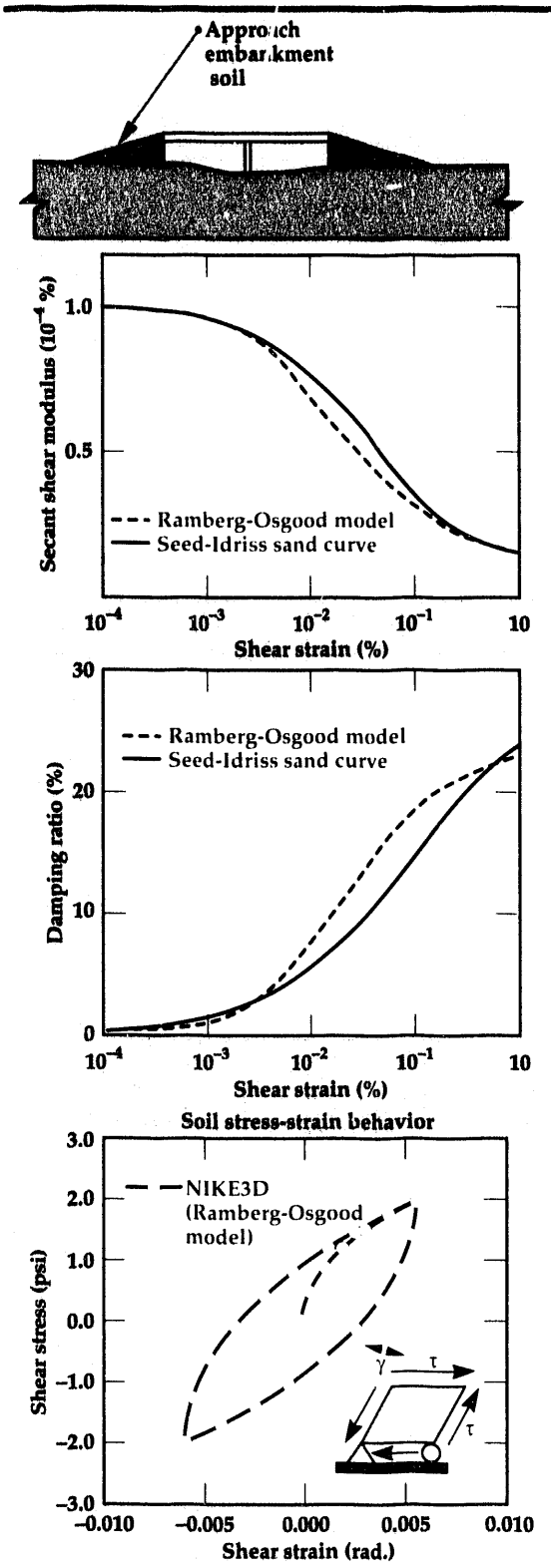


Figure 6. Simple nonlinear soil characterization for finite element model.

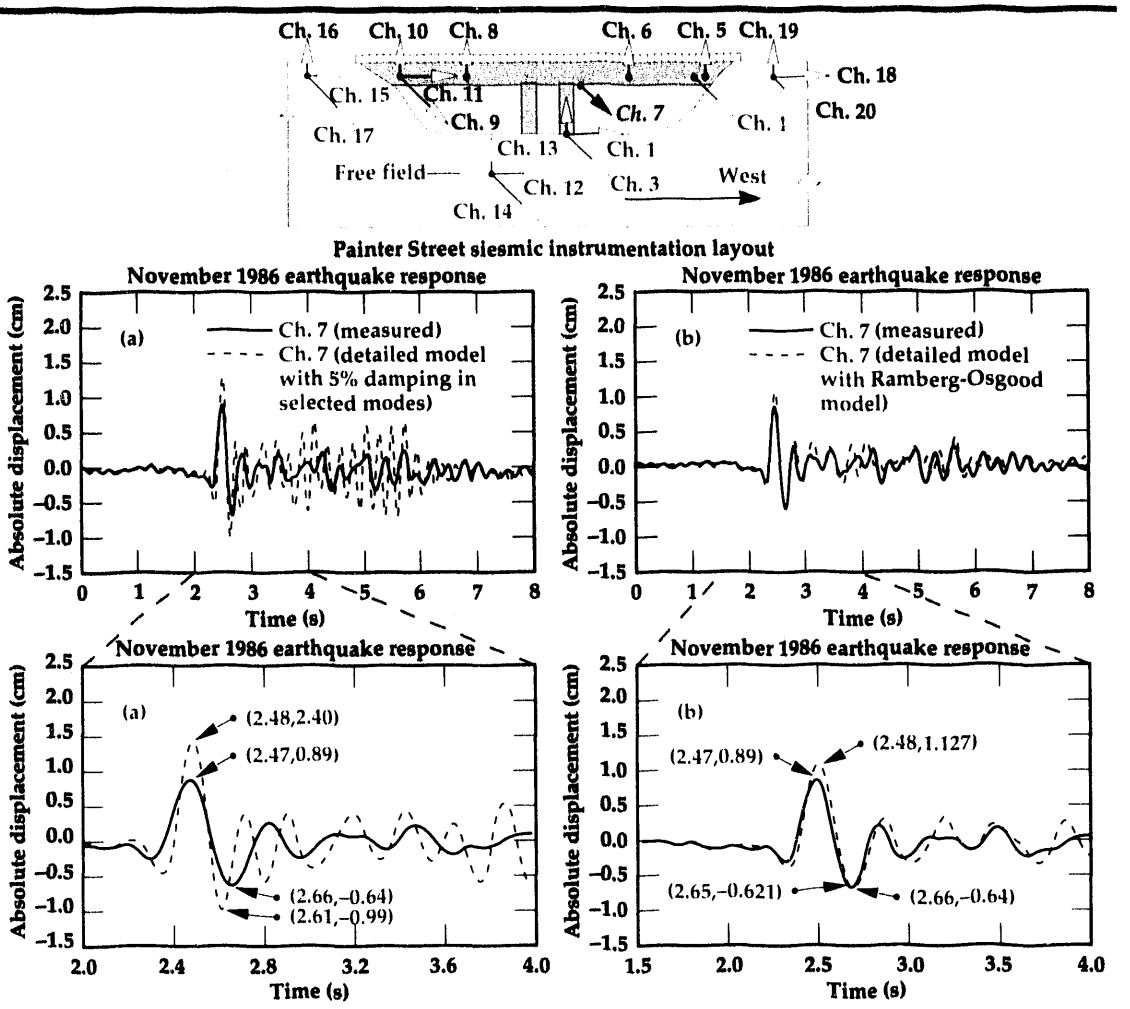


Figure 7. Response predictions with (a) linear and (b) nonlinear models.

is shown in **Fig. 7b**. This model also used mass-proportional Rayleigh damping, in which the damping in the first transverse mode was set to 10%. The nonlinear model exhibits significant improvement over the linear model. The nonlinear model displays appropriate softening and energy dissipation in the system, such that the frequency content and amplitude are more representative of the actual structural response.

Future Work

Significant progress has been made in the study of the Painter Street overcrossing site. Construction of the seismological model and the structural model have been completed, and calculations have been generated with both models. Additional field measurements of Green's functions from future microearthquakes will continue to enhance the site seismological model, and laboratory experimental data will improve the soil characterization in the finite element model of the bridge/abutments. The site-soil characterization will also allow site-response analysis to transform bedrock motion to soil surface motion. Specific tasks that we intend to perform during the next year include:

(1) Use of the nonlinear model of the bridge/abutment system to predict the response of the bridge to the April 1992 Petrolia earthquake. The predicted response will be compared to the actual bridge response measured by CDMG. This earthquake should have resulted in significant nonlinear behavior of the bridge/abutment system, and this analysis will allow us to further verify the ability of the nonlinear model to accurately predict bridge/abutment response.

(2) Based on measured Green's functions, the seismological model will generate a final suite of 25 time histories for the April 1992 Petrolia magnitude-7 earthquake.

(3) The bedrock-motion time histories will be transformed to surface motion with a site-response analysis, and the suite of surface time histories will be compared to the actual free field motion measured at the site by CDMG.

(4) The suite of predicted free field responses will be run through the structural model, and response statistics will be compared to the actual response from the April 1992 magnitude-7 earthquake.

The ultimate goal of our project is to allow accurate site-specific estimates of structural response for a specified earthquake on a specified fault. For practical applications of this methodology, it will be essential to decide how the structural engineer may best use the information provided by the suite of time histories developed by the seismological portion of the study. It will generally be impractical to perform 25 time history analyses (or more if multiple faults/multiple rupture zones are considered) for a large structural model. It is necessary to consolidate the information obtained from the time histories into a simplified form (e.g., a representative response spectrum and corresponding single time history) to achieve practical application.

The Painter Street site study will allow a critical evaluation of the accuracy of the method that is being developed, and a demonstration of our technology in all segments of the methodology chain. It will also provide an opportunity for interaction between structural analysts and seismologists, so that appropriate procedures for using the earthquake ground motion in structural response calculations can be developed.

1. L. Hutchings, *Modeling Earthquake Ground Motion with an Earthquake Simulation Program (EMPSYN) That Utilizes Empirical Green's Functions*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-ID-105890 (1992).
2. L. Hutchings, *Bull. Seism. Soc. Am.* **81** (5), (1991).
3. L. Hutchings, *J. Geophy. Res.* **95** (B2), (1990).
4. P.B. Schnabel, H.B. Seed, and J. Lysmer, *SHAKE—A Computer Program for Earthquake Response Analysis of Horizontally Layered Sites*, Earthquake Engineering Research Center, University of California, Berkeley, California, Report EERC 72-12 (1972).
5. J.H. Prevost, *DYNAFLOW, User's Manual*, Department of Civil Engineering, Princeton University (1992).
6. B.N. Maker, R.M. Ferencz, and J.O. Hallquist, *NIKE3D: A Nonlinear, Implicit, Three-Dimensional Finite Element Code for Solid and Structural Mechanics, User Manual*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-MA-105268 (1991).
7. D.B. McCallen, K.M. Romstad, and G.L. Goudreau, "Dynamic Response of a Reinforced Concrete Box-Girder Bridge," *Engineering Research, Development, and Technology*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-91, 2-12 (1992).
8. Personal communication between F. Heuze (LLNL) and Kenneth Cole (CDOT) (1992).
9. J.C. Wilson, and B.S. Tan, *ASCE J. Eng. Mech. Div.* **226** (8), (1990).
10. S.D. Werner, J.L. Beck, and M.B. Levine, *Earthquake Eng. and Struct. Dyn.* **15**, (1987).
11. F.E. Heuze and R.P. Swift, *Seismic Refraction Studies at the Painter Street Bridge Site, Rio Dell, California*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-ID-108595 (1992).
12. H.B. Seed, R.T. Wong, I.M. Idriss, and K. Tokimatsu, *Moduli and Damping Factors for Dynamic Analysis of Cohesionless Soils*, Earthquake Engineering Research Center, University of California, Berkeley, California, Report EERC-84/14 (1984).
13. T.S. Ueng, and J.C. Chen, *Computational Procedure for Determining Parameters in Ramberg–Osgood Elastoplastic Model Based on Modulus and Damping Versus Strain*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-ID-111487 (1992).
14. D.B. McCallen and K.M. Romstad, *Dynamic Response of a Reinforced Concrete, Box-Girder Bridge*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-ID-110640 (1992).

Reinforced Concrete Damage Modeling

Sanjay Govindjee and Gregory J. Kay
Nuclear Explosives Engineering
Mechanical Engineering

The modeling of reinforced concrete structures is currently performed by empirical codified formulae and linear elastic calculations. This state of the practice, however, can lead to both non-conservative designs on the one hand and to over-designed and costly structures on the other. This wide range of outcomes arises from the lack of an adequate constitutive model to describe the behavior of concrete as it cracks under applied loads. This report briefly describes work at Lawrence Livermore National Laboratory in the development of an appropriate constitutive model for concrete damage.

Introduction

In the modeling of reinforced concrete structures, the current state of the practice involves the use of codified empirical formulae and linear calculations. While these methods are very useful, they can also produce unwanted results. When using empirical formulae, there is risk involved in applying them to a situation that is not absolutely identical to the tests from which they were deduced. In particular, formulae for limit loads scale in a rather non-linear fashion and must be applied with care and experience to avoid a non-conservative design. On the other hand, one does not want to have to over-build a structure and hence make it overly costly because of uncertainties in modeling.

A vast improvement to the design cycle is obtained if some of the empirical formulae currently used are replaced by analytical models. The main unknown that most of the empirical formulae try to address involves the behavior of the concrete itself as it cracks under various loading conditions with different reinforcement patterns. Thus, the thrust of our work has been to develop a constitutive model that describes the behavior of damaging concrete. Because this work is being performed for the Computational Earthquake Initiative at Lawrence Livermore National Laboratory (LLNL), the level of complexity of the model has been chosen to be commensurate with that needed to model critical sections of large reinforced concrete structures under seismic loading conditions. This requires the constitutive model to be able to track the progression of damage induced by arbitrary three-dimensional (3-D) loading histories in complex 3-D geometries. Because of these requirements, the model has been developed as a 3-D damage theory that is suitable for large-scale finite element calculations.

Such thinking is not new to the modeling of reinforced concrete structures.[1] This original work, and almost all that has followed since, has been confined to two-dimensional (2-D) problems. Under seismic excitations, however, one must look at the more general situation that includes 3-D effects, because of the high likelihood of complex loading paths. There does exist a handful of 3-D models.[2,3,4] However, none of these models is suitable for the present problem. The first two models and others like them are only suitable for isotropic compressive type behavior, and the third, while promising, still requires some development. The present model takes advantage of the insights and developments of this previous work and extends them to a new framework for damage modeling. The framework we have developed most closely resembles the framework proposed by Ortiz.[4]

Progress

Progress for FY-92 has been made on many different aspects of the problem: choosing an appropriate class within which to develop the model; developing the features to incorporate into the model; developing appropriate numerical algorithms to efficiently perform finite element calcu-

lations; and determining how reinforcing bars should be modeled in conjunction with the cracking concrete.

Model Class and Features

Model class refers to the basic style of the model: plasticity-like or damage-like. In plasticity-like models, material unloading takes place elastically with a stiffness equal to the initial elastic stiffness of the material (**Fig. 1a**). In a damage-like model, material unloading takes place elastically with a degraded stiffness (**Fig. 1b**). The plasticity-like models have strong appeal for a number of reasons, but mainly because their algorithmic properties are reasonably well understood and are known to be suitable for finite element calculations. The true behavior of cracking concrete, however, resembles more closely damage-like model behavior.

Nevertheless, at the beginning of this project, we used a plasticity-like model to examine some of the numerical and theoretical issues that are unique to materials displaying softening behavior like that shown in **Fig. 1**. The main use of this model class was to examine the issue of ill-posed boundary-value problems. Materials displaying stress-strain behavior like that in **Fig. 1** often generate ill-posed boundary-value problems.[5] While there are several ways around this issue, for concrete the most physically realistic one is the notion of constraining the amount of energy dissipated in the system on a per-unit-volume basis to equal that dissipated on a per-unit-area basis when opening new crack faces. This type of constraint results in the appearance of a characteristic length in the model formulation. For the development of the present model, the continuum formulation[6] was used to render the present formulation well-posed for both the plasticity and damage model classes.

In the domain of damage models, there is a wide variety of model choices. To choose the appropriate one usually requires a fair amount of insight into the micromechanical mechanisms of the observed damage and their relationship to the free energy density of the material. In the case of concrete with Mode I-, II-, and III-type cracks, such information is not available. Therefore, several general hypotheses of continuum mechanics have been used instead to generate a complete model.

The basic premise of the model is that the damage state of the material will be represented by the rank 4 stiffness tensor of the material. Hence, as is known to occur in other damaging systems,[7] the 'elastic stiffness' of the material is allowed to evolve with the loading history. To determine the evolution law for this degraded stiffness, the notion of maximum dissipation is used. To use this idea, one first postulates restrictions on the admissible stress or strain states of the material. For the concrete, two restrictions are postulated. The first restriction states that the normal tractions across cracks in the system must be below a given critical value and that the critical value evolves as the damage increases. The second restriction states that the shear tractions across cracks in the system must be below a given critical value, which also evolves with progressing damage. Cracks are assumed to nucleate in the material when the maximum principal stress at a point exceeds a given value. Using these two restrictions and the concept of maximum dissipation, an evolution law can be derived for the rank 4 stiffness tensor of the material that gives the correct anisotropic structure to the damaged stiffness tensor.

The other dominant phenomenological features of cracking concrete that have been incorporated into the model are:

(1) The choice of restrictions on the admissible stress states in the material provides

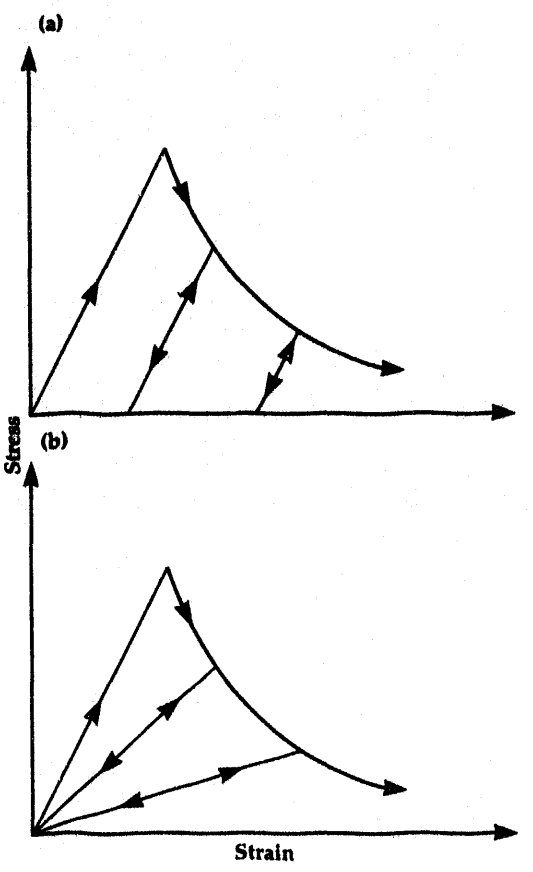


Figure 1. Material unloading in (a) plasticity-like and (b) damage-like model classes.

for Mode I-, II-, and III-type crack growth (damage evolution).

(2) The notion of crack closure has been included by monitoring the tractions across crack faces. When the traction across a crack face becomes compressive (negative) and the shear tractions are below their critical value, the material behaves as though it is undamaged (up to the compressive yield limit of the concrete).

(3) The notion of shear retention is built into the model by limiting the amount of shear degradation allowed in the system.

(4) The softening evolves with an exponential character.

(5) The damage evolution is anisotropic.

Algorithms

The algorithmic implementation of the proposed model in a finite element setting has involved the development of several novel algorithms. Of foremost importance for softening models has been the development of a characteristic-length interpolation scheme for 3-D problems. While an interpolation scheme for 2-D problems has been presented,[6] a straightforward extension of this method to 3-D leads to singular characteristic lengths for certain crack orientations. In our work, a new interpolation method has been developed that does not have these singularities and yet remains faithful to the original characteristic-length idea.

The other algorithmic issues that have been addressed deal with local and global integration algorithms. On the local level, a concave (as opposed to convex, as in metal plasticity) optimization problem governs the stress point calculation. Because of the concave nature of the problem, a unique answer to the stress point calculation does not exist; there are two answers, with one being inadmissible. However, by picking a suitable starting value, the stress point algorithm can be made to always produce the admissible answer. On the global level, the nonlinear balance equations of the boundary-value problem have multiple bifurcation paths that lie extremely close to each other and cause global convergence difficulties. To circumvent these well-known convergence difficulties, an aggressive, automatic time-stepping scheme has been developed. The scheme uses logarithmic-based time step control in conjunction with a special oscillating norm check. The combination of these two ideas greatly enhances the ability of the global solvers to achieve equilibrium.

Reinforcing Bars

Since using fixed rebar bars (i.e., compatible displacements between concrete and rebar) gives reasonable results, to date only a small effort has been devoted to rebar issues. Our results are, however, slightly non-conservative. To address this, some preliminary work has been done on rebar release methods. Force- and damage-based slideline release methods have been used, as have bond-link elements. The damage-based slideline release has been found to be superior to the force-based model and the bond-link element for accuracy against experimental data. However, the best overall robustness for these methods (after the fixed rebar model) is given by the bond-link element, which is a node-on-node contact element with a displacement-based release law.

Examples

Two examples are shown to partially demonstrate the proposed model. The first example involves the 3-point bending of a lightly reinforced beam; the second example involves the 3-point bending of a heavily reinforced beam.

In the first example, the beam is 12 feet long with a 8 × 20 in. cross section that contains two #8 rebars in the lower fibers. The load deflection curve at mid-span is shown in **Fig. 2**. Overall agreement is seen to be quite good. At point (A), the concrete starts to crack, and load is transferred into the rebars. Cracking progresses up through the cross section with more load being transferred into the rebars until at point (B) the rebars yield. These observations from the simulation are consistent with experimental observations.[8]

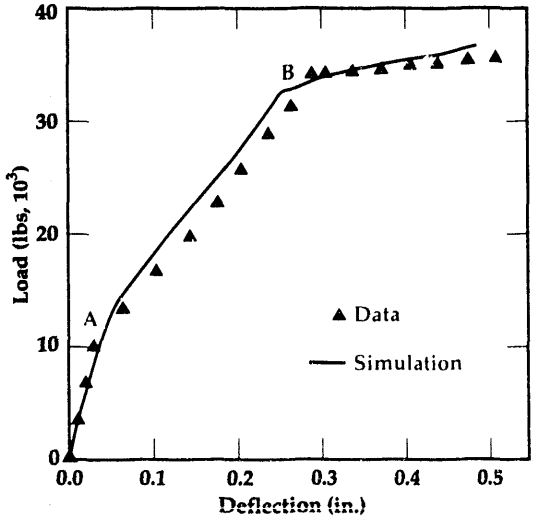


Figure 2. Load deflection curves at mid-span for beam with two #8 rebars in the lower fibers. The damage initiation point (A) and the point of yield (B) are marked.

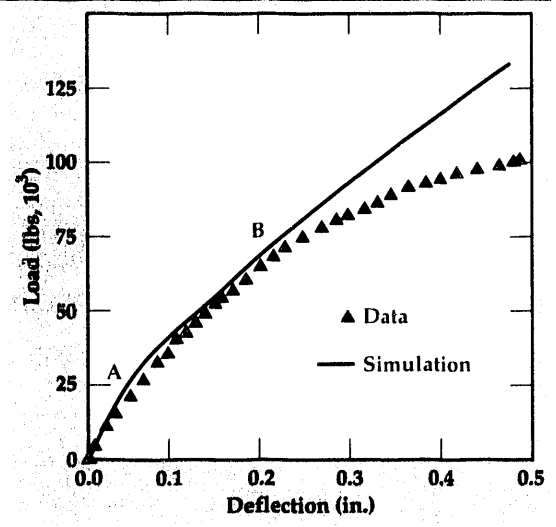


Figure 3. Load deflection curves at mid-span for beam with four #9 rebars in the lower fibers and two #4 rebars in the upper fibers. The damage initiation point (A) and the rebar-concrete interface failure initiation point (B) are marked.

In the second example, the beam is 12 feet long with a 12 × 21.75 in. cross section that contains four #9 rebars in the lower fibers of the beam and two #4 bars in the upper fibers of the beam. Additionally, there are #2 stirrups every 8.25 in. along the length of the beam. **Figure 3** shows the load deflection curve at mid-span for the experiment[9] and the calculation. At point (A), the concrete starts to crack, and there is a large load transfer to the #9 rebars. The #4 rebars do not carry much of the load. Vertical cracks develop along the span and grow upwards and towards the centerline of the beam. At point (B), the calculation deviates from the data because rebar release was not included in the simulation.

Future Work

Future work will focus on making the local stress point algorithm more robust and efficient. In addition, a few new features will be added, such as compressive flow of the concrete and crossing cracks.

Acknowledgements

The authors wish to acknowledge Dr. B. Maker of LLNL, Prof. R.L. Taylor of the University of California, and Prof. J.C. Simo of Stanford University for their help and interest in carrying out this work.

1. Y.R. Rashid, *Nuc. Eng. Des.* **7**, 334 (1968).
2. D.C. Drucker and W. Prager, *Q. Appl. Math.* **10**, 157 (1952).
3. F.L. DiMaggio and I.S. Sandler, *J. Eng. Mech.* **97**, 935 (1971).
4. M. Ortiz, *Mech. Mat.* **4**, 67 (1985).
5. L.J. Sluys, *Wave Propagation, Localization and Dispersion in Softening Solids*, Ph.D. Dissertation, Delft University of Technology (1992).
6. J. Oliver, *Int. J. Numer. Meth. Eng.* **28**, 461 (1988).
7. S. Govindjee and J.C. Simo, *J. Mech. Phys. Solids* **39**, 87 (1991).
8. N.H. Burns and C.P. Seiss, University of Illinois Civil Eng. Studies SRS No. 234, (1962).
9. B. Bresler and A.C. Scordelis, *J. Am. Concr. Inst.* **60**, 51 (1963).

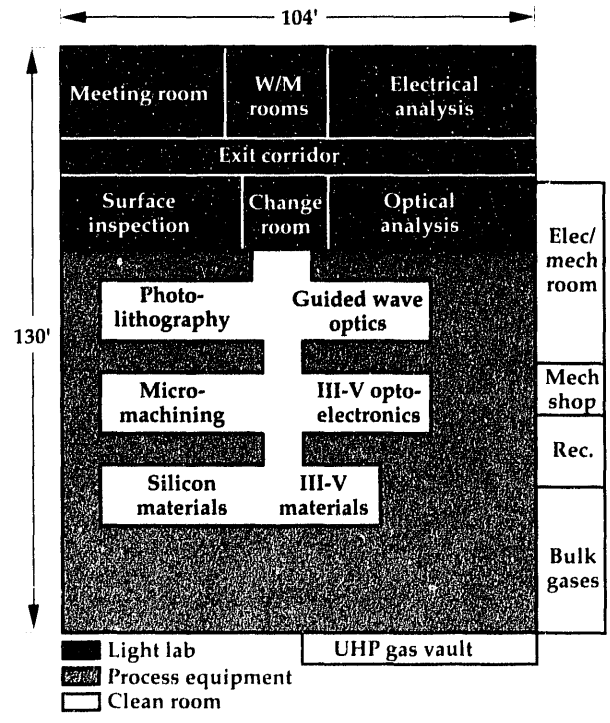
Diagnostics and Microelectronics

The Diagnostics and Microelectronics thrust area conducts activities in semiconductor devices and semiconductor fabrication technology for programs at Lawrence Livermore National Laboratory. Our multidisciplinary engineering and scientific staff use modern computational tools and semiconductor microfabrication equipment to develop high-performance devices. Our work concentrates on three broad technologies of semiconductor microdevices: (1) silicon and III-V semiconductor microelectronics, (2) lithium niobate-based and III-V semiconductor based photonics, and (3) silicon-based micromachining for application to microstructures and microinstruments.

In FY-92, we worked on projects in seven areas, described in the reports that follow: (1) novel photonic detectors; (2) a wideband phase modulator; (3) an optoelectronic terahertz beam system; (4) the fabrication of microelectrode electrochemical sensors; (5) diamond heatsinks; (6) advanced micromachining technologies; and (7) electrophoresis using silicon microchannels.

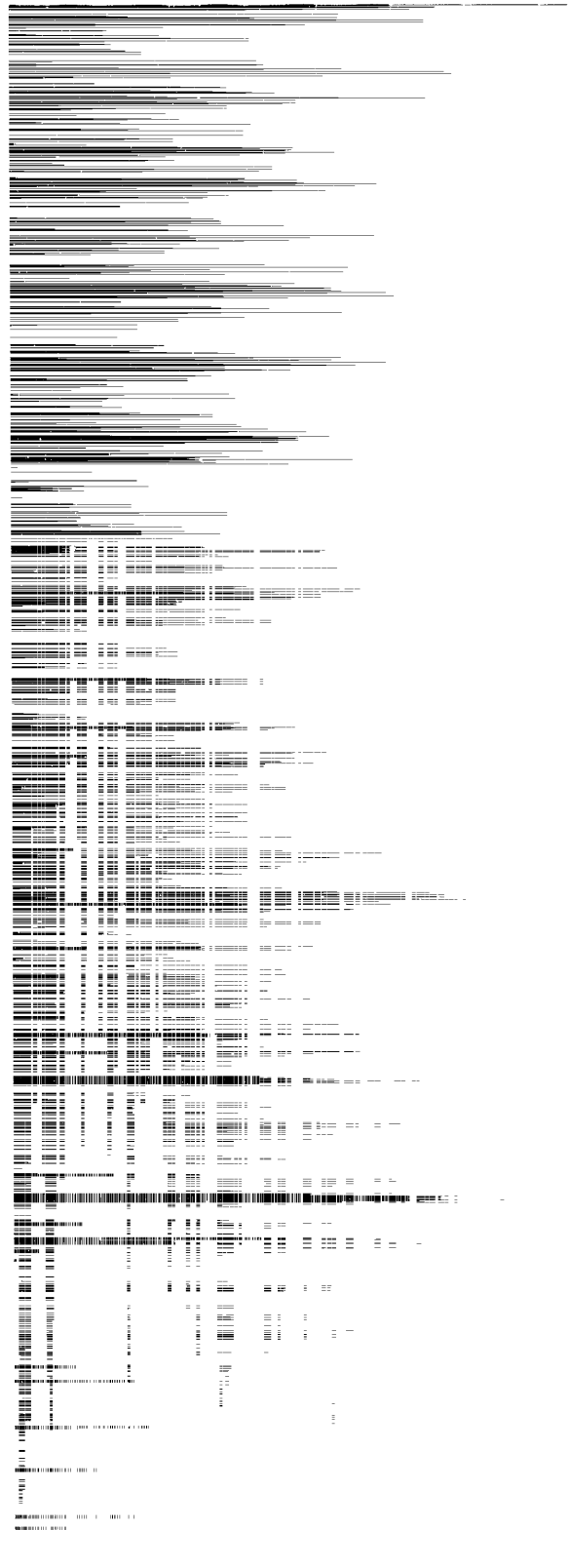
In FY-92 construction of the new Micro-Technology Center was completed. This new state-of-the-art facility includes 7,500 sq. ft. of Class 10-1000 cleanrooms and three large dry laboratories. The building was specifically constructed for the Laboratory to exceed all federal and state safety codes and regulations. All toxic gases are stored in autopurge gas cabinets located in separate earthquake resistant vault. All air handling machinery is mounted on a separate foundation vibrationally isolated from the cleanroom laboratories. The dry laboratories are used for microscopic inspection, packaging, and electrical and optical testing of devices. While the Micro-Technology Center is primarily a solid state device research facility, the emphasis of the thrust area is to solve problems for internal and external customers relating to diagnostic and monitoring instrumentation in a variety of scientific investigations.

Joseph W. Balch
Thrust Area Leader



B153: Micro-Technology Center

Section 3



3. Diagnostics and Microelectronics

Overview
Joseph W. Balch, Thrust Area Leader

Novel Photonic Detectors
Raymond P. Mariella, Jr., Gregory A. Cooper, Sol P. Dijaili, Robert Chow, and Z. Liliental-Weber 3-1

Wideband Phase Modulator
Charles F. McConaghy, Sol P. Dijaili, and Jeffrey D. Morse 3-5

Optoelectronic Terahertz Beam System: Enabling Technologies
Jeffrey D. Morse 3-9

Fabrication of Microelectrode Electrochemical Sensors
Dino R. Ciarlo, Jackson C. Koo, Conrad M. Yu, and Robert S. Glass 3-13

Diamond Heatsinks
Dino R. Ciarlo, Jick H. Yee, Gizzing H. Khanaka, and Erik Randich 3-15

Advanced Micromachining Technologies
Wing C. Hui 3-19

Electrophoresis Using Silicon Microchannels
Jackson C. Koo, J. Courtney Davidson, and Joseph W. Balch 3-21

Novel Photonic Detectors

Raymond P. Mariella, Jr., Gregory A. Cooper, and Sol P. Dijaili
Engineering Research Division
Electronics Engineering

Z. Liliental-Weber
Lawrence Berkeley Laboratory
Berkeley, California

Robert Chow
Materials Fabrication Division
Mechanical Engineering

This project had two parts for FY-92: (1) to fabricate a photocathode that could respond to infrared (IR) light; and (2) to fabricate a diode laser that would function as an x ray-to-light converter.

Although IR-sensitive photocathodes are not available commercially, there are numerous Lawrence Livermore National Laboratory and Department of Defense applications for such devices, including a 1.3-μm streak camera and radiation-hard IR sensors. The key part of our work on an IR-sensitive photocathode is the use of molecular beam epitaxy (MBE) to grow high quality semiconductor layers that can absorb IR light and transport the resulting charge carriers to the opposite surfaces of this electrical device. During this last year, as part of a separate research project, we discovered a new kind of photocathode and, thus, we centered our activities for both projects on fabricating and testing devices that incorporated it.

Little data had been published on the direct effects of x rays on diode lasers, and our idea was to use the absorption of x rays within the gain medium itself to modulate the optical output from a diode laser. The advantage of this device was expected to be that it should have picosecond response times since, at least in a simple double-heterostructure laser, there would be no time delay due to carrier transport. To test the device experimentally, we used a pulsed x ray source, which was a plasma that was created by a pulsed laser focused onto a metal surface. Although we did observe the direct conversion of x rays to optical output on a fiber optic, we were unable to make an accurate determination of the ultimate time response of the device.

Introduction

Photocathode

Of all materials tested, p-type GaAs, coated with cesium and a form of cesium oxide (hereafter we shall simply refer to this as 'activated'), has shown the highest quantum yield, η, for detection of visible and near-infrared (IR) light; state-of-the-art commercial GaAs photocathodes can have $\eta = 10\%$ for wavelengths out to 0.9 μm. For light with longer wavelengths, however, the GaAs is essentially a transparent material, and its use for direct detection of IR light is not practical. While all-solid-state detectors with high values of η exist for light wavelengths longer than 0.9 μm, photocathodes offer greater radiation hardness and are well-suited to photon counting and two-dimensional imaging when used in conjunction with electron multipliers, such as microchannel plates. A radiation-hard detector with sensitivity to 1.06-μm light is desirable for LIDAR applications; atmospheric viewing can be achieved in the 1.3- to 1.8-μm band; and a photocathode with sensitivity to 1.3-μm light would find application in a streak camera that could be used for remote monitoring of physics experiments.

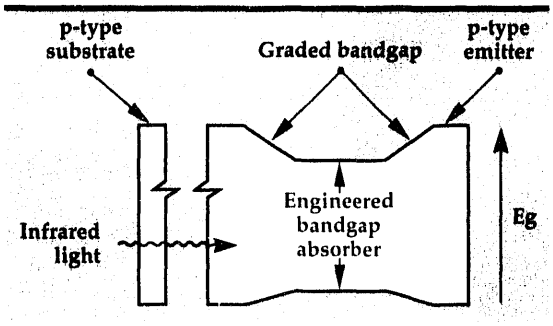


Figure 1. Schematic drawing of the band structure of an IR-photocathode. The smaller bandgap region absorbs the IR light, to which the substrate is transparent. Light absorption promotes an electron from the lower energy level (the valence band) to the upper level, the conduction band. Since the conduction band of the layers on either side of the IR absorber is higher than that of the IR absorber, this unbiased device cannot transport the electron out of the IR absorber.

Photocathodes of $Ga_{1-x}In_xAs$ or $GaAs_{1-x}Sb_x$ (for small values of x) have smaller bandgaps than GaAs and can absorb longer-wavelength light and have been fabricated elsewhere, but their values of η fall very quickly as a function of x. These photocathodes, in which a single material functions as both the IR absorber and the electron emitter, have not shown useful single-shot sensitivity to light with wavelengths beyond 1.1 μm.

A more advanced concept, originally demonstrated by Varian EOSD, was to use a two-part semiconductor photocathode, with one part as the IR absorber (GaInAsP) and the other part as the electron emitter (InP). This two-part approach has been what we have pursued. Our idealized IR photocathode would, therefore, contain a material that could absorb the infrared light, create electron-hole pairs, and (under applied electrical bias) separate and move the electrons without loss to an activated surface for emission. This is shown schematically in **Fig. 1**.

We started investigating strained-layer $In_yGa_{1-y}As$ on GaAs substrates with $In_x(Al_wGa_{1-w})_{1-x}As$ as an emitter, but because this combination was intrinsically limited in its long-wavelength response to about 1.3 μm and because we had recently invented a new photocathode (GaAlSb), we concentrated on this latter system, which we are in the process of patenting.

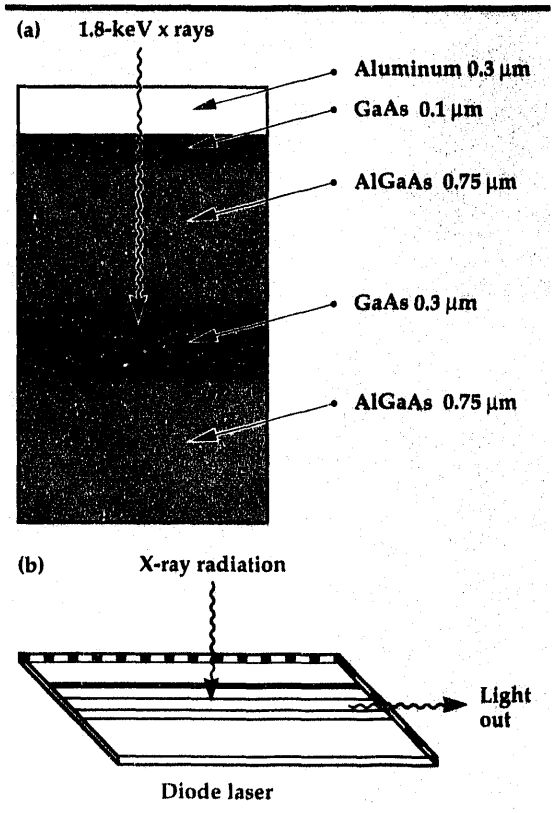


Figure 2. Simple representation of the physical structure of the x ray-to-light converter. (a) Schematic drawing of the layers in the device and calculations of absorption of 1.8-keV x rays. (b) Sketch of device.

X ray-to-Light Converter

In the research area of x-ray diagnostics, one desirable feature of a detector is to encode the temporal information about the intensities directly onto a coherent optical beam, which is transmitted on an optical fiber for remote recording. One approach, originally proposed by J. Koo,[1] had been to combine a solid-state photoconductor with a diode laser, where the electrical carriers generated in the photoconductor would be used to modulate the output of the laser; this was successfully tested.[2] A limit to the high-speed response of such a detector is the time it takes to move electrical carriers into the gain region of the diode laser.

We proposed using the excess carriers, which are generated by the absorption of x rays in the gain region itself, as the source of modulated laser output. This has the advantage, at least for a simple double-heterostructure laser, that no time is lost for carrier transport. Since the time for the 'hot' x ray-generated carriers to thermalize has been calculated to be less than one picosecond, the overall time response for such a device should be limited only by the stimulated emission lifetime of the carriers, which can easily be a few tens of picoseconds or less. The difficulty in designing and fabricating such a device is that the x rays must pass through the top cladding layer of the laser before they can be absorbed in the gain region to create useful electron-hole pairs to modulate the optical output (**Fig. 2**). If the x rays have little absorption in the cladding, they will also have little absorption in the gain medium. Similarly, if they are intensely absorbed in the gain medium, they will be intensely absorbed in

the cladding and will not reach the gain medium. Penetration depths of x rays vary rapidly with x-ray energy, so we expected our devices to be sensitive to only a rather narrow energy range of x rays in the low keV energy range. This is a range commonly produced by focused-laser plasmas.

Progress

Photocathode

During FY-92, we used our molecular beam epitaxy (MBE) system to grow heteroepitaxial structures with $In_x(Al_wGa_{1-w})_{1-x}As$ emitters and $In_yGa_{1-y}As$ absorber, including strained-layer superlattice buffers between the GaAs substrate and the IR absorber with electron emitter. Because activation and testing of photocathodes is a slow process, we began concentrating our efforts on $Ga_xAl_{1-x}Sb$ ($x \approx 0.3$) emitters as soon as we discovered that it worked about as well as GaAs (**Fig. 3**). The main appeal of the $Ga_xAl_{1-x}Sb$ emitter is that it is lattice-matched to single-crystal IR absorbers, which span the range of wavelengths from 0.9 μm to more than 10 μm. GaSb absorbs out to 1.7 μm, InAs reaches 4 μm, and $InAs/Ga_wIn_yAl_{1-y-w}Sb$ superlattices have been shown to absorb out to 12 μm. We have grown, activated, and tested numerous photocathodes of $Ga_xAl_{1-x}Sb$, and we have grown superlattices of InAs/GaSb, which we expect to absorb in the 0.9- to 2-μm range. We are still investigating doping levels in the various layers to minimize dark currents. Dark currents degrade the performance of these devices.

X ray-to-Light Converter

To fabricate an appropriate device, we used our MBE to grow a simple DH laser with 70% aluminum in the cladding. This high aluminum content increased the overlap of the optical beam with the carriers in the gain region and also allowed more x rays to pass through the cladding and enter the gain region for absorption. Because our simple modeling showed that our devices, with 0.7-μm thick cladding and 0.3-μm thick gain region, would exhibit peak sensitivity for x rays with energies of a few keV, we had to discontinue using our normal top-side metallization of titanium/platinum/gold and substitute pure aluminum. (The gold and platinum would have absorbed virtually all of the desired x rays.) This required our developing a new metallization procedure and a wire bonding procedure for the aluminum contact.

We also designed and built a heat sink to mount this device, and the heat sink had to be designed so that the epoxy that we used for F-O pigtailing would not flow onto the laser facet, yet would allow accurate

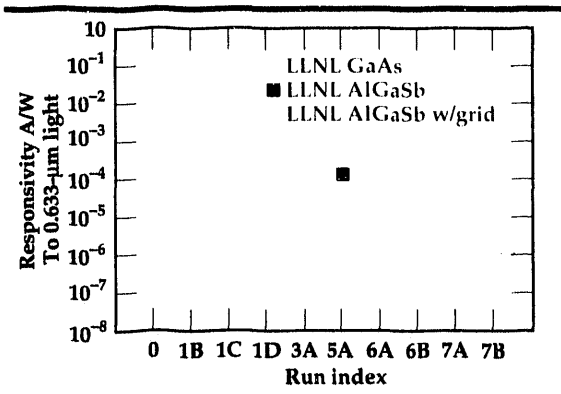


Figure 3. Plots of photoresponse for a variety of GaAs and GaAlSb photocathodes that were grown on our MBE and activated in our test apparatus.

placement of the F-O proximal to the emitting facet. We then mounted the assembled laser on a vacuum flange with a fiber-optic connection through the flange. Using x rays from a focused-laser plasma, we did observe the direct conversion of x rays to coherent optical light, as hoped (**Fig. 4**). Because the laser pulses were multi-nanosecond in duration, we could not determine the shortest time response of our detector. However, we did learn that the pulse-to-pulse variations in x rays that were generated by the focused laser caused far more variation in the output of our detector than in the simple photoconductive detectors. This, again, is due to the fact that our absorbing region is 0.8 μm beneath the surface of the laser (0.7-μm-thick cladding with 0.1-μm-thick electrical contact layer).

The data shown in **Fig. 4** represent the average of 100 laser pulses. When we tried the same experiment with a subpicosecond x ray source with less total fluence, we were not able to detect x rays.

Future Work

We are patenting the new $Ga_xAl_{1-x}Sb$/IR-absorber photocathode. An Engineering Research

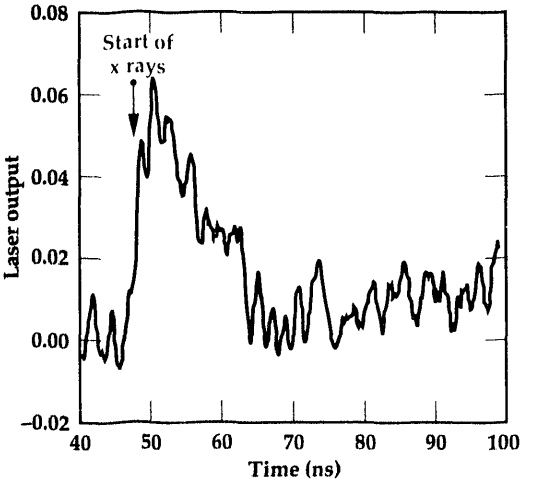


Figure 4. Plot of laser output vs time for our x ray-to-light converter, averaged over 100 laser pulses.

Division project at Lawremce Livermore National Laboratory is working on its development and seeking an industrial partner. The x ray-to-light converter is not currently being pursued further.

1. J.C. Koo, Private communication, Lawrence Livermore National Laboratory, Livermore, California (1988).
2. C.L. Wang, *Appl. Phys. Lett.* **54**, 1498 (1989).

Wideband Phase Modulator

Charles F. McConaghy,
Sol P. Dijaili, and
Jeffrey D. Morse
Engineering Research Division
Electronics Engineering

Lithium niobate integrated optics work has been an ongoing effort at Lawrence Livermore National Laboratory (LLNL) for many years. We have delivered completely packaged Mach–Zehnder modulators with bandwidths to 20 GHz, extinction ratios over 40 dB, and losses as low as 4 dB, to LLNL programs. These devices have traditionally been used to intensity-modulate laser sources in high-speed analog links. During the past year, we have been doing research on a very broad bandwidth, integrated-optic phase modulator. Such a device would have immediate applications to stimulated Brillouin scattering suppression in both optical fibers and glass amplifiers. In addition, these devices can be used to generate very short optical pulses from long or even cw laser pulses (pulse compression). Although neither of these applications is new, what is unique here is the efficient, integrated-optical phase modulator used to implement these techniques.

Introduction

Pulse Compression

Current commercial capability in generating picosecond optical pulses usually involves large, expensive table-top laser systems, which greatly restricts the applications of these systems. A picosecond, compact, inexpensive pulse compressor suitable even with cw laser sources can be built with an ultra-high bandwidth, $LiNbO_3$ phase modulator together with a dispersive element such as a grating or a fiber. In fact, these optical pulse generators can be built at any wavelength where suitable optical waveguides can be built. Previous attempts to compress and generate picosecond pulse trains using phase modulators used bulk devices and, hence, required many kW of microwave power to achieve picosecond pulses and optical bandwidths of 600 GHz.[1] One other author has tried a guided-wave device. However, a low drive power and an inefficient electrode structure limited the bandwidth to 12 GHz.[2]

The bandwidth of a phase modulator is directly proportional to the drive voltage and inversely proportional to $V\pi$, which is the voltage required to produce a phase change of 180° in the optical carrier. By using integrated optics as opposed to bulk optics, $V\pi$ can be reduced to 10 volts or less. Higher electrode voltages can be achieved for a given drive voltage by using resonant electrodes. We have been working to achieve a Q-factor on the order of 100 in a microwave transmission line resonator designed in an integrated fashion with the optical waveguide. At resonance, the voltage applied to the electrodes is approximately the source voltage multiplied by the square root of the Q-factor. We estimate that bandwidths on the order of 500 GHz can be achieved with microwave power levels on the order of several watts. Using a dispersive element, this bandwidth can give rise to picosecond pulses from optical sources. For a transform limited pulse, the minimum pulse width obtainable is given by .8/(bandwidth). Therefore, for 500 GHz of bandwidth, 1.6 ps pulses can be obtained. **Figure 1** shows the concept for pulse compression.

Stimulated Brillouin Scattering Suppression

Stimulated Brillouin scattering (SBS) is a nonlinear optical effect that limits the maximum optical power that can be transmitted in both glass amplifiers and glass fibers. For example, at 800 nm, experimental evidence exists that shows optical power

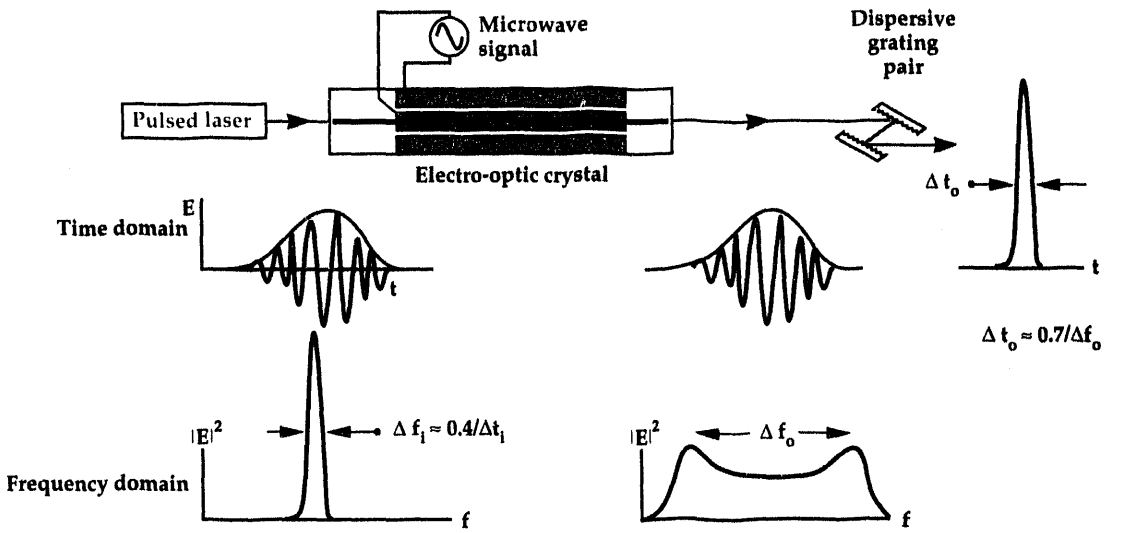


Figure 1. Optical pulse compression with a wideband phase modulator to chirp the incoming laser.

saturation at about 100 mW in fiber. The SBS is inversely proportional to optical linewidth. Therefore, it is more difficult to obtain power transmission with a narrow linewidth laser. To suppress the negative effects of SBS, the linewidth can be widened with phase modulation. Experiments are planned to see how the wideband phase modulator can be used to minimize the SBS problem.

Progress

The goal of FY-92 work has been to design and fabricate a high-Q electrode structure. We have gone through three iterations of electrode design this year. To maximize the overlap of the optical and electrical waves, the microwave electric field must be confined to a gap of about 10 μm. A tight gap electrode structure can be limited in Q since the microwave current bunches in the conductors close to the gap. We have both modeled and built microwave resonant electrodes from symmetric coplanar and asymmetric coplanar lines. We have studied shorted and open-ended lines. In addition, we have studied how the microwave energy is coupled to the resonant electrode.

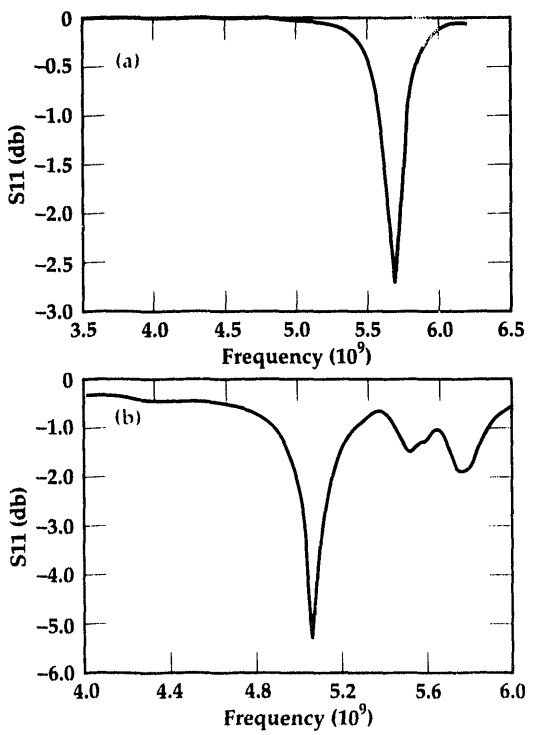


Figure 2. Plots of (a) calculated and (b) measured reflection coefficient, S11.

In our first iteration, two electrode structures were packaged in one box for test purposes. We discovered that the microwave energy excited not only the electrode of interest, but the additional electrode in the same box. The first iteration used both an asymmetric line and an asymmetric input-coupling scheme. In our second iteration, symmetric lines were used for both the input coupling line and the resonator. These devices were tested with high frequency rf probes and a network analyzer. Network analyzer measurements of these structures indicated two poorly defined resonances. Further computer modeling with a commercial electromagnetics program (Sonnet EM) indicated that resonances were achieved at slightly different frequencies for waves on either side of the resonant electrode. This was probably due to the perturbation of symmetry from the center feed point on one side of the electrode.

Another electrode pattern that was identical except that it had a 100-μm gap, did not have this problem. However, the wide gap is incompatible with building a high efficiency modulator. The third iteration produced a well defined notch in the reflection coefficient. This indicated that the electrode was indeed resonant and that a substantial amount of input power was coupled to it.

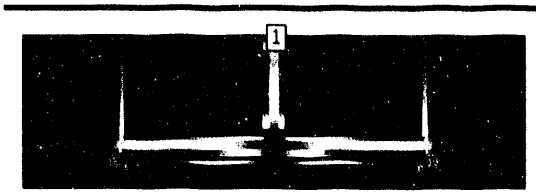


Figure 3. Current distribution on electrode at resonance.

In this electrode design, the input coupling consisted of a symmetric coplanar line, and the resonant electrode was an asymmetric coplanar line. This alleviated the problem that existed in the completely symmetric resonator.

Figure 2 shows both the calculated and measured reflection coefficient, S11. The resonance has a Q of about 25, which corresponds to a 5-times voltage enhancement. The depth of the notch is no greater than 5 dB, indicating that about one-third of the incident power is not coupled into the resonant electrode. A new mask with a wider coupling gap has improved this notch depth to about 10 dB or 90% coupling. The slight differences in notch depth and resonant frequency between the modeled and measured data can probably be accounted for by the fact the modeling code is two-dimensional and does not take into account the effect of electrode thickness.

Figure 3 shows a current distribution picture at resonance, with most microwave current crowded near the tight 10-μm gap.

Future Work

Currently, the modified asymmetric electrode structure shown in **Fig. 3** is being electroplated on top of an 800-nm optical waveguide. Once fabricated, the devices will have their bandwidths evaluated with an optical spectrometer. If sufficient bandwidth, in the neighborhood of 500 GHz, is achieved, the chirped pulses will be compressed with a grating to achieve picosecond optical pulses. In related integrated-optics work, we are exploring the use of modulators at superconducting temperatures. We would like to explore the possibility of building one of the resonant electrode phase modulators with niobium electrodes to determine what type of Q can be achieved at superconducting temperatures.

1. T. Kobayashi, H. Yao, K. Amano, Y. Fukushima, A. Morimoto, and T. Sueta, *IEEE JQE* **24** (2), 382 (1988).

2. B.H. Kolner, *Appl. Phys. Lett.* **52** (14), 1122 (1988).

Optoelectronic Terahertz Beam System: Enabling Technologies

Jeffrey D. Morse
Engineering Research Division
Electronics Engineering

In FY-92, we investigated the photoelectronic properties of semiconductor materials and structures for implementation as photoconductive dipole antennas in a terahertz beam system. We have measured optically generated electrical pulses propagating on-chip having temporal resolution as short as 1 ps. Furthermore, our devices have been used in a terahertz beam system to generate and detect electromagnetic pulses traveling through free space, with durations as short as 500 fs.

Introduction

With the advent of sub-picosecond laser sources, optoelectronic switching devices can be used to emit broadband electromagnetic (EM) pulses into free space. Integrated metal-semiconductor-metal photoconductive devices are capable of radiating EM pulses from monolithically integrated antenna structures.[1,2] The basic system concept is illustrated in **Fig. 1**. A static electric field is stored across the electrodes of the highly resistive, photoconducting antenna structure. When an optical pulse of intensity E_{opt} is incident between the electrodes, the conductance of the photoconductor increases, and the relation between the radiated electric field and the static electric field strength is described by[3]

$$E_r = -E_s \sigma_s \eta_0 / \left[\sigma_s \eta_0 + \left(1 + \sqrt{\varepsilon}\right)\right], \qquad (1)$$

where ε is the dielectric constant of the antenna material, η_0 is the free space impedance, E_s is the static electric field applied across the photoconductor, and σ_s is the conductivity of the photoconductor due to the photogenerated carriers. The conductivity is described by

$$\sigma_s = q\, \mu_{eff}\, (1 - r)\, E_{opt} / h\nu, \qquad (2)$$

where q is the electronic charge, r is the reflectivity, μ_{eff} is the effective mobility, hν is the photon energy, and E_{opt} is the optical energy density. From **Eqs. 1** and **2**, it can be seen that the effective mobility of the photoconductor material directly relates to the efficiency of the radiating antenna element. Therefore, this research will focus on optimizing the mobility of materials used as photoconducting antenna structures, which are suitable for compact arrays of emitters for ultra-wideband radar and remote applications. The advantages of using integrated optoelectronic antenna structures include ultra-wide bandwidth, high power, excellent directionality, low cost, and compact, durable elements conducive to large, photonically controlled arrays.

Progress

The power radiated by the photoconductive antenna element is directly related to the electronic transport properties of the semiconductor material being used. In particular, a semiconductor material that retains high mobility for the

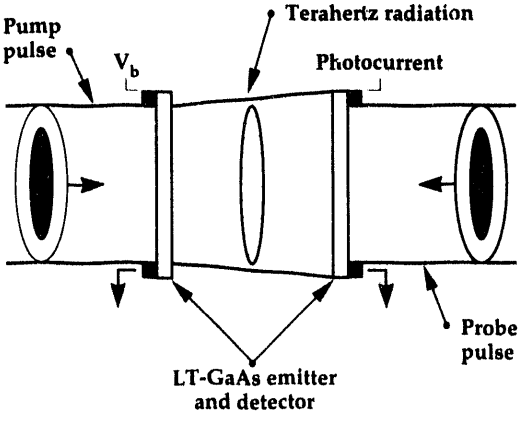


Figure 1. Schematic diagram of photoconducting antenna emitter and detector.

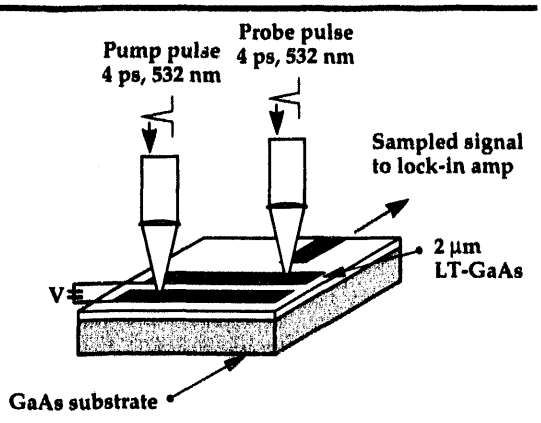


Figure 2. Autocorrelation circuit configuration.

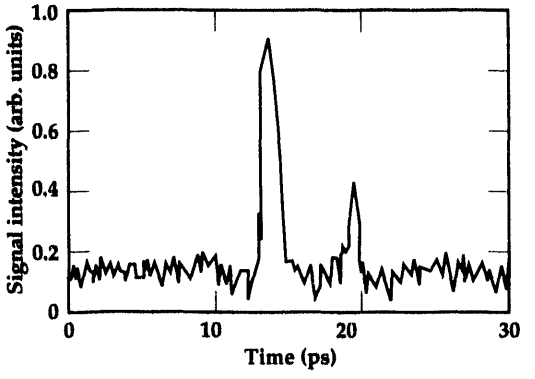


Figure 3. Impulse response of LT-GaAs photoconductor measured by reflective electro-optic sampling.

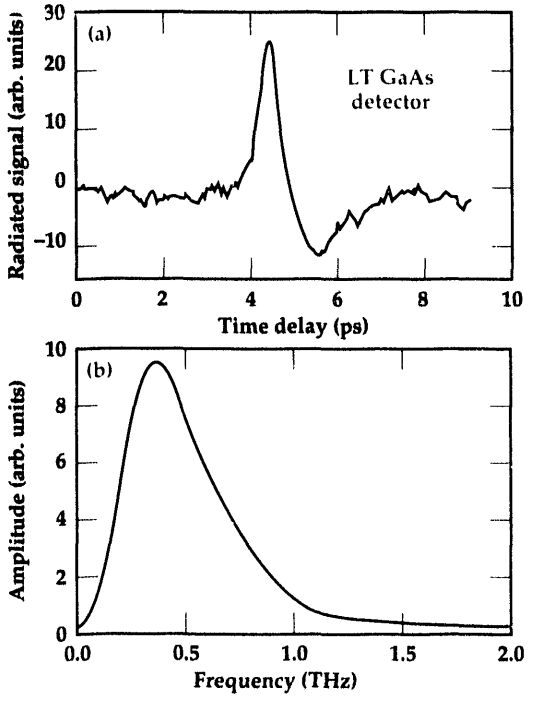


Figure 4. (a) Waveform and (b) corresponding Fourier Transform of radiated pulse detected with LT-GaAs detector in terahertz beam system.

photogenerated carriers while providing fast recombination lifetimes is desirable. In general, these two electronic properties conflict. Recently, it has been found that the growth of GaAs by molecular beam epitaxy (MBE) at low substrate temperatures (190° to 400°C) introduces large concentrations of point defects to the crystal structure through the incorporation of excess arsenic into the lattice.[4] During thermal annealing of the epitaxial layer at temperatures ranging from 580° to 800°C, the excess arsenic diffuses to form precipitates. Depending on the anneal time and temperature, the resulting arsenic precipitates have diameters ranging from 2 to 20 nm and spacings ranging from 5 to 50 nm. It is believed that these defects are metallic in nature,[5] behaving as fast recombination centers and resulting in sub-picosecond photoconductive response times. Furthermore, since the epitaxial LT-GaAs retains excellent crystalline quality, higher mobility is exhibited, which translates to higher sensitivity in comparison to alternative materials for picosecond photoconductivity.

Photoconductive autocorrelation circuits, illustrated in **Fig. 2**, have been fabricated from LT-GaAs grown by MBE at 190°C. This circuit uses the 'sliding' contact configuration,[6] which consists of two balanced coplanar lines with 5-μm width and 10-μm separation. With an electric field applied across the coplanar lines, an electrical transient signal can then be launched onto the lines by shorting the gap between them with an optical pulse. The photoconductive pulse can be generated at any point along the coplanar lines, hence 'sliding' contact. This is especially useful for characterizing the dispersive effects of the coplanar lines. To measure the transient electrical pulse, a sampling element is positioned along one of the lines in the coplanar pair (**Fig. 2**). The sampling can be either photoconductive[7] or electro-optic.[8] By varying the relative delay between the generating and sampling optical pulses, the photoconductive transient response is measured. **Figure 3** illustrates the electrical pulse measured for this material, by the reflective electro-optic sampling technique.[9] The optical pulse width is ~ 600 fs at 820-nm wavelength, and the signal has propagated approximately 100 μm on the transmission line. The response is < 1 ps full width at half maximum. Calibration of the incident optical intensity gives an estimated mobility for this material of ~ 120 cm /Vs, which is a factor of 4 to 10 times larger than that of other materials used for picosecond photoconductivity.[7]

Initial measurements of our devices in a terahertz beam system have been conducted by researchers at Columbia University.[10] Results from these experiments have demonstrated that our devices are capable of detecting transient electric fields having amplitudes in excess of 1 kV/cm with temporal resolution of 600 fs. The measured

response is illustrated in **Fig. 4a**, the corresponding frequency domain response in **Fig. 4b**. From these results, it can be seen that these pulses have useful frequency content beyond 1 THz. This renders terahertz beam systems useful for further applications such as far-infrared spectroscopy, imaging, and ultra-wideband communications.

Future Work

This research has demonstrated the suitability of our devices as high-performance, photoconducting antenna elements. The next step is to implement photonically controlled phased arrays based on this technology. This will be done by implementing integrated optics technology as the active system component to embed the rf signal, and phase modulation on the optical carrier to achieve beam-steering functionality. The combination of these technologies will make this system extremely useful for airborne and space-based applications.

Acknowledgements

This work would not have been possible without the contributions of Dr. Raymond Mariella (Lawrence Livermore National Laboratory) and Dr. Michael Spencer (Howard University) in materials growth, and Dr. X.-C. Zhang (Renselaer Polytechnic Institute), J.T. Darrow (Raytheon), and Dr. D.H. Auston (Columbia University) who provided terahertz beam system measurements.

1. A.P. DeFonzo and C.R. Lutz, *Appl. Phys. Lett.* **51**, 212 (1987).
2. C.H. Fattinger and D. Grischkowsky, *Appl. Phys. Lett.* **53**, 1480 (1988).
3. X.-C. Zhang, B.B. Hu, J.T. Darrow, and D.H. Auston, *Appl. Phys. Lett.* **56**, 1011 (1990).
4. F.W. Smith, H.Q. Le, V. Diadiuk, M.A. Hollis, A.R. Calawa, S. Gupta, M. Frankel, D.R. Dykaar, G. Mourou, and T.Y. Hsiang, *Appl. Phys. Lett.* **54**, 890 (1989).
5. A.C. Warren, J.M. Woodall, J.L. Freeouf, D. Grischkowsky, D.T. McInturff, M.R. Melloch, and N. Otsuka, *Appl. Phys. Lett.* **57**, 1331 (1990).
6. M.B. Ketchen, *Appl. Phys. Lett.* **48**, 751 (1986).
7. D.H. Auston, in *Picosecond Optoelectronic Devices*, C.H. Lee (Ed.), Academic Press (London, England), 1984.
8. J.A. Valdmanis, G. Mourou, and C.W. Gable, *Appl. Phys. Lett.* **41**, 211 (1982).
9. L. Min and R.J.D. Miller, *Appl. Phys. Lett.* **56**, 524 (1990).
10. J.T. Darrow, X.-C. Zhang, D.H. Auston, and J.D. Morse, *IEEE JQE* **QR-28** (6), 1607 (1992).

Fabrication of Microelectrode Electrochemical Sensors

Dino R. Ciarlo,
Jackson C. Koo, and
Conrad M. Yu
Engineering Research Division
Electronics Engineering

Robert S. Glass
Materials Division
Chemistry and Materials Science Department

We are using integrated circuit technology to fabricate microelectrode electrochemical sensors. These sensors have improved performance compared to those that use a single macroelectrode. The near-term application for these new sensors is for environmental monitoring, especially for heavy metal contamination.

Introduction

An electrochemical sensor consists of a pair of dissimilar electrodes immersed in a solution containing unknown ions, as shown in **Fig. 1**. The relationship between the current in the working electrode (I_W) and applied potential (V_W) referenced to a reference electrode, depends on the ions in solution and on the composition of the electrodes. This sensor is particularly well suited for the measurement of heavy metal contamination and pH as needed in environmental monitoring.[1,2] The goal of this project was to use integrated circuit (IC) microfabrication technology to fabricate multielectrode electrochemical sensors.[3] The advantages of using microelectrodes in electrochemical sensors are: (1) immunity from uncompensated resistance effects because of low currents used; (2) high rates of mass transfer and hence higher sensitivity; (3) higher signal-to-noise ratios; (4) the potential for extremely fast experiments; and (5) the extension of normal electrochemical background limits.

In addition, the use of a matrix of different electrode materials improves the information content of the measurement. Also, because of the mass-production capability of microfabrication, reproducible sensors can be produced inexpensively and used in a disposable fashion.

Progress

This past year, we used the photolithography and vacuum evaporation capabilities available in the MicroTechnology Center of Lawarence Livermore National Laboratory to fabricate microelectrode electrochemical sensors. **Figure 2** shows a computer drawing of the sensor electrodes. In one design, silver was used for the reference electrode, platinum for the counter electrode, and the four working electrodes were platinum, platinum, iridium oxide, and silver. In one application, the iridium-oxide working electrode is used as a pH sensor: one of the platinum working electrodes is coated with a mercury thin film

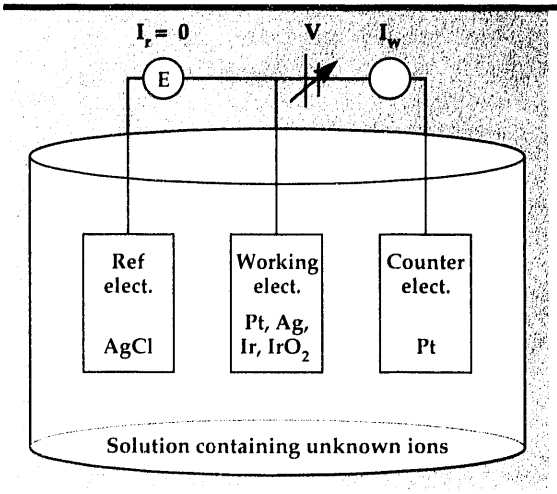


Figure 1. Typical electrochemical sensor arrangement. The relationship between the current I_W and the applied potential V_W depends on the electrode materials and the ions in solution.

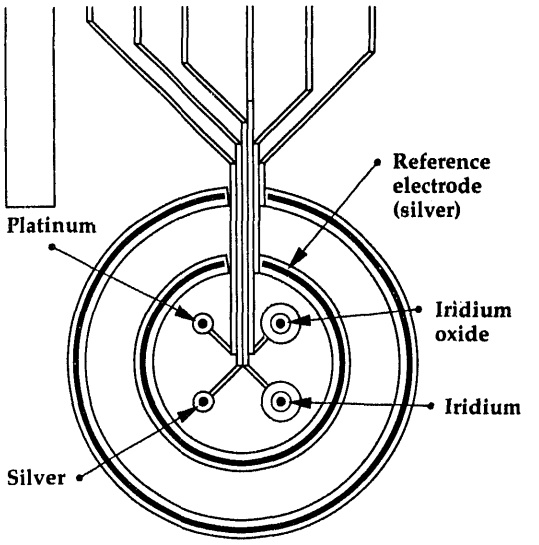


Figure 2. Electrode configuration of the microelectrode electrochemical sensor.

to detect lead, cadmium, zinc, and copper ions; the other platinum working electrode is coated with a polymer to detect heavy metals; and the silver working electrode is used to detect chlorine.

We devoted considerable effort to the development of a reliable fabrication process. Silicon wafers, 1-mm-thick, were used as substrates so that conventional IC processing equipment could be used. The thickness was chosen so that they would be robust enough to allow handling without breakage, and so that commercial connectors could be used to interface the sensors with the processing electronics. A combination of 5000 Å of thermal oxide plus 2000 Å of silicon nitride was used to electrically isolate the sensor films from the silicon substrate. Two sensors were fabricated on each wafer, and we could process six wafers at a time in the vacuum system. When completed, the sensors were cut with a dicing saw to their final size of 0.5 in. × 1.5 in.

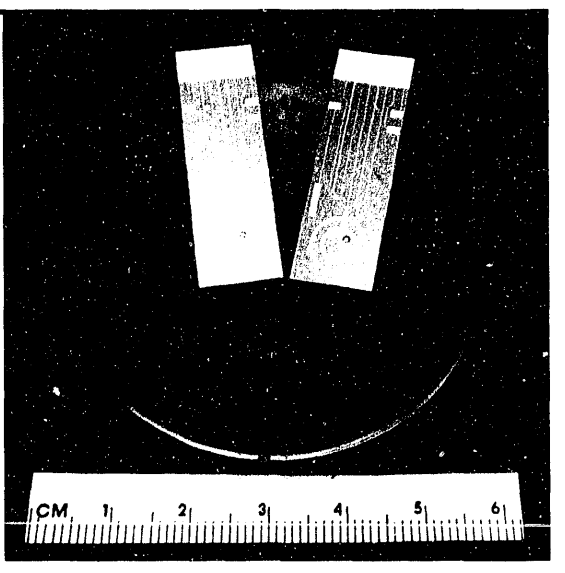


Figure 3. Photograph of two completed microelectrode electrochemical sensors. The pads at the top interface to a commercial connector.

Since the materials used for this sensor fabrication were not those that one typically finds in IC processing, we had to deal with new problems of film adhesion, cracking, and compatibility. Originally, we tried an etch procedure to define the sensor material, but this was difficult because of the incompatibility of the resist with some of the etches. In addition, it was difficult to completely etch away the films, and this caused some conductivity between the various electrodes. We eventually settled on an all-lift-off procedure for the sensor fabrication. With this approach, openings are patterned in the photoresist layer, the sensor material is then evaporated onto the entire wafer, and the unwanted material is lifted off by dissolving the resist in acetone. This eliminated the need for any chemical etching. After the sensor materials were defined in this manner, a photosensitive layer of polyimide was applied to the wafer and patterned with the openings required for the sensor electrodes and the connector pad area. The typical circular area of the exposed working electrode had a diameter of 50 μm. **Figure 3** is a photograph of two completed sensors.

Following fabrication, the sensors were interfaced to a data collection system, and experiments were performed. In many cases, a 'textbook' response was obtained from the sensors for the ions of interest.

Future Work

Future work will involve experiments with electrodes having different sizes and shapes, to try to optimize the performance of the sensor. We will also work with other sensor materials that are more specific to the ions of interest. The fabrication process will be refined to improve the overall yield of useful sensors. We will also modify the vacuum evaporator so that we can simultaneously fabricate 24 instead of 12 sensors. Finally, we will start working with an industrial partner, since some early versions of this sensor appear ready for commercialization.

1. R.S. Glass, S.P.Perone, and D.R. Ciarlo, *Anal. Chem.* **62**, 1914 (1990)).

2. R.S. Glass, K.C. Hong, W.M. Thompson, R.A. Reibold, J.C. Estill, D.W. O'Brien, D.R. Ciarlo, and V.E. Granstaff, *Electrochemical Array Sensors for Plating Waste Stream Monitoring*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-108819 (1992).

3. R.S. Glass, S.P. Perone, D.R. Ciarlo, and J.F. Kimmons, "Electrochemical Sensor/Detector System and Method," U.S. Patent #5,120,421, June 9, 1992.

Diamond Heatsinks

Dino R. Ciarlo,
Jick H. Yee, and
Gizzing H. Khanaka
Engineering Research Division
Electronics Engineering

Erik Randich
Materials Division
Chemistry and
Materials Science Department

We are studying patterned diamond films for use as heatsinks to cool solid-state laser diodes. We have etched diamond slabs using our chemically assisted ion beam etcher. An inductively coupled plasma torch has been set up for the high rate deposition (> 50 μ/h) of diamond films onto patterned silicon wafers. Our modeling effort was used to design the optimum dimensions for both types of heatsinks.

Introduction

For some time, Lawrence Livermore National Laboratory has been using silicon microchannel heatsinks to cool solid-state laser diodes. Very intricate microchannels have been etched into the surface of silicon wafers to provide paths for cooling water. The packaging has been designed so that diode bars can be stacked together to maximize the radiated flux.[1] In some designs, a heat dissipation approaching 3000 W/cm^2 has been achieved.

To expand our heatsink options, we have been studying the use of patterned diamond as a heatsink material. Diamond makes an excellent heatsink because it has the highest thermal conductivity of any known material at room temperature, i.e., 20 W/cm°C vs 1.5 W/cm°C for silicon. It is also an excellent electrical insulator (1×10^{16} Ω-cm) and will not corrode. Until recently, diamond had not been used extensively for heatsinks because of its high cost. However, recent advances in the deposition of diamond films using chemical vapor deposition (CVD) techniques has lowered its cost. These CVD films are now commercially available from several vendors and are being used as heatsinks. The thermal conductivity of CVD diamond is somewhat lower than that of natural diamond, i.e., 14 W/cm°C vs 20 W/cm°C, but it is still high enough to make the material very attractive. All of the commercially available diamond is in the form of flat slabs. Our emphasis is on patterned diamond slabs. The patterning can be used for water flow channels or for slots into which laser diode bars are inserted.

Progress

During FY-92, we worked on three aspects of the diamond heatsinks problem: modeling, etching, and film deposition. The modeling effort consisted of refining a code, originally developed by Landram,[2] to make it more user friendly and more efficient for analyzing silicon and diamond. **Figure 1** shows the heatsink configuration used in the modeling work. In this Figure, a heat generating device is shown bonded to a microchannel heatsink. There are five thermal impedances in this structure that limit heat flow: (1) μ_{spread} (μ_{sp}), the spread of heat from a point source generator; (2) μ_{bulk} (μ_{bu}), the flow of heat through the bulk of the heat generating device; (3) $\mu_{interface}$ (μ_{in}), the flow of heat across the eutectic bonding material; (4) $\mu_{convection}$ (μ_{conv}), the flow of heat from the eutectic bonding material to the cooling fluid; and

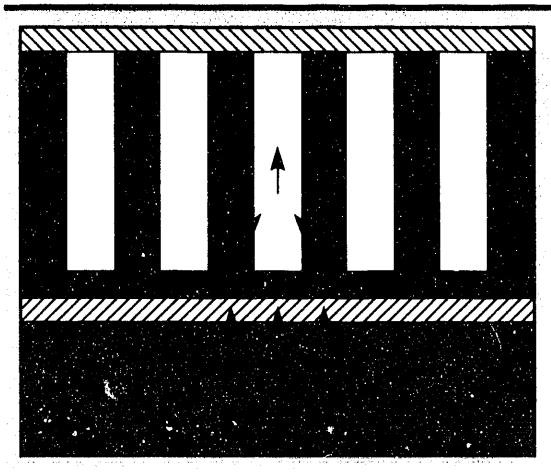


Figure 1. Cross section of a solid-state device bonded to a microchannel heatsink.

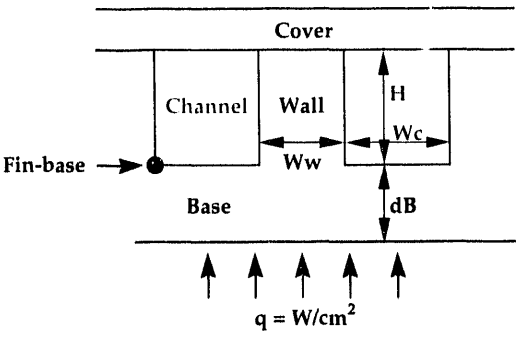


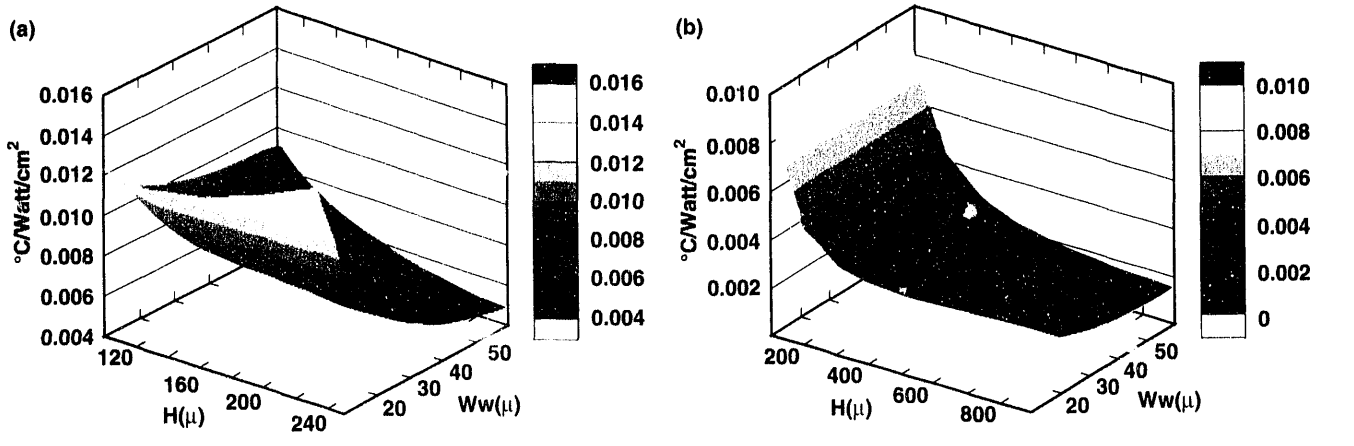
Figure 2. Cross section of a microchannel heatsink illustrating the design parameters.

(5) $\mu_{caloric}$ (μ_{cal}), the removal of heat by the cooling fluid. Our modeling effort concentrated on μ_{conv}, which concerns the optimum design for the microchannels. We were able to compare microchannels fabricated from silicon and diamond.

Figure 2 is a cross section of a microchannel heatsink illustrating the parameters used in the modeling. The model calculates the thermal impedance for the microchannel heatsink, μ_{conv}. Knowing μ_{conv}, we can immediately determine the difference between the average temperature of the cooling fluid and the temperature of the fin-base, as identified in **Fig. 2**. This temperature difference is given by $\Delta T = (\mu_{conv})^*(q)$, where q is the heat flux applied to the heatsink in W/cm^2. Thus, a low value for μ_{conv} is desired, and the optimum values for the dimensions of the microchannels are determined by those that give the lowest μ_{conv}.

Figures 3a and **3b** are three-dimensional plots of the thermal impedance for silicon and diamond microchannel heatsinks. Both are for a channel width (W_c) of 20 μ. For silicon heatsinks with a wall thickness (W_w) of 20 μ, the optimum channel height (H) is ~ 180 μ. Beyond this, not much is gained. For diamond, also with a W_c of 20 μ and a W_w of 20 μ, H is ~ 700 μ. If these values could be achieved, the diamond heatsink would out-perform the silicon heatsinks by a factor of four. This height, however, is rather impractical from a fabrication point of view, but our modeling has shown that if diamond heatsinks were fabricated with the same wall height as silicon, the thermal impedance would be lowered by a factor of two.

To create flow channels in diamond, we performed a number of etching experiments using our chemically assisted ion beam etcher (CAIBE). The diamond was purchased from two different vendors. One vendor supplied free-standing slabs that were 5 × 5 mm and 0.3-mm thick. The diamond was deposited by CVD. The other vendor supplied diamond bonded to a 1.5-mm-thick tungsten carbide substrate. This film was 600-μ thick, and the diameter of the part was 26 mm. It was deposited by a hot pressed technique and then polished smooth. We deposited a 1000 Å-thick chromium film on both types of parts, patterned the chromium using photolithography and a wet chemical etch, and then etched the diamond in our CAIBE. The etching experiments followed work by Geiss.[3] In this experiment, the diamond is bombarded with xenon ions that have been accelerated to an energy of 700 eV. At the same time, the sample is flooded with a source of oxygen, such as N_2O or NO_2 gas. The bombarding xenon ions promote chemical etching and, since they are collimated and directional, the etching proceeds in a directional manner. Under these conditions, an etch rate of approximately 200 Å/min is achieved. When the xenon ion energy was increased to 1800 eV, the etch rate doubled to 400 Å/min. Unfortunately, the chromium mask also erodes away, limiting how deep one can etch. With a 1000 Å-



Figures 3a and 3b. Three-dimensional plots of the thermal impedance of (a) silicon and (b) diamond microchannel heatsinks. The channel width is 20 μ in both cases.

thick chromium mask, we could only etch to a depth of 0.5 μm. We are looking at other mask materials such as oxides that may be more durable to the ion beam.

We also studied the deposition of diamond films onto patterned silicon substrates. The plan was to deposit thick diamond onto a silicon substrate that already had deep flow channels etched into its surface. The silicon could then be etched away, leaving a patterned diamond film. **Figure 4** shows cross sections of the silicon wafers prepared for the deposition. The etched grooves are all 20-μ wide, on 40-μ centers. The depth-to-width ratios used were 20/20, 40/20, 100/20, and 200/20. In May of 1992, we gained access to an inductively coupled plasma (ICP) torch to be used for the diamond film deposition. This equipment has been used by others to deposit diamond films at rates as high as 50 μ/h.[4] The high-velocity directional flow of the gases should make this technique ideal for the deposition of diamond into preformed grooves. **Figure 5** shows a diagram of this machine. It uses argon, hydrogen, and methane as source gases and is powered by a 50-kW, 4-MHz generator. From May to October of FY-92, we worked on the gas control system, and built cooling chambers and wafer holders. A number of calibration runs were made to adjust the plasma operating conditions. Actual film depositions are planned for early FY-93.

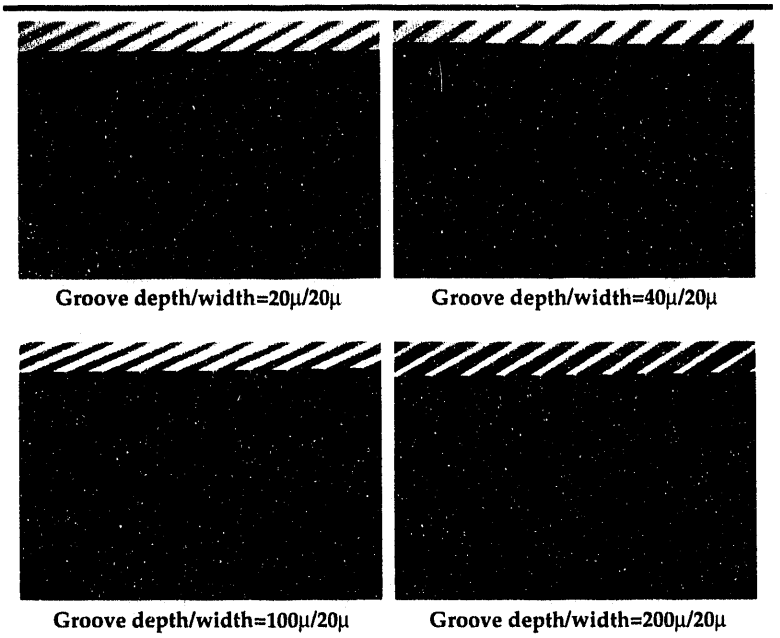


Figure 4. Scanning electron microscope cross sections of silicon wafers to be coated with thick diamond films.

Future Work

Our modeling effort needs to be extended to include the other four thermal impedances discussed above. This will help designers optimize the complete diode package instead of only the

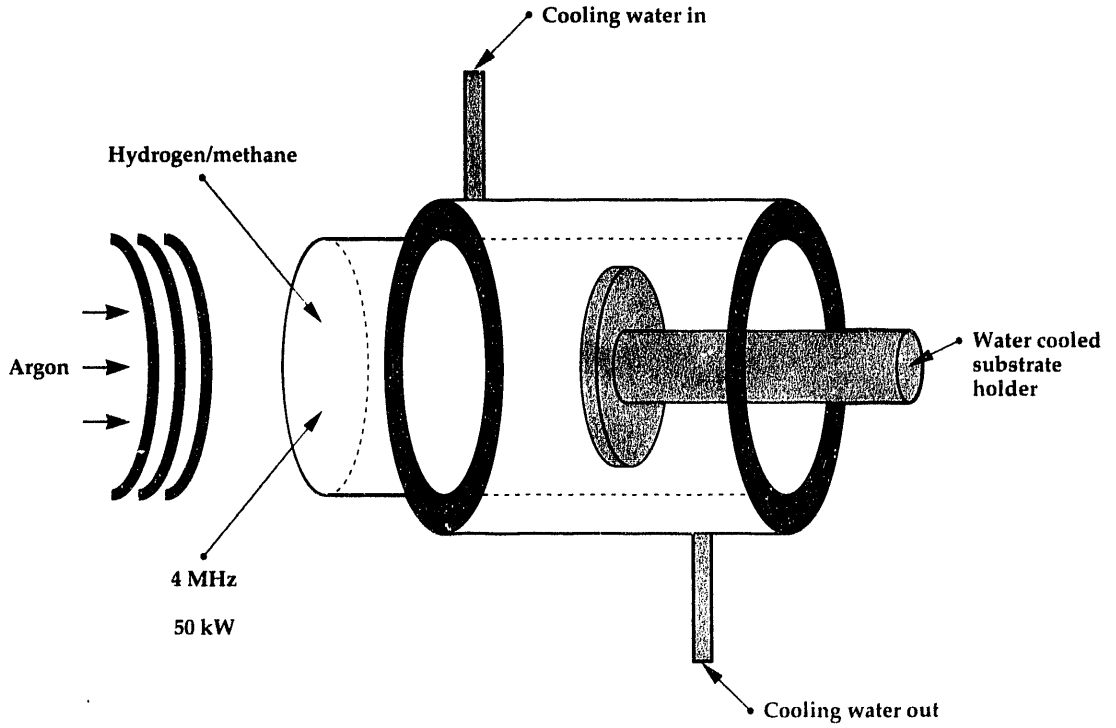


Figure 5. Diagram of the ICP torch used for high-rate diamond film deposition.

heatsink itself. Experiments need to be conducted with other mask materials in our CAIBE to find those with low etch rates so that we can etch deeper structures. Finally, and most important, we need to use the ICP torch to deposit thick diamond films into silicon grooves and then etch away the silicon. If successful, this will be the first report of a diamond slab with deep vertical channels.

1. G. Albrecht, R.J. Beach, and B. Comaskey, *Energy and Technology Review*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-52000-92-6, 7 (1992).

2. C.S. Landram, *An Exact Solution for Conjugate Longitudinal Fin-Fluid Heat Transfer in Internal Flow Including Optimization*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-103249 (1990).

3. N.N. Efremow, M.W. Geiss, D.C. Flanders, G.A. Lincoln, and N.P. Economou, *J. Vac. Sci. and Technol.* **B3** (1), (1985).

4. M.A. Capelli, T.G. Owano, and C.H. Kruger, *J. Mater. Res.* **5** (11), 2326 (1990).

Advanced Micromachining Technologies

Wing C. Hui
Chemical Sciences Division
Chemical and Materials Science Department
and Engineering Research Division
Electronics Engineering

We have developed several innovative micromachining techniques that will facilitate the future development of high-tech microtechnologies, such as microelectronics, microstructures, microactuators, microsensors, and microinstruments. The Corner-Protection Technique will produce sharp convex corners and clear scribe lines in any anisotropic etching process. The Circular Etching Process alone, or combined with Selective Wet Chemical Etching for Boron Nitride Film, can be used to fabricate numerous forms of new, round features that were previously unattainable.

Introduction

Over the last two decades, single-crystal silicon has been increasingly used in a variety of new applications besides microelectronics. Single-crystal silicon is not just a good semiconductor material; it is also an excellent mechanical material for microscale devices, such as microstructures, microactuators, microsensors, and microinstruments. To facilitate the development of these new high-tech devices, newer and better micromachining technologies have to be created for the fundamental fabrication processes.

Recently, we have developed several new micromachining processes. These new processes will allow us to make microscale features that were previously unattainable. The Corner-Protection Technique will allow us to make precise sharp-corner features, without rounding the corner by undercutting. The Circular Etching Process, which uses isotropic etching with boron nitride as the masking film, is well engineered to fabricate circular thin film membrane windows or circular microstructures. Combined with Selective Wet Chemical Etching for Boron Nitride Film, this Circular Etching Process can build clear circular microstructures with or without additional films.

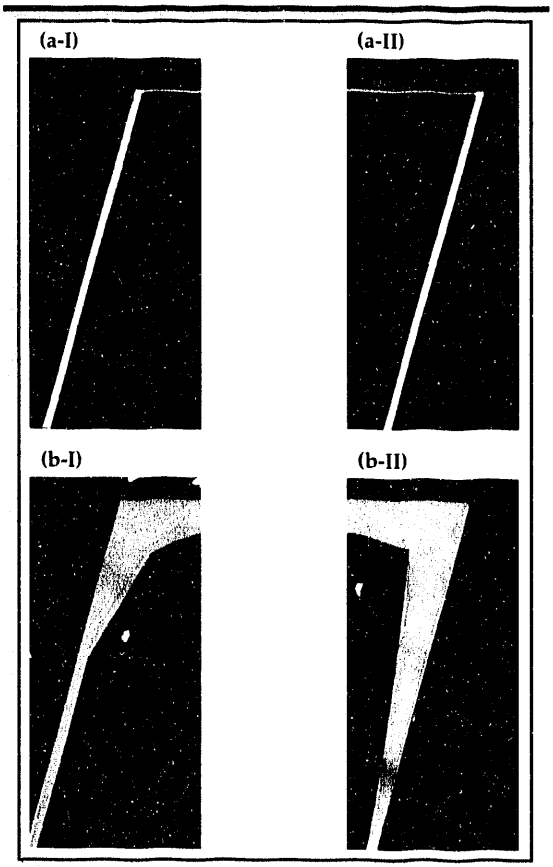


Figure 1. Comparison of the Lawrence Livermore National Laboratory (LLNL) Corner-Protection Technique and the regular etching technique on the anisotropic etching of a (110) silicon wafer: (a-I) the clear 109.4° corner etched by the LLNL technique; (a-II) the sharp 70.6° corner etched by the LLNL technique; (b-I) the undercut 109.4° corner etched by the regular technique; and (b-II) the undercut 70.6° corner etched by the regular technique.

Progress

Corner-Protection Technique

Microscale features on silicon wafers are very often achieved by means of anisotropic wet chemical etching. However, most of these etching pro-

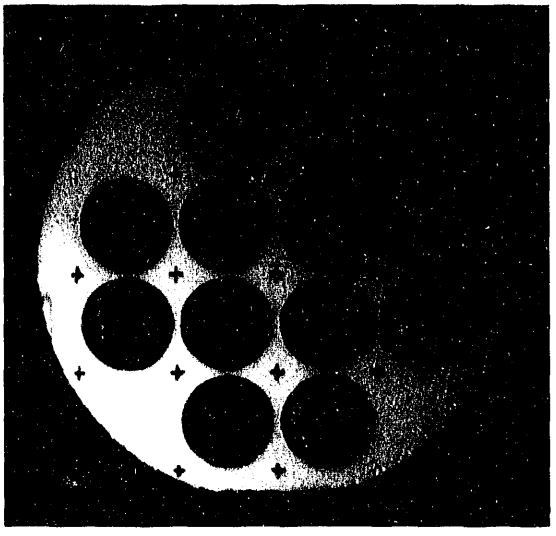


Figure 2. An example of the round features etched by the Circular Etching Process.

cesses have a severe undercutting problem at any convex or outer corner of a chip or device feature. This undercutting problem will limit the compactness and effectiveness in the overall design.

We have successfully developed a novel method for protecting these convex corners with very little space. This technique can produce sharp convex corners and clear scribe lines at any desirable etching depth. It can make many previously impossible geometries possible.

Last year, we demonstrated the technique on (100) silicon wafers to make narrow-frame thin film membrane windows, also known as 'thin-wall windows.' This year, we extended the technique to (110) silicon wafers. **Figure 1** compares results with and without the Corner-Protection Technique. This successful demonstration has provided a possible licensing opportunity with Endevco, a designer and manufacturer of instrumentation for vibration, shock, and pressure measurement.

Circular Etching Process

Most micromechanical devices rely on traditional anisotropic KOH etching for fabricating the microfeatures. This etching technique will produce only features with straight boundaries. Since it is desirable to have round features in many applications, we have put a great deal of effort into the development of special circular etching techniques to create new and unique microstructures.

First, we carefully engineered an isotropic etching process with $HF/HNO_3/CH_3COOH$ to produce even etching in all directions. Boron nitride thin film was used as the etching mask because of its chemical compatibility with the etchant. Circular boron-nitride

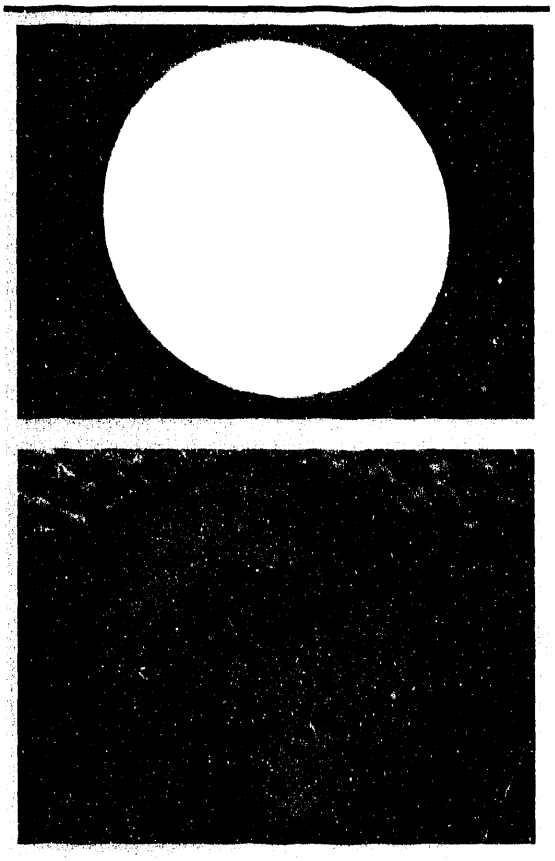


Figure 3. Comparison of the boron nitride-coated, round silicon disks etched by LLNL Selective Wet Chemical Etching and Dry Plasma Etching: (a) the shiny, small, round silicon disk after Selective Wet Chemical Etching; and (b) the foggy, small, round silicon disk after dry plasma etching.

thin film windows and other round microstructures were fabricated in this way (**Fig. 2**).

For this circular etching technique to be more useful as a tool when making round features for general applications, it is sometimes desirable to remove the masking boron nitride thin film. Conventionally, the boron nitride film can be removed only by dry plasma etching. However, plasma etching is not very selective—it also etches silicon nitride film, silicon-based substrate or film, and even gold film.

Our contribution to the solution of this problem is the development of the first wet chemical process that will selectively etch only boron nitride, but not coatings or substrates of silicon, silicon nitride, and silicon dioxide. The etchant is a very strong oxidizing reagent of sulfuric acid and hydrogen peroxide. It can remove the boron nitride film very selectively and smoothly, without leaving any over-etched surfaces, as the plasma etching process does (**Fig. 3**).

This modified round-etching process was also demonstrated to be very useful in the development of the microcapillaries for the Miniaturized Gas Chromatography Project of Conrad M. Yu and the author.

Electrophoresis Using Silicon Microchannels

Jackson C. Koo,
J. Courtney Davidson, and
Joseph W. Balch
Engineering Research Division
Electronics Engineering

We are developing electrophoresis techniques in microchannels that can be micromachined in substrates. We have developed a model of electro-osmotic flow in free-solution capillary electrophoresis when an external electric field is applied to the walls of the capillary. Also, we have begun development of gel-filled microchannels to be used for electrophoresis of biological materials.

Introduction

Electrophoresis is the separation of charged ions or molecules in a solution based on their differential migration in an applied electric field. It is widely used in modern analytical chemistry to separate charged particles in ionic solutions, and in biochemistry to separate biomolecules. The emergence of the biotechnology industry has greatly increased the interest in various electrophoresis methods. Our objective is to develop methods for novel electrophoresis of liquid ionic solutions and charged biological material, in microchannels in silicon and other substrate materials. Silicon microchannels for electrophoresis have potential advantages over conventional quartz capillaries due to (1) enhanced thermal dissipation of heat generated during electrophoresis, because the thermal conductivity of silicon is nearly 100 times that of quartz; and (2) the construction of a high-density array of microchannels on a single substrate in silicon and other materials, using modern microfabrication technology. Silicon-based microfabrication technology also provides a promising way to incorporate field plates around an electrophoresis microchannel to control the electro-osmotic flow of solution in free-solution electrophoresis. Control of electro-osmotic flow in free solution electrophoresis is a promising means to improve separation resolution and to allow separation of solutes in relatively short capillaries (i.e., 1 to 10 cm).

Progress

Our previous efforts[1] demonstrated control of electro-osmotic solution flow in round quartz capillaries and rectangular silicon microchannels, by applying an external electric field perpendicular to the inner walls of the electrophoresis capillary or rectangular channel. The electro-osmotic flow of the solution is caused by the force of an axial electric field upon a diffuse, dipole charge sheath in the solution adjacent to the capillary wall. This sheath is created by the electrostatic attraction of charge surface states of the capillary wall on solvated ions in the electrolyte. As the mobile part of the sheath is moved along the capillary wall by the applied electric field, the solvated ions of the diffuse layer transfer momentum to the remainder of the electrolyte solution. Therefore, the whole solution moves with the diffuse layer, and plug-like flow is created in the capillary. We can control the mobility of the electro-osmotic flow by applying an external electrostatic field perpendicular to the capillary wall, so the diffuse dipole layer in the electrolyte is modulated or eliminated.

Figure 1 shows a typical experimental result we obtained previously for the electro-osmotic mobility in a quartz capillary as a function of the external voltage applied perpendicular to the capillary walls. Our experimental setup was quite similar to that of the University of Maryland, where two concentric quartz tubes each filled with elec-

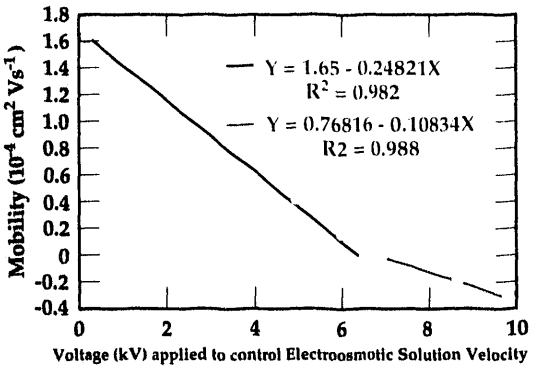


Figure 1. Measured electro-osmotic mobility of an electrolyte solution (2 mM concentration) in a 50-μm internal diameter capillary as a function of voltage applied across the 167.5-μm-thick capillary wall.

trolytes were used.[2,3] The inner tube was for the electro-osmotic flow; the outer tube was used to provide the external bias electric field. In our experiment, we continuously increased the bias voltage up to 10 kV. At about 7 kV, the electro-osmotic flow changed its direction. Here, we noticed that the rate of change of the mobility was not exactly a linear function of the externally applied voltage. However, we could fit the curve with two linear lines: one when the mobility is positive and another when the mobility is negative. This is expected because, as the direction of electro-osmotic flow changes, the moving cations are replaced with anions.

Model for External Electric Field Control of Electro-osmosis

This year, we developed a theoretical model of external electric field control of electro-osmotic flow to gain additional insight into the physical mechanisms of electro-osmosis. This model allows us to relate the functional dependence of measured electro-osmotic mobility vs the applied external voltage, to the density of surface states on the quartz wall and the ionic concentration of the electrolyte. Our model of the capacitive and interface-charging phenomena in a conductor-insulator-electrolyte capillary structure is adapted from a model widely used for similar phenomena in Metal-Insulator-Semiconductor capacitors used in microelectronics.[4] This type of model can easily be interpreted in terms of an electrical equivalent circuit. **Figure 2** shows this model's equivalent circuit for our experiment, where two concentric capillaries are each filled with electrolytes. In this equivalent circuit, the externally applied voltage is shared among a series of capacitors.

An extensive set of equations was derived that relates the various elements in the equivalent circuit to important physical quantities such as electrolyte concentration, viscosity of the electrolyte, the thickness and dielectric constant of the quartz wall, the surface state density of the quartz wall, and the applied voltage that controls the electro-osmotic flow. From these equations, we are able to relate the change in electro-osmotic mobility as a function of applied voltage (i.e., the slope of the curves in **Fig. 1**) to the density of surface states on the quartz wall for a given electrolyte concentration. For experimental results such as shown in **Fig. 1**, our analysis shows that when the anion traps are replaced by cation traps, there is a significant change in the density of the surface state per volt, while the total surface state densities per square centimeter (Ns) under the two different conditions are nearly the same. This indicates that the total number of active surface states, Ns, on the silicon dioxide are determined by the initial chemical conditions, and the active sites are bipolar in nature. The calculated value (2.4×10^{12} cm^{-2}) of the

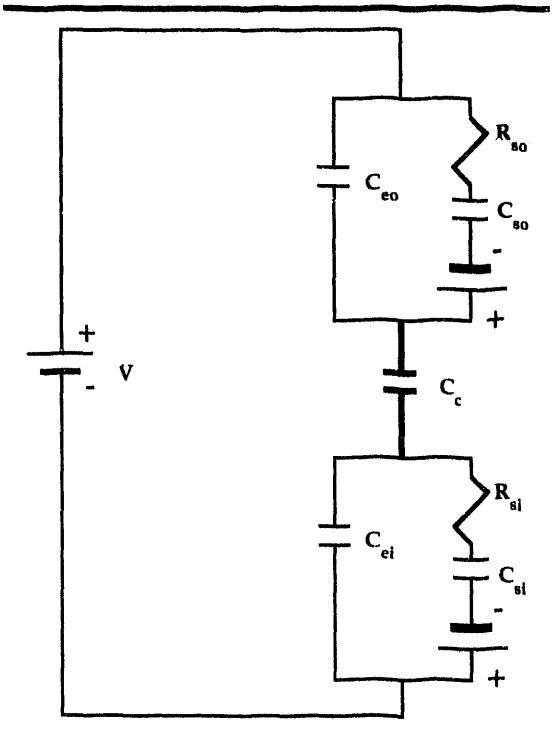


Figure 2. The electrical equivalent circuit model of two annular capillaries, each filled with electrolytes. The outer capillary is used to apply an external electric field to the walls of the inner capillary to control electro-osmosis in the inner capillary. C_{eo} represents the capacitance of the electrostatic diffuse layer between the capillary wall and the outer aqueous interface. C_c represents the capacitance of the inner capillary tube. C_{ei} represents the capacitance of the electrostatic diffuse layer between the inner capillary wall and the inner aqueous interface. R_{si} and R_{so} are surface resistances for the inner and outer surfaces, respectively. The surface capacitances C_{so} and C_{si} represent the surface state densities per volt on outer and inner surfaces of the capillary.

surface state densities is, however, much smaller than that normally accepted (5×10^{14} cm^{-2}). This is so because the values measured are the number of active sites, which are shown to be dependent upon not only the surface condition but also the different pH's, ionic strength, and liquid densities. From experimental data,[1] we calculated the values for Ns and found that the values of Ns under experimental conditions similar to ours were of the same order of magnitude, but increased close to one order of magnitude as the electrolyte concentration increased from 1.5 mM to 50 mM.

Electrophoresis of Biological Materials

This year, we also began an investigation of electrophoresis of biological materials in gel-filled silicon microchannels. Gel electrophoresis is widely used for separating biological materials such as DNA fragments. Almost all clinical, molecular, or forensic projects that involve the characterization of DNA are dependent upon the separation and/or purification of DNA fragments by one or more methods. By far, the most common method is based upon electrophoresis. Since the DNA double-helix backbone is negatively charged, fragments of DNA migrate toward the anode when placed in an electric field. If the DNA is caused to migrate through a sieving matrix such as agarose or polyacrylamide, fragment mobility is a function of fragment size, i.e., the smaller fragments migrate faster than larger fragments. In our experiments, we filled microchannels that had been etched in silicon substrates, with agarose gel. **Figure 3** shows the results obtained using a standard 4-mm-thick agarose gel compared to those for a thinner gel supported in a 1-mm-thick etched silicon channel. These results indicate that not only is it possible to separate the fragments in a structure of these dimensions, but more important, both the resolution and speed of the separation are enhanced. We expect that the significantly higher thermal conductivity of silicon, compared to that of standard glass materials used in conventional gel electrophoresis, will enable electric field separations to be done at much higher electric fields to achieve faster separation.

More experimentation is required to verify the expected improvement in separations by going to smaller (i.e., thinner and narrower) gels. We have performed experiments that indicate the definite trend in improved resolution and speed in even narrower gels, down to 0.25 mm. However, difficulty arises in proper and repeatable sample injection in these narrower gels. This warrants further research and development in novel, precise, high-density, small-sample-volume injection.

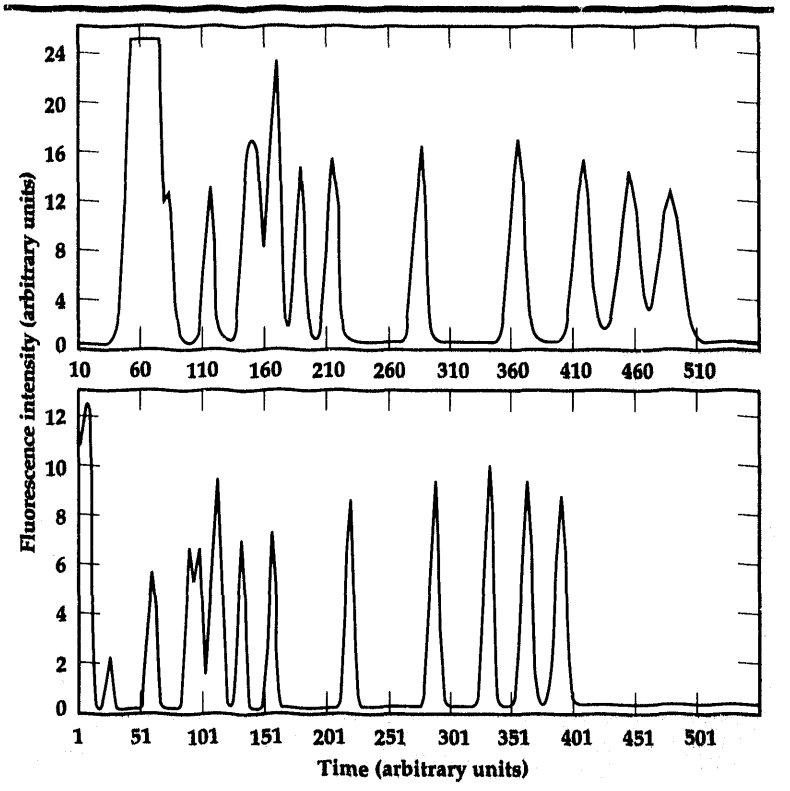


Figure 3. A comparison of DNA fragment separation results obtained for (a) a standard agarose slab gel, 4 mm thick and 4 mm wide and (b) an agarose gel 1 mm thick and 4 mm wide as formed by an etched silicon channel. Note the increase in both peak resolution and speed of separation in the thinner gel supported by the silicon substrate.

1. J.W. Balch, J.C. Davidson, and J.C. Koo, "Capillary Zone Electrophoresis Using Silicon Microchannels," *Laboratory Directed Research and Development FY91*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53689-91, 40 (1991).
2. C.S. Lee, W.C. Blanchard, and C.T. Wu, *Anal. Chem.* **62**, 1550 (1990).
3. C.S. Lee, D. McManigill, C.T. Wu, and B. Patel, *Anal. Chem.* **63**, 1519 (1991).
4. S.M. Sze, "Metal-Insulator-Semiconductor Diodes," *Physics of Semiconductor Devices*, Wiley-Interscience (New York), Chapter 9, 1969.

Emerging Technologies

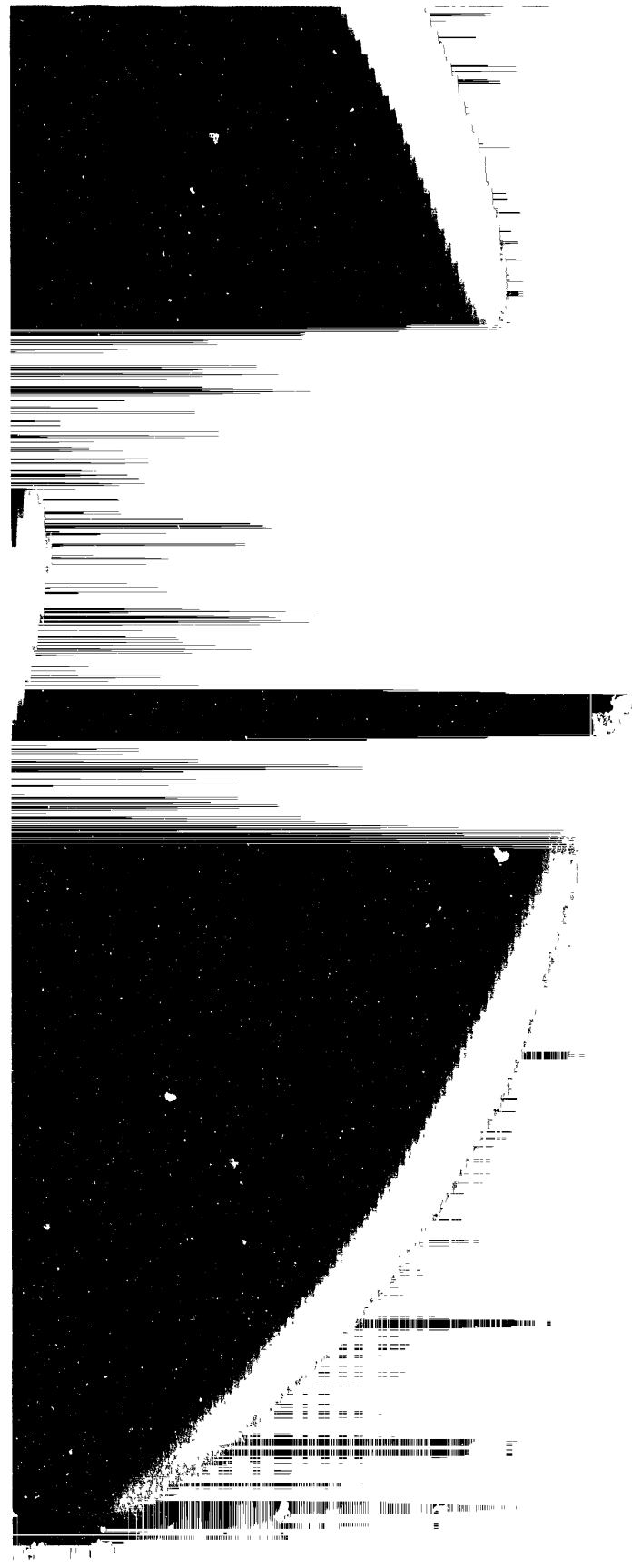
The mission of the Emerging Technologies thrust area at Lawrence Livermore National Laboratory is to help individuals establish technology areas that have national and commercial impact, and are outside the scope of the existing thrust areas.

We continue to encourage innovative ideas that bring quality results to existing programs. We also take as our mission the encouragement of investment in new technology areas that are important to the economic competitiveness of this nation.

In fiscal year 1992, we have focused on nine projects, summarized in this report: (1) Tire, Accident, Handling, and Roadway Safety; (2) EXTRANSYT: An Expert System for Advanced Traffic Management; (3) Odin: A High-Power, Underwater, Acoustic Transmitter for Surveillance Applications; (4) Passive Seismic Reservoir Monitoring: Signal Processing Innovations; (5) Paste Extrudable Explosive Aft Charge for Multi-Stage Munitions; (6) A Continuum Model for Reinforced Concrete at High Pressures and Strain Rates: Interim Report; (7) Benchmarking of the Criticality Evaluation Code COG; (8) Fast Algorithm for Large-Scale Consensus DNA Sequence Assembly; and (9) Using Electrical Heating To Enhance the Extraction of Volatile Organic Compounds from Soil.

Shin-yee Lu
Thrust Area Leader

Section 4



4. Emerging Technologies

Overview
Shin-yee Lu, Thrust Area Leader

Tire, Accident, Handling, and Roadway Safety
Roger W. Logan 4-1

EXTRANSYT: An Expert System for Advanced Traffic Management
Rowland R. Johnson 4-9

Odin: A High Power, Underwater, Acoustic Transmitter for Surveillance Applications
Terry R. Donich, Scott W. McAllister, and Charles S. Landram 4-13

Passive Seismic Reservoir Monitoring: Signal Processing Innovations
David B. Harris, Robert J. Sherwood, Stephen P. Jarpe, and David C. DeMartini 4-17

Paste Extrudable Explosive Aft Charge for Multi-stage Munitions
Douglas R. Faux and Russell W. Rosinsky 4-21

A Continuum Model for Reinforced Concrete at High Pressures and Strain Rates
Kurt H. Sinz 4-23

Benchmarking of the Criticality Evaluation Code COG
William R. Lloyd, John S. Pearson, and H. Peter Alesso 4-27

Fast Algorithm for Large-Scale Consensus DNA Sequence Assembly
Shin-yee Lu, Elbert W. Branscomb, Michael E. Colvin, and Richard S. Judson 4-29

Using Electrical Heating To Enhance the Extraction of Volatile Organic Compounds from Soil
H. Michael Buettner and William D. Daily 4-31

Tire, Accident, Handling, and Roadway Safety

Roger W. Logan
Nuclear Explosives Engineering
Mechanical Engineering

We are developing technology for an integrated package for the analysis of vehicle handling and of vehicle impact into roadside features and other vehicles. The program involves the development and use of rigid-body algorithms and the finite-element codes, DYNA and NIKE. Our goal is a tool for use by highway engineers at the Federal Highway Administration and state Departments of Transportation that allows good quantitative results at the workstation level. Our work has involved integration of handling and deformation codes, development of material and tire models, and comparisons of our results to test data.

Introduction

Our Tire, Accident, Handling, and Roadway Safety (TAHRS) project at Lawrence Livermore National Laboratory (LLNL) provides technical advances to be used in an externally funded program for Vehicle Impact Simulation Technology Advancement (VISTA), to begin on a small scale in FY-93.

The goal of the TAHRS initiative is to develop the technical capability to accurately model vehicle/barrier crash and post-crash behavior (**Fig. 1**). An improved analysis capability will improve highway barrier (and possibly vehicle) designs to minimize risk to occupants, and the hazards due to post-crash vehicle motion. These technical developments will become an integral part of the VISTA program. The goal of VISTA is to integrate the entire state of technology, including DYNA3D,[1] NIKE3D,[2] TAHRS, and other worldwide developments, into a user-friendly highway design tool useful at various levels of expertise.

The current state of the art in barrier design and post-crash dynamics involves a mixture of actual testing using instrumented vehicles, and empirical/numerical modeling using small, personal computer-based codes. These codes have been developed over many years; their empirical aspects have been tuned against crash test data. They are useful tools, but their relative lack of physics leaves them open to technical or legal doubt when extrapolation is involved. As an alternative, LLNL's three-dimensional (3-D) NIKE and DYNA codes could be used separately or coupled to analyze the vehicle/barrier crash interaction. Recently, a Federal Highway Administration (FHWA) contract on this topic concluded that DYNA3D is the code of choice on which to form a VISTA program. The incentive for VISTA is quite strong, since about 40,000 traffic deaths occur each year in this country. As a direct consequence, about 40 billion dollars worth of lawsuits are active at any given time. Often state Departments of Transportation (DOT's) are the targets of these lawsuits. More than half of the fatal accidents typically involve only one vehicle. Thus, the ability to model and analyze barrier crash and post-crash motion with physics-based tools like NIKE, DYNA, and an integrated real-time handling (RTH) capability, could provide a strong supplemental tool for sorting out areas of responsibility.

Progress

The TAHRS technical efforts are organized into four overlapping areas: (1) vehicle handling and interfacing; (2) roadside features and component modeling; (3) vehicle models and integrated analysis; and (4) test data and validation. Highlights of progress in each area are summarized below.

Vehicle Handling and Interfacing

This technical area involves developments in the simulation of vehicle handling, linkage of RTH

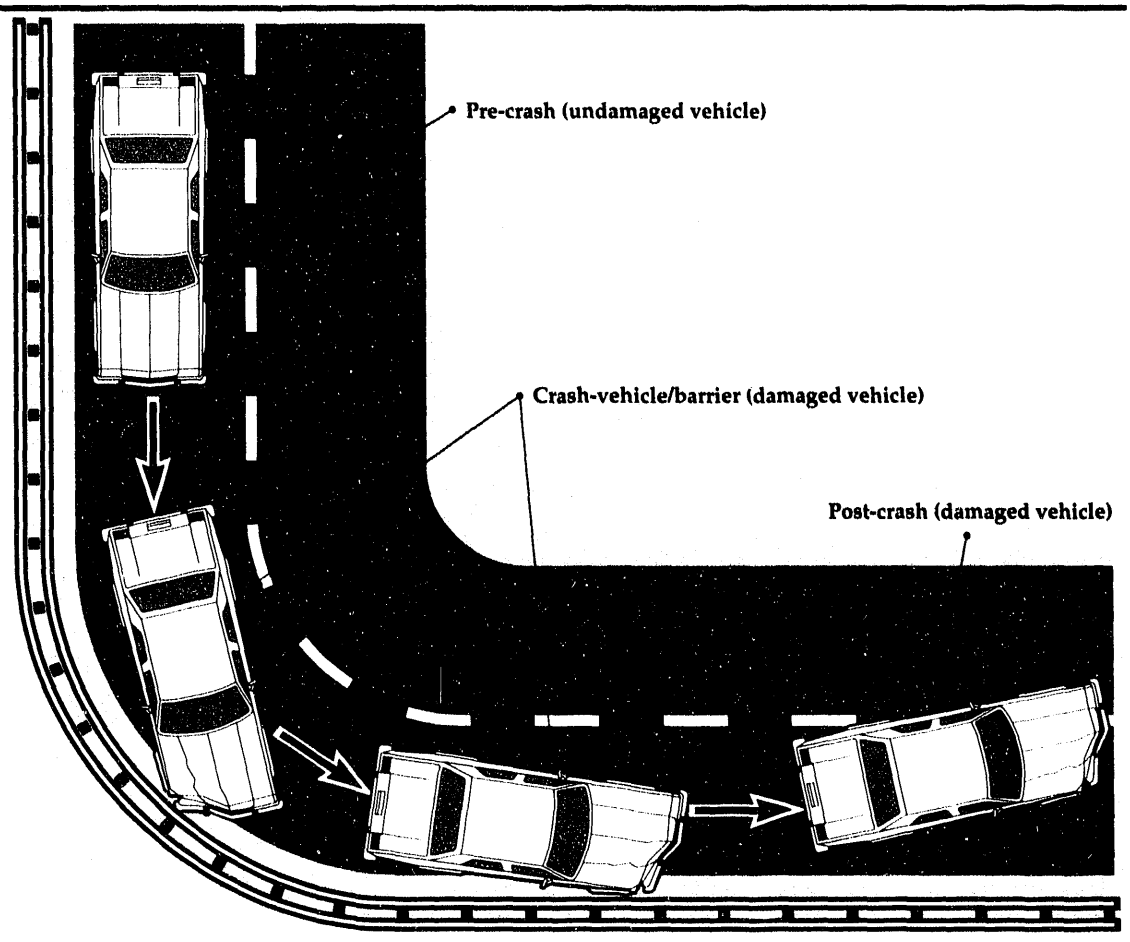


Figure 1. Illustration of the total handling/impact scenario to be addressed by the TAHRS technology and VISTA package.

and finite element mesh (FEM) codes, and development of an informative user interface.

In preparation for the linkage of RTH and FEM codes, a steering force boundary condition is being added to NIKE3D. The lateral forces generated by each tire can be computed by the inclusion of a tire model subroutine in NIKE3D. The vector diagram in **Fig. 2a** illustrates how NIKE3D computes the lateral load on the tire. Simulating the road as a stone wall, NIKE3D first determines the vertical load on the tire. Then, using the user-input driver steering angle, θ, the vehicle orientation direction, **A**, and velocity, **V**, NIKE3D determines the tire slip angle, α. NIKE3D then uses a complex tire model to determine the lateral load, **L**, as a function of these variables. **Figure 2b** shows top views of a car model during two simulations. Identical driver input is used: the wheel is turned first to the left, then to the right. The 25-mph simulation results in a circular path; in the 45-mph case, the car skids into an unstable oversteer condition.

We have also developed a rigid-body vehicle handling code called AUTOSLED to demonstrate the linkage both to NIKE, DYNA, and the user interface. This code and others can be used to simulate the vehicle dynamics before and after the contact, and subsequent deformations of the vehicle. Upon contact detection, data from the model is passed to the finite element code, which simulates the dynamics of the vehicle during the collision. If the accident is such that the vehicle disengages from contact with the barrier, the new, deformed vehicle configuration and the dynamic conditions can be passed back to the rigid-body model for continued simulation. The vehicle configuration is fully 3-D, rigid-body, with 10 degrees of freedom. There is one sprung mass and four independently suspended unsprung masses (wheels). The wheels, which are connected to the sprung mass with a spring and damper, are constrained to move perpendicular to the vehicle. Another set of springs and dampers are used to model the deformation of the wheels, which are free to leave the ground surface. The ground, however, is limited to a flat plane. The vehicle velocity can be controlled by specifying driving forces or a desired velocity. Steering can be accomplished either by specifying a table of steering angles or by designating a path that the steering control will attempt to follow. This results in a 10-degrees-of-freedom model, with each suspension element represented by a spring and damper, as

is each unsprung mass. The tire forces are modeled using the Dugoff tire model and are limited using the friction circle concept. Control of the vehicle during a run is currently accomplished by completing tables in a data file. Vehicle velocity can be controlled by specifying a desired speed or by inputting a table of driving forces vs time. Steering, likewise, can be controlled in two ways. One way is to specify a table of steering angles vs time, and the other is to specify a table of x–y coordinate pairs representing the desired path of the vehicle. An imbedded steering controller will then attempt to follow the path as closely as possible.

To provide the interfaces among the user, AUTOSLED, NIKE, and DYNA, a simulation program has been developed to read an output file from AUTOSLED, and display pertinent information in an interactive graphics environment. A typical session using this simulation program is shown in **Fig. 3**. In the upper left corner of the window, an oval racetrack is shown with a rectangle representing the car. The user has the option of displaying or not displaying the racetrack. In addition, outlines of some of the car's previous positions are shown. The frequency with which these outlines are shown is another option controlled by the user. This view displays both position and yaw of the vehicle. To the right of the track are four gauges. These display suspension forces on the tires as the car follows the path. Next to the gauges are friction circles, which convey information about the normal, longitudinal, and lateral forces experienced by the tires. A constant diameter circle is based on initial forces on the tire when the car is stationary. Below this, steering angles and slip angles are shown. At the lower left corner, a rear view of the car is shown, giving the user information about the roll angle. To the right of this, a speedometer displays vehicle speed in miles per hour. The maximum speed on the speedometer is based on the maximum speed the vehicle reaches during the simulation. A pop-up shell next to the speedometer lets the user create strip charts using any of the 48 variables from the AUTOSLED output file. For example, one could plot lateral forces at tire 1 vs pitch of the vehicle. In the figure shown, the y coordinate of the center of gravity is plotted against time.

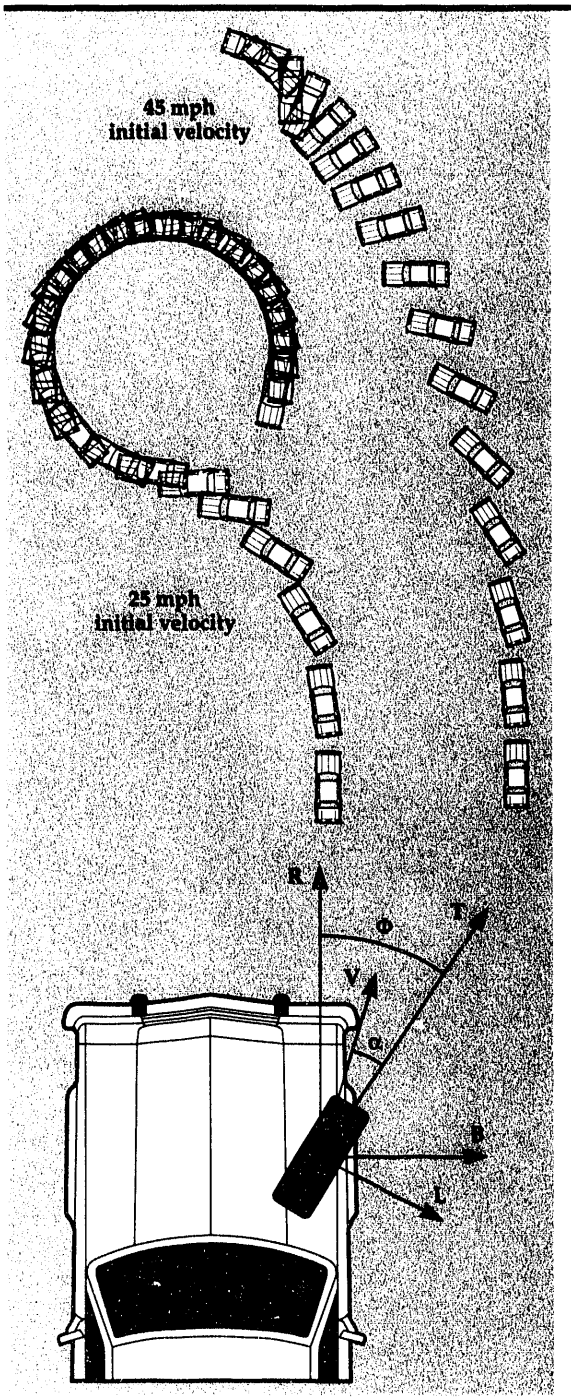


Figure 2. Implementation of steering algorithm and lateral tire force in NIKE3D. Steering is 20 degrees left at t = 0.2 s, then 20 degrees right at t = 2.0 s. Vehicle responds differently as a function of velocity.

Roadside Features and Component Modeling

Before embarking on a 'big picture' analysis of vehicle and roadside barrier under linked handling and impact conditions, it is necessary to consider the FEM deformation analysis of smaller components. This work was begun with an analysis of a rigid bogey developed through a collaboration between the California Department of Transportation (CalTrans) and the University of California Davis. The bogey has a crushable steel box-structure front end resembling a coarse honeycomb, as modeled with DYNA3D in **Fig. 4**. This analysis was run at slow velocities to approximate the static crush test conducted on the actual structure. The mesh was kept coarse in the spirit of workstation level

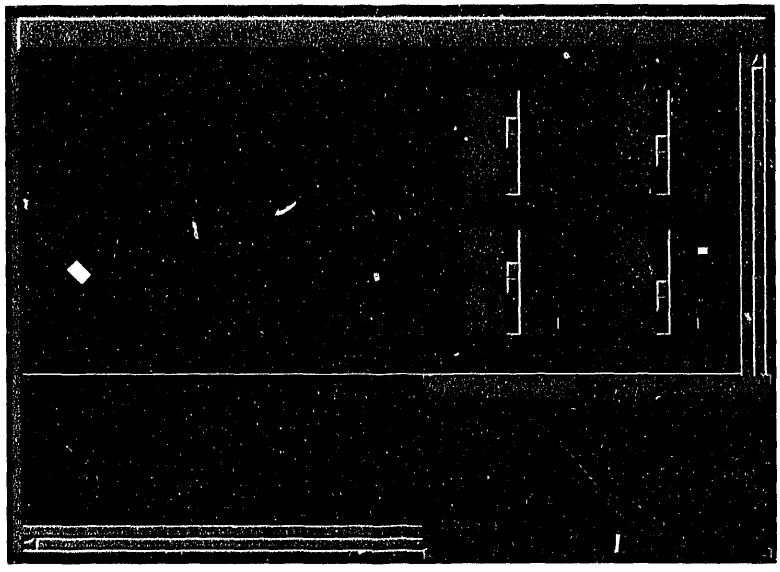


Figure 3. User interface for AUTOSLED handling program. Information includes position, speed, roll angle, original and current friction circles, normal and lateral tire loads, and slip angles.

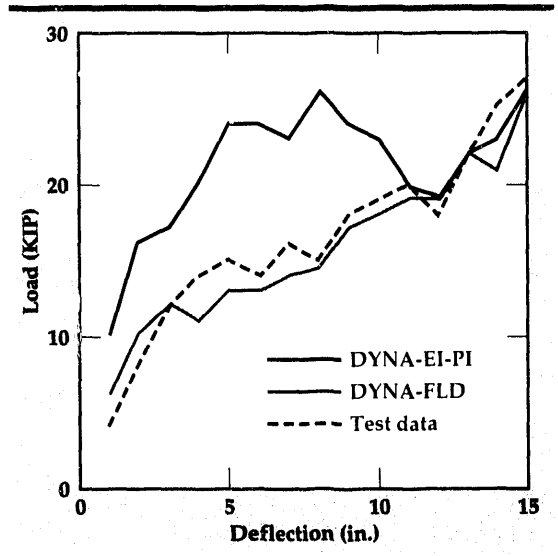
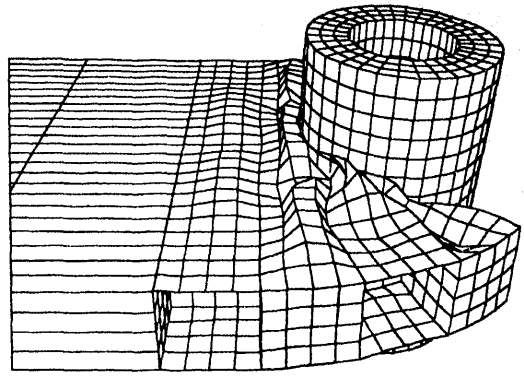


Figure 5. Load deflection for bogey crush into pole. Static test data matches DYNA3D if FLD failure model is used.

Figure 4. Workstation-level DYNA3D mesh of bogey front crush area on impacting rigid pole.



models, and load deflection was compared against the test data, as shown in **Fig. 5**. The first runs with DYNA3D used an elastic perfect-plastic material model. This type of material behavior gives a numerically well-posed problem that is not too dependent on the mesh size. However, the calculated load-deflection (DYNA-El-Pl line) is too stiff during early stages of the crushing process. Use of the augmented Forming Limit Diagram concept[3] with rate-dependent flow and failure allows a match to be achieved (DYNA-FLD) with the test data. The effectiveness of advanced material models under development at LLNL is demonstrated here for isotropic flow and failure. Related studies involve the integration of anisotropic flow and failure theories for analysis of metallic and non-metallic materials, such as deep drawing steels or chopped fiber composites. The type of simulation in **Figs. 4** and **5** is neither predictive nor post-predictive, but is still of value in learning the techniques and meshing needed to match real tests.

Another matching exercise at the component level involved a small car hitting a modified bullnose median barrier. Crash test data for a Honda Civic hitting a modified bullnose median barrier head-on at 60 mph are documented by a report prepared by the Southwest Research Institute for the FHWA.[4] This test was chosen for simulation both because of the availability of test data and because damage to the car was relatively small, allowing a simple car model and a focus on barrier deformation. Since model size and run time were limited, and since many model parameters (especially material properties) had to be estimated, the model is simplified and contains many estimates of relevant parameters. **Figure 6** shows the DYNA model and a sequence of plots as the car plows into the barrier. The car was modeled as rigid. The barrier nose slit was not modeled; rather, the car was 'caught' by constraining the vertical displacements of the front bumper and the bottom edge of the bullnose. The zigzag cross section of the thrie-beam rail was approximated by a rectangular strip with the same moment of inertia and weight-per-unit length as the thrie-beam. Since the deforming rail kinks at the posts where it is fastened, short lengths of thin, 12-gage strip were used near the posts to capture this kinking. The posts were modeled with tie-breaking slidelines, so that they broke off at ground level (as they did in the test), with a region of elastic-plastic material just above ground level

to allow some energy dissipation during post breakage.

With the addition of self-contact and an approximation of the plastic hinge development at the posts, it is possible to match the vehicle trajectory to the test data, as shown in **Fig. 7**. However, predictive or even post-predictive analysis will require further study of the thrie-beam and post components. We need to learn how to make the approximate, coarse-mesh models shown here provide behavior consistent with more detailed models, without having to deal with increased computational requirements.

Vehicle Models and Integrated Analysis

A forerunner of analyses to follow was performed this year in a joint effort involving LLNL, the FHWA, and University of Alaska faculty.[5] We developed a working model of a 1991 domestic sedan.

Our goal was to define the car in sufficient detail to capture its pre-crash, impact, and post-crash behaviors, and yet keep the model simple enough for analyses to be run overnight on a workstation. In the light pole impact example (**Fig. 8**), the car model is rigid material aft of the firewall. Underhood features are modeled as simple rigid bodies. The vehicle model consists of 20 parts, 2406 nodes with six degrees of freedom at each node, 10 beam elements, 1575 plate elements, and 224 solid elements. This is one example of problems we hope to eventually run routinely: large deformations of both vehicle and roadside features, with possible coupling to vehicle handling, in a workstation environment.

The model above was then used in post-predictive mode to demonstrate DYNA3D's crash modeling capabilities. These analytical predictions are compared with crash test results obtained from the National Highway Transportation Safety Administration, where this 1991 domestic sedan was impacted against a rigid wall at a velocity of 57.5 km/h. Although all major structural components of the car were accounted for, the soft crush characteristics of the bumper area were not accounted for in the vehicle model used here and in **Fig. 8**. To compensate for that, a clear distance of 0.5 m between the structural bumper and the rigid wall was allowed. **Figure 9** compares DYNA3D prediction and crash test results of the time/acceleration history of the engine block (upper plot) and rear seat area (lower plot). Given the coarse FEM of the model, the agreement is remarkably good.

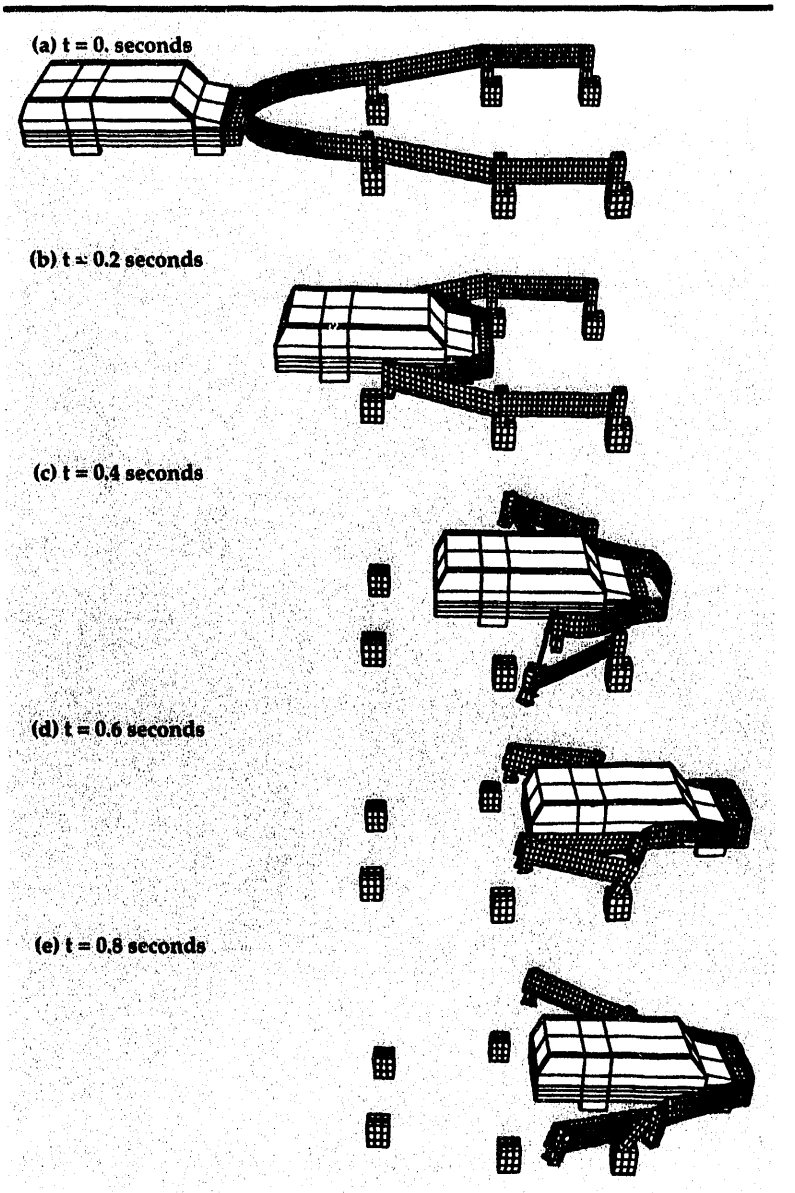


Figure 6. Time sequence of impact of simplified vehicle into modified bullnose barrier. Meshing is again at the workstation level.

Test Data and Validation

A vehicle model of a Ford Fairmont is being constructed. An instrumented test of this vehicle is planned to demonstrate the potential for and effectiveness of an integrated program of analysis, measurements, and vehicle testing. This may lead to further tests or parts of tests at LLNL.

Future Work

This year, we have demonstrated the effectiveness of integrated vehicle/barrier impact analysis at the workstation/FHWA/DOT level, and have identified the needs for additions and refinements

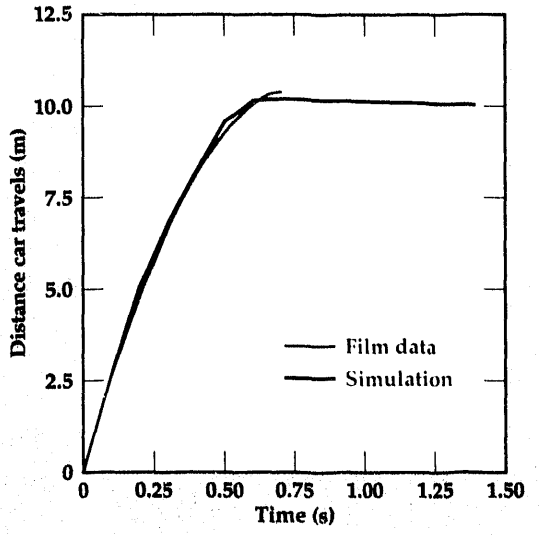


Figure 7. Vehicle temporal position for impact into bullnose of Fig. 6. Although this simulation is not a predictive mode, DYNA3D is able to match the data, even with a very coarse simple model.

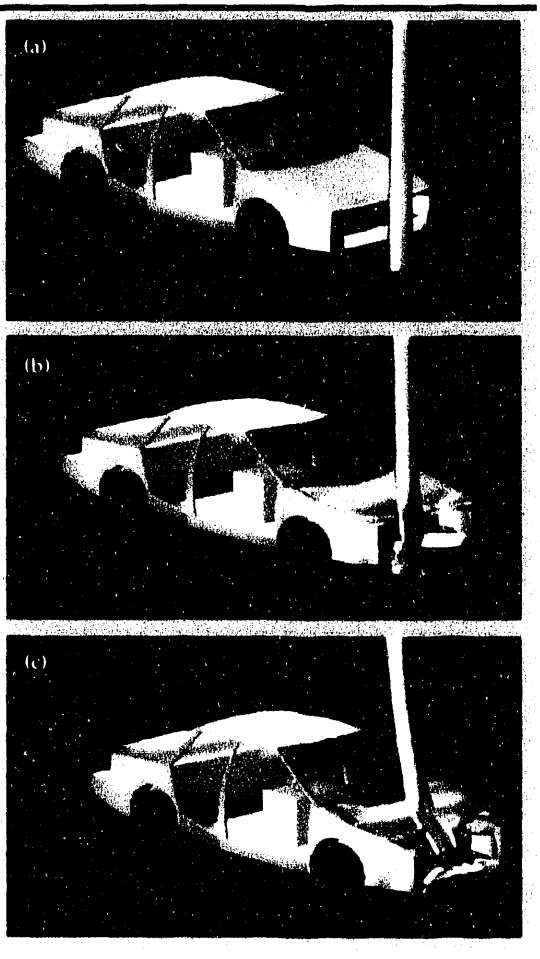


Figure 8. Time sequence of domestic sedan impacting luminaire support. Pole failure is modeled with LLNL's SAND technology.

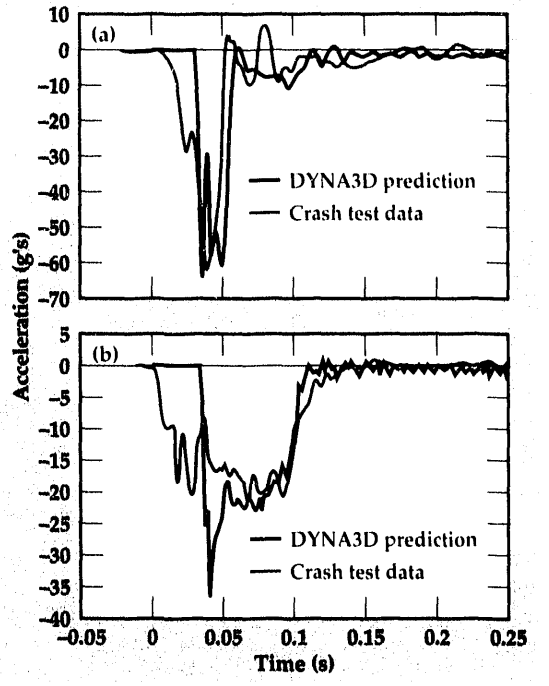


Figure 9. Acceleration history of vehicle model of Fig. 8 for a 30-mph rigid-wall impact. Comparison to NHTSA supplied data is done in post-predictive mode. Agreement is good for only a 2000-node vehicle.

leading to a complete package. Our goals for the future focus on the four technology areas established. We will work toward full linkage of the AUTOSLED RTH code to NIKE and DYNA, and development of compatible tire models for all the codes.

Continued study of both roadside and vehicle structural sections will continue at the component level to ensure that model simplification is efficient yet accurate compared to more refined meshes. A more complete suite of vehicle models and roadside hardware will be developed, making use of material model improvements for flow and crush of aluminum and fiber composite materials, including features such as anisotropy, forming limit, and composite damage. These will be used in future lightweight designs such as Calstart's Neighborhood Electric Vehicle. We will continue the close integration of our analysis package with test data obtained at LLNL and elsewhere.

Acknowledgements

The author wishes to acknowledge the many contributors to the TAHRS/VISTA program this year. Of special note are the contributions of B.N. Maker (**Fig. 2**), D.D. Dirks and M.C. Seibel (**Fig. 3**), and S.J. Wineman (**Figs. 4** and **5**). The author appreciates the close cooperation of Prof. A. Frank

at the University of California Davis on the bogey analysis (**Figs. 6** and **7**), and of Prof. J. Wekezer of the University of Alaska on the domestic sedan model (**Figs. 8** and **9**).

1. R.G. Whirley, *DYNA3D: A Nonlinear, Explicit, Three-Dimensional Finite Element Code for Solid and Structural Mechanics—User Manual*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-MA-107254 (1991).
2. B.N. Maker, R.M. Ferencz, and J.O. Hallquist, *NIKE3D: A Nonlinear, Implicit, Three-Dimensional Finite Element Code for Solid and Structural Mechanics—User Manual*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-MA-105268 (1991).
3. R.W. Logan, *Implementation of a Pressure and Rate Dependent Forming-Limit Diagram Model into NIKE and DYNA*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-ID-105760 (1992).
4. L. Meczkowski, *A Thrie-Beam Bullnose Median Treatment*, U.S. Department of Transportation, Federal Highway Administration Publication Nos. FHWA-RD-88-004 and FHWA-RD-88-005 (1987).
5. J.W. Wekezer, M.S. Oskard, R.W. Logan, and E. Zywicz, "Vehicle Impact Simulation," *J. Transportation Eng.* (in press).

EXTRANSYT: An Expert System for Advanced Traffic Management

Rowland R. Johnson
Engineering Research Division
Electronics Engineering

Coordination of traffic signal systems is carried out at present by a signal timing plan that uses a relatively primitive computer program, TRANSYT. To deal with the difficulties in using TRANSYT, our project is developing an expert system called EXTRANSYT that encodes the knowledge of an expert TRANSYT user. The project is a collaborative effort among (1) Lawrence Livermore National Laboratory, in the lead and providing the computers and computer science expertise; (2) the University of California Berkeley, Institute of Transportation Studies, providing the TRANSYT/traffic engineering expertise; and (3) the City of San Jose, California, Department of Streets and Traffic, providing the testbed for the system.

Introduction

Coordination of traffic signal systems is the primary means by which congestion, pollution, and fuel consumption caused by city traffic is reduced. A coordinated system can be either a single artery or a grid and typically consists of between 10 and 50 intersections. An intersection phase is the time duration for which the traffic lights at the intersection remain fixed. Each intersection has a controller that causes the intersection to cycle through its set of phases.

Coordination is achieved by the use of a signal timing plan wherein the controller at each intersection in the system has the same cycle length. That is, there is a background cycle during which each intersection cycles through each of its phases. The signal timing plan also specifies the offset for the beginning of each phase at each intersection.

One strategy for efficient coordination is achieved by good platoon progression. Platoon progression is the situation whereby a set of closely spaced vehicles (i.e., a platoon) progresses from intersection to intersection, and the platoon is given a green light as it arrives and passes through the intersection. Platoon progression also has the psychological benefit of drivers perceiving that they are moving faster through the system. Another strategy for efficient coordination is achieved by preventing multiple acceleration/de-acceleration cycles. For example, vehicles should be delayed (although the drivers perceive no apparent reason) until a platoon arrives that they can merge with. Usually travel time is the same, and a reduction in pollution and fuel consumption is realized because of the reduction in acceleration/de-acceleration cycles.

Signal timing plan design is usually done by using a computer program called TRANSYT that can (1) simulate the operation of a coordinated system, and (2) find the optimal signal timing plan based on some combination of congestion, pollution, and fuel consumption. Typically, a traffic

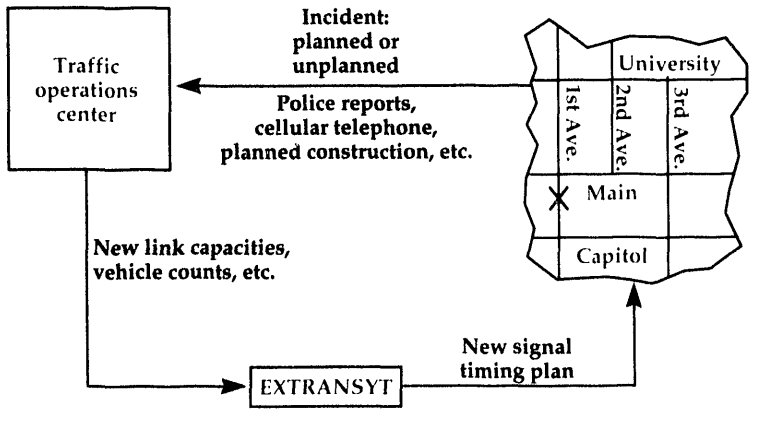


Figure 1. Real-time incident response. A traffic incident has occurred on a city street and has been reported to a central traffic operations facility. Operations personnel determine the impact on vehicular flow capacities and vehicular flow demands. This information is then routed to EXTRANSYT, which quickly determines an appropriate signal timing plan and downloads it to the traffic light controllers.

engineer will provide a description of a set of intersections and streets as well as traffic flow capacities and traffic flow requirements. The TRANSYT model is then calibrated against actual traffic flow conditions, followed by the search for the optimal signal timing plan.

TRANSYT has several limitations that are described below. However, the reality is that it is the only analysis tool of its kind and is likely to remain so for at least five years.

TRANSYT was originally developed in the 1960's when input to computer programs consisted of a punched card deck, and the output device was a lineprinter. In response to the primitive nature of TRANSYT, several peripheral programs have been developed that make it easier to use TRANSYT. However, these efforts do not appear to be adequate since we have found that 30 to 50% of the 'fielded' signal timing plans have errors. Since the original development of TRANSYT, there have been several advances in intersection controller hardware that are not directly modeled by TRANSYT. However, it is possible for an expert user to derive useful results from TRANSYT about coordinated systems that use the newer controllers.

The difficulties in using TRANSYT results in an error-prone, lengthy process to develop a signal timing plan for a coordinated system. A traffic engineer not accustomed to using TRANSYT can require up to four months to develop a signal timing plan for a moderately complicated grid.

Furthermore, the resulting signal timing plan will often have errors that need to be 'tuned out' in the field, resulting in more time required and a sub-optimal signal timing plan.

Progress

To deal with the difficulties in using TRANSYT, our project is developing an expert system called EXTRANSYT that encodes the knowledge of an expert TRANSYT user. The project is a collaborative effort among (1) Lawrence Livermore National Laboratory (LLNL), in the lead and providing the computers and computer science expertise; (2) the University of California Berkeley, Institute of Transportation Studies, providing the TRANSYT/traffic engineering expertise; and (3) the City of San Jose, California, Department of Streets and Traffic, providing the testbed for the system.

The input to TRANSYT specifies a set of intersections, streets connecting them, and the length of each street. It does not specify the location of each intersection. That is, the original Euclidean 2-space specification of the grid is reduced to an undirected graph. A particular undirected graph will have an infinite set of realizations in Euclidean 2-space. Therefore, it is impossible to present the traffic engineer with the two-dimensional (2-D) layout of the intersections and streets that yielded the TRANSYT input. This fact results in many input errors that are never discovered.

In practice, TRANSYT users usually use one of several intersection/street numbering schemes. EXTRANSYT uses heuristics to determine if such a scheme is being used and the Euclidean information derivable from it. Other heuristics about likely intersection/street configurations (e.g., a city street is unlikely to pass over another city street) are also used. As a result, EXTRANSYT is able to determine a likely 2-D layout. In practice, this layout is almost always close enough to the actual intersection/street configuration that the traffic engineer can easily discover input errors.

EXTRANSYT also uses another set of heuristics to discover probable errors not related to the geometry of the grid. For example, the situation where the speed limit in one direction on a street is not the same as the speed limit in the other direction on the same street is flagged as a probable error. As another example, many existing TRANSYT input sets have errors pertaining to the existence, non-existence, and direction of one-way streets. EXTRANSYT has proven to be very effective in finding these types of errors.

Future developments in EXTRANSYT will include heuristics to determine phase sequencing for each intersection. For example, should a particular approach be given the left turn before or after through traffic is allowed to move. Also included will be heuristics to determine which intersections should be in a coordinated system. Closely related to this will be heuristics to determine if an intersection should be fully actuated, semi-actuated, or non-actuated.

As described above, TRANSYT is used to design signal timing plans. Potentially, TRANSYT could also be used to respond to an incident occurring on a city street. As an example, consider an accident that causes the capacity of a street to be reduced and, further, that reduced capacity will exist for one hour. A modified signal timing plan based on the reduced capacity due to the accident would (1) take advantage of reduced demand downstream of the accident, and (2) accommodate extra demand on the alternate routes chosen by drivers upstream of the accident. The problem with this approach is that the modified signal timing plan must be derived quickly. Typically, 15

minutes are required to first download a signal timing plan and then switch to the new plan. In this example, to obtain 30 minutes of improved traffic flow, the modified signal timing plan must be derived in 15 minutes.

Future Work

The current version of EXTRANSYT has been installed at Department of Streets and Traffic in the City of San Jose, California for the purpose of developing a real-time incident response system. The deployed system in San Jose is linked to the development system at LLNL via high-speed modem lines. EXTRANSYT is being used to help analyze existing traffic situations in San Jose. This in turn is used to provide a better understanding of how to implement the heuristics described above. A real-time incident response version of EXTRANSYT will be operational in October 1993.

Odin: A High-Power, Underwater, Acoustic Transmitter for Surveillance Applications

Terry R. Donich and Scott W. McAllister
Defense Sciences Engineering Division

Charles S. Landram
Nuclear Test Engineering
Mechanical Engineering

The Odin project staff has performed an engineering assessment of an underwater acoustic projector using impulse-driven, split-ring-projector technology in an ocean surveillance, anti-submarine-warfare application. An Odin projector system could be engineered to meet the system requirements for output power and acoustic beam control; however, the final projector size raises serious issues about its compatibility with existing deployment platforms. This problem and the fact that the submarine threat has changed have led the project team to defer further work on this application and to focus on the air-deployable, impulse-driven projector being funded by the Navy.

Introduction

Split ring projectors (SRP) are acoustic transmitters for underwater use in active sonar systems to detect submarines. In FY-90, Lawrence Livermore National Laboratory (LLNL) developed the idea of using the combustion of chemical fuels to drive a SRP element. When a chemical fuel combusts inside the cylinder, the resulting inward pressure pulse drives the shell outward, loading strain energy into the split ring shell. The split ring shell 'rings' down, converting the strain energy into acoustic energy, as illustrated in **Fig. 1**. 'Impulse-driven split ring projector' is the phrase used to describe this system.

The chemical fuel-driven SRP overcomes the acoustic power limitation encountered when piezoelectric ceramics drive the split ring shell. At LLNL, our capabilities in numerical modeling of combustion and detonation and in structural modeling give us a unique capability to assess the feasibility of impulse-driven SRP's.[1]

In FY-92, we were funded to assess the feasibility of the impulse-driven SRP concept, scaled to a ship-towed surveillance system, as illustrated in **Fig. 2**. The hypothesis put forth in the reviews was that impulse-driven SRP technology could create the acoustic power required by the surveillance community in a reasonably sized, hydrodynamic package that could be towed with greater ease than the existing projector arrays. The basis for the hypothesis was twofold: (1) package size would be reduced, since the direct conversion of chemical energy into strain energy was more efficient than converting the chemical energy into electricity, conditioning that electricity, and creating strain energy with magnetostrictive or piezoelectric materials; (2) the SRP, being long and slender, provided a more hydrodynamic shape than other piezoelectric or magnetostrictive projectors.

Progress

This project comprised four tasks: (1) the enhancement of our fluid-loaded SRP codes, SOFA[2] for the frequency domain and SOTA[3] for the time domain; (2) a parameter study of surveillance-scale

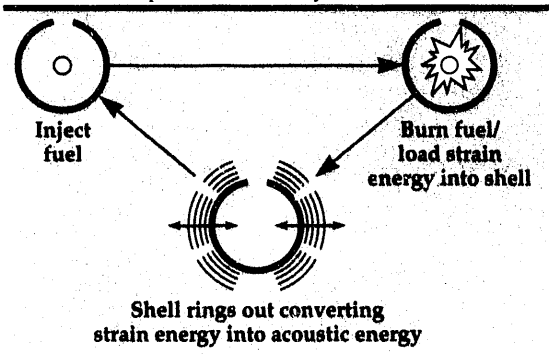


Figure 1. The cycle for a chemical fuel-driven SRP element.

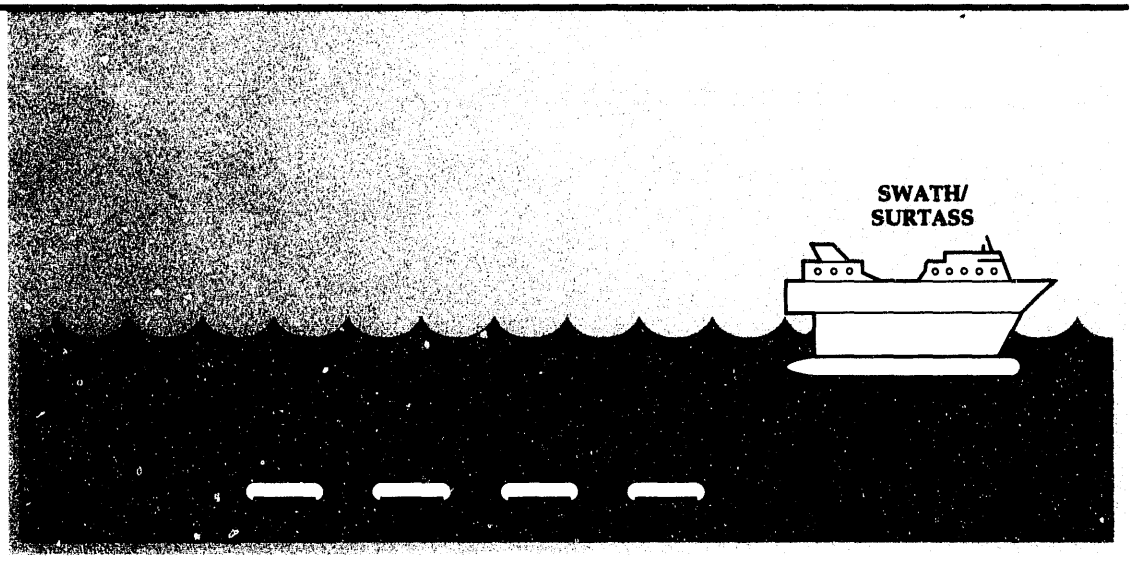


Figure 2. Configuration of Odin projector and SWATH/SURTASS.

SRP's; (3) an assessment of the feasibility of the projector parameters in an actual application; and (4) marketing activities for this project and related projects.

Our parameter study yielded the following results:

(1) The optimal radius at 20 Hz is 1 m; the corresponding shell thickness for this frequency is 24.4 cm. The shell material is steel.

(2) The chemical-to-acoustic energy conversion efficiency is low (≈ 1.5% for L = 20 m).

(3) Acoustic powers of ≈ 363 kW or 226.4 dB are attainable with an input of 310 MJ of chemical energy. The peak stress is within the elastic range for high quality steels.

(4) The projector acoustic output has the temporal characteristic of $P_0 e^{-\alpha t} \sin\omega t$. Our experimentally validated analytical model predicts $\alpha = 0.095\ s^{-1}$ for this shell radius and projector frequency.

From our engineering assessment,[4] we determined that in an array of these large SRP's, the 2σ individual projector timing specification must be 0.005 s. This specification is realizable from an engineering point of view. It will take approximately 60 s to recycle a projector after it fires. This time includes the time to purge the exhaust and reload the fuel/oxidizer for the next shot.

A five-element projector array with a per element power of 226.4 dB will enable detection ranges in excess of 140 nautical miles.[5]

Five 2-m-dia-×-20-m-long projectors are massive enough to question the ability of existing platforms to recover such an array.

Our marketing activities are influenced by the status of the Navy surveillance community. Redefinitions of their mission include the transition of the surveillance environment from deep water to shallow water; and the transition of surveillance platforms from vulnerable ships to air-deployable disposable systems.

The Navy is not able to fund this project in FY-93.

LLNL staff were asked to witness a Navy test series off San Clemente Island to assess the performance of the current generation of surveillance projector technology.[6] This information has been used to submit additional white papers for reimbursable work for the Navy.

The Office of Naval Research (ONR) has expressed interest in the techniques used in SOFA and SOTA, our two fluid-coupled split-ring modeling codes. A white paper has been prepared, on the basis of the ONR interest, to assess some of the fundamental questions that arise when modeling the acoustic radiation from a complex stiffened structure.[7]

In summary, although the Odin projector concept is feasible from an engineering point of view, a large funded project is precluded because of the major redefinition of the missions of surveillance communities. However, our marketing has uncovered other potential funding sources for related projects.[8]

Acknowledgements

We thank the staff of LLNL's Military Applications and Advanced Conventional Weapon Systems for their support. Our thanks also go to Tom Reitter for his numerical computations, using CALE to determine the pressure time histories generated internal to an underwater cylinder by an explosive charge; to Ensign Hal Perdew for his work in the

systems studies of SRP arrays; to Barry Bowman for marketing assistance; to S. Christian Simonson, Robert Tipton, and Rich Couch for computations; to Kent Lewis for mode conversion; and to Clark Sauers who formulated the equation of state for non-explosive energetic material.

1. C.S. Landram, "SOFA," 1991–1992.
2. C.S. Landram, "SOTA," 1991–1992.
3. T.A. Reitter, "CALE Calculations of Small-Charge Explosions in Underwater Pipes," TF92-34, April 2, 1992.
4. H.G. Perdew, "Timing Requirements on Split Ring Projector Arrays," August 1992
5. H.G. Perdew, "Performance Assessment of Split Ring Projector Arrays in Shallow Water," September 1992.
6. S.W. McAllister, "Lessons Learned From NCCOSC, NRAD Tests of 9/16/92 and 9/17/92," September 18, 1992.
7. C.S. Landram, "ONR Proposal on Mode Conversion," TF-92-76, September 21, 1992.
8. S.W. McAllister, "Electrically Initiated - Frequency Dispersive Sources, A Requirements Document," October 1, 1992.

Passive Seismic Reservoir Monitoring: Signal Processing Innovations

David B. Harris and Robert J. Sherwood
Engineering Research Division
Electronics Engineering

Stephen P. Jarpe
Earth Sciences Department

David C. DeMartini
Shell Development Company
Houston, Texas

We have extended our matched field processing capability in mapping acoustic emissions associated with hydraulic fracturing. In our new approach, we generate elastic matching fields for a range of source types, and match the best linear combination of these fields, against the observed data. We have begun work with Shell Development Company, applying our methods to data from their monitoring wells.

Introduction

Hydraulic fracturing is a widely used well completion technique for enhancing the recovery of gas and oil in low-permeability formations. Hydraulic fracturing consists of pumping fluids into a well under high pressure (1000 to 5000 psi) to wedge open and extend a fracture into the producing formation. The fracture acts as a conduit for gas and oil to flow back to the well, significantly increasing communication with larger volumes of the producing formation. While typical treatment costs exceed $100,000 per well, hydraulic fracturing may double or triple production. Such returns justify extensive use of the technique. In the interval from 1949 to 1981, more than 800,000 treatments were completed.[1] In tight gas sands and diatomite oil reservoirs,[2] virtually all new wells are hydraulically fractured.

Field engineers need diagnostics for the height, length, and orientation of fractures to design the proper spacing of wells in the field and to design individual fracture treatments. The diagnostics must be inexpensive (10% of treatment cost), fast, and reliable. Diagnostics that are available in a few hours can be used to plan successive stages of a multistage fracturing operation in a single well; diagnostics that take minutes could be used in real-time controls of pumping rates and fluid composition.

The best diagnostics that fully map a fracture use transient microseismic signals emitted from micro-fracture events along the fracture surface.[3,4] These signals are detected by sensors placed in adjacent monitoring wells or in the treatment well. The arrival times of the signals are measured (usually manually), then used to triangulate the sites of emission. The 'cloud' of locations for several hundred discrete emissions delineates the fracture. This method is slow due to the need for manual picking of arrival times, and has potentially limited application when an insufficient number of high signal-to-noise ratio transient signals are detectable.

We have adapted matched-field processing methods to the problem of imaging fractures, using continuous microseismic emissions.

Progress

In FY-92, we extended our earlier results,[5] which used an acoustic model for propagation, to the

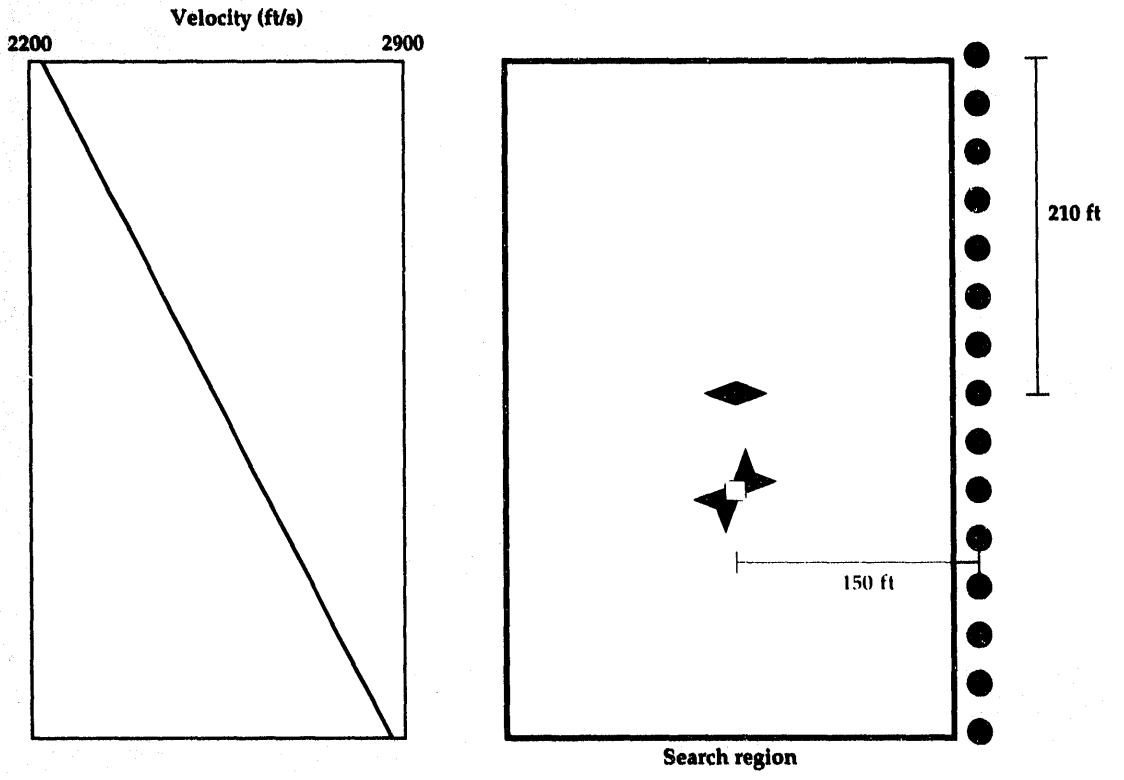


Figure 1. Source and array configuration for matched field processing test. The shear wave velocity as a function of depth is displayed on the left. On the right are two sources adjacent to a measurement array of vertical geophones. The top source is a horizontal dipole (intended to simulate a pulsating crack), and the bottom source is a double couple (intended to simulate a microearthquake).

case of full elastic propagation. We have developed two elastic field simulators; one to produce test data, and another that is a highly efficient narrowband code to generate matching fields for the array of sensors. The latter code involves innovations in paraxial wave field extrapolation,[6] that have potential application to oil prospecting and ocean acoustic modeling. We have also developed a matched field processor with the ability to match a wide range of source types. Hydraulic fracture microseismic sources may come in a variety of forms, such as an opening crack caused by pumping, or microearthquakes caused by slip between adjacent blocks in the pre-stressed medium.

We simulated these two types of sources for the source and sensor configuration, shown in **Fig. 1**, for a medium intended to approximate the conditions in the Shell Belridge oil field.[7] **Figure 1** shows a vertical array of vertical-axis geophones in a monitor well, and two simulated sources 150 feet away. The array has 15 geophones spaced at 30-ft intervals, which is similar to the Shell Belridge sensor configuration. The velocity structure is non-uniform, consisting of a gradient with shear wave speeds ranging from 2200 ft/s to 2900 ft/s over the aperture of the array.

The fields radiated by the two source types are markedly different and require a new processing strategy for situations where the source type is unknown. **Figure 2** shows the fields generated by the opening crack and the slip type sources. The two sources radiate energy away from the source location with very different patterns as a function of direction. The signals received by the array are correspondingly unique. This presents a problem if the matching field is not chosen appropriately, as shown in **Fig. 3**. The first two reconstructions of the two-source test case are reconstructed with theoretical fields corresponding to a single source type. In both cases, one of the sources is missing in the reconstruction.

Our response to this problem was to develop a modification of the matched field processing approach, which we call multiple-field matching (MMFP). In this approach, we generate matching fields for a range of possible source types, and match the best linear combination of these fields point for point in the search regions, against the observed data. The result of MMFP for our test case, shown in **Fig. 3**, is an image containing both source types.

Future Work

We have entered into an agreement with Shell Development Company to apply matched field

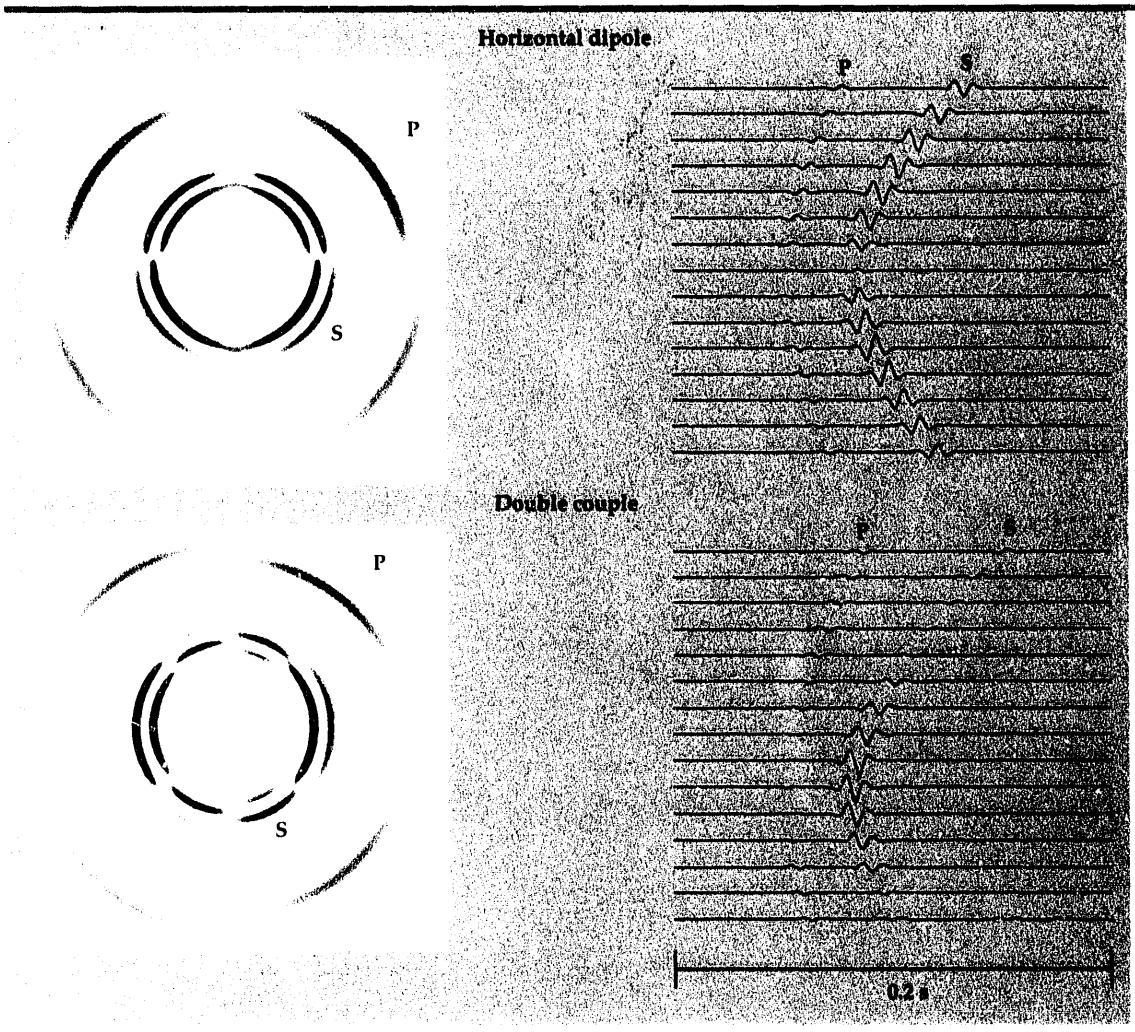


Figure 2. The fields generated by the two representative source types. On the left are the fields displayed at a single instant of time as a function of depth and horizontal distance. Note that both compressional (P) and shear (S) waves propagate. The two source types produce fields with different patterns of sign reversals and nulls in different directions. On the right are the time-dependent waveforms recorded by the geophones in the array.

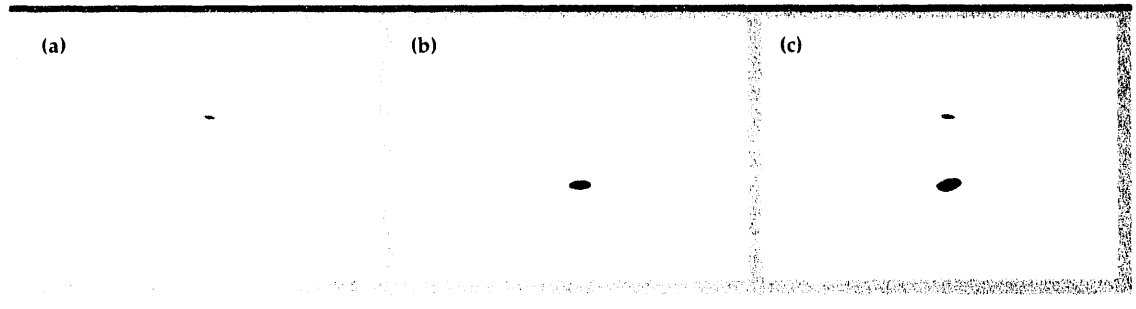


Figure 3. Three reconstructions of the source distribution made from the sum of the two sources shown in Fig. 2. The reconstructions are for the search region outlined in Fig. 1. Reconstruction (b) uses a double couple field, and misses the horizontal dipole source. Reconstruction (c) is the MMFP reconstruction that uses a best linear combination of both fields point for point in the search region, and picks up both sources.

processing methods to their Belridge hydraulic fracture data set.[7] The Shell data set is the best available data for testing hydraulic fracture imaging diagnostics. It includes two multi-stage fracture operations recorded by three monitor wells. The vertical geophone sensors were grouted into the wells, largely suppressing the tube waves that confound hydrophone recordings in fluid-filled monitor wells. Preliminary analysis of the data has shown us the necessity of using the multiple field extension of matched field processing with real data. We anticipate that our analysis will provide a definitive test of the value of matched field processing in the coming year.

1. B. Waters, *J. Pet. Tech*, 1416 (August 1981).
2. "Frac Attack," *ChevronWorld* **Spring/Summer**, 24 (1991).
3. J. Fix, R. Adair, T. Fisher, K. Mahrer, C. Mulcahy, B. Myers, J. Swanson, and J. Woerpel, *Development of Microseismic Methods To Determine Hydraulic Fracture Dimensions*, Teledyne Geotech, Garland, Texas, Technical Report No. 89-0116 (1989).
4. B. Thorne and H. Morris, *SPE Form. Eval.*, 711 (December 1988).
5. D. Harris, R. Sherwood, S. Jarpe, and P. Harben, *Mapping Acoustic Emissions from Hydraulic Fracture Treatments Using Coherent Array Processing: Concept*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-ID-108262 (1991).
6. D.B. Harris, *Wide-Angle Fourier Wavefield Extrapolators for Laterally Heterogeneous Media*, in preparation.
7. H. Vinegar, P. Wills, D. DeMartini, J. Shiyapobersky, W. Deeg, R. Adair, J. Woerpel, J. Fix, and G. Sorrels, *J. Pet. Tech.*, **44** (1) 28 (January 1992).

Paste Extrudable Explosive Aft Charge for Multi-Stage Munitions

Douglas R. Faux and
Russell W. Rosinsky
Nuclear Explosives Engineering
Mechanical Engineering

Our development project for a paste extrudable explosive (PEX) aft charge is a multi-year effort with the goal of demonstrating the technology in a multi-stage munition. In FY-92, we studied PEX borehole fill characteristics and PEX initiation schemes.

Introduction

Multi-stage conventional munitions typically have a two-stage warhead: a forward-shaped charge that produces a borehole in the target, and an aft charge that enters the borehole and then detonates, destroying the target. The aft charge is usually a steel-encased explosive that either enters the borehole by its own kinetic energy or is 'driven' into the borehole by a rocket or velocity augmenter.

To increase the versatility and reduce the weight of a portable, multi-stage munition, a paste extrudable explosive (PEX) aft charge that injects PEX into the borehole formed by the forward charge replaces the steel-encased aft charge. The PEX aft charge can be used with a smaller borehole and provides greater coupling of the explosive with the target.

Progress

The PEX Aft Charge project is a multi-year effort with the goal of demonstrating the technology in a multi-stage munition.

The PEX aft charge is a proposed, pre-planned product improvement for the penetration augmented munition (PAM) currently being developed for U.S. Special Operations Forces and future multi-stage munitions. **Figure 1** illustrates a conceptual drawing of a standoff destruct munition (SODM, or 'flying PAM') using a PEX aft charge.

Our FY-92 development work on the PEX aft charge involved two areas: PEX borehole fill characteristics and PEX initiation schemes. Computer modeling of the PEX extrusion through a nozzle and of the borehole fill process has been completed and will be validated by forthcoming tests.

The two-dimensional hydrodynamic code CALE has been used to model the PEX borehole fill process (**Fig. 2**). The complete simulation required the coupling of a DYNA2D analysis of the PEX extrusion through a nozzle, to the CALE analysis of a borehole fill.

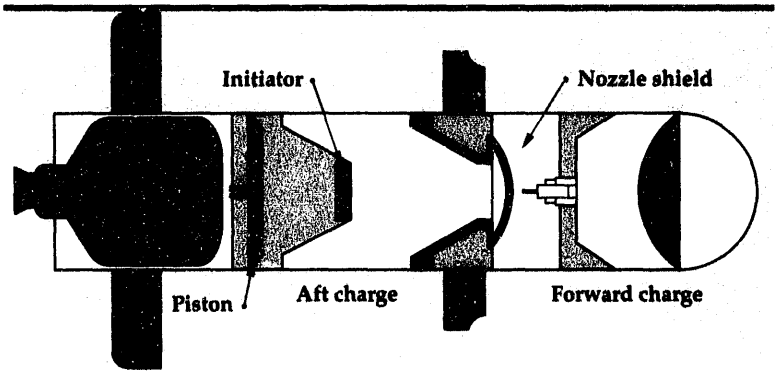


Figure 1. Conceptual sketch of a SODM with a PEX aft charge.

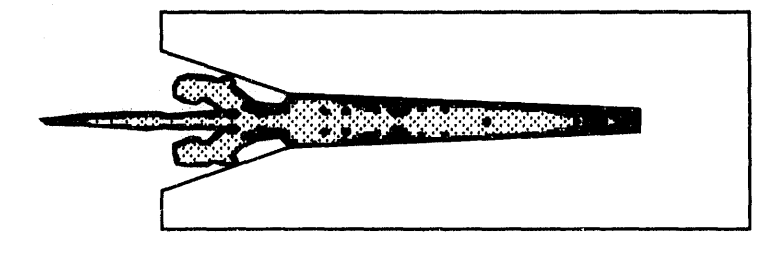


Figure 2. CALE Simulation of a PEX borehole fill.

Future Work

Four tests are scheduled: two tests will evaluate PEX flow characteristics during borehole fill and potential sympathetic detonation of the PEX; two tests will investigate PEX initiation schemes. The first initiation scheme involves the use of a detonation chord and carrier vane to pull the detonation chord into the borehole with the PEX; the second initiation scheme involves the use of three chemically delayed detonators that will flow with the PEX into the borehole.

A Continuum Model for Reinforced Concrete at High Pressures and Strain Rates

Kurt H. Sinz
Earth Sciences Department

We are studying the behavior of concrete at high pressures (200 kb) and strain rates (10^4/s) and report on a computer model used for this purpose. Applications include a predictive capability for the damage done to concrete when it is subjected to attack by demolition munitions or penetrators.

Introduction

Concrete is one of the most common building materials in the world. In 1992, the U.S. alone used an estimated 265 million cubic yards.[1] Consequently, the need arises for occasional demolition or perhaps even destruction, such as in the event of armed conflict. Recent advances in small-scale munitions make it possible to consider the effects of a successful point attack against concrete even when it contains heavy steel reinforcement ('rebar'), as in bridge piers or bunkers. The damage mechanism is very different from that in seismic events or in ground shocks induced by nuclear explosions where damage results from large-scale flexure and fracture or rebar pull-out.

Payoff of Predicting Concrete Damage

To date, no computer model exists that can predict the damage envelope in concrete resulting from an interaction with a demolition munition or a penetrator. A calculational model that predicts the damage done to concrete subjected to point attack is highly desirable for a number of reasons. With the help of such a model, experiments could be more effectively designed to yield specific information, thus increasing the 'leverage' of experiments that are performed. Aspects of experiments and tolerances of design that are not 'laboratory perfect' could be evaluated by computer. This might include non-ideal standoffs, oblique angles of incidence, and structural peculiarities. This capability in turn would permit the design of a new munition to be more effectively optimized, and the survivability of penetrators could be calculated. Development time of new munitions would be shortened and would require fewer experiments. The resulting cost savings and product improvements are obvious.

Problem

The problem we pose is to develop a continuum model for concrete that explains the results of experiments performed for the penetration augmented munition (PAM) program. A rebar-cutting charge for the PAM has been designed to specifically attack concrete and to cut near-surface rebar up to No. 11 in size.[2] The minimum diameter of this size rebar (when ignoring any ribbing) is between 3.3 and 3.4 cm. Data from rebar-cutter tests exist that do not seem to be obscured by complicated hydrodynamic motion. Inhomogeneities are small compared to the effects of scale, and rebar spacing is of the order of the damage scale. The configuration is therefore amenable to analysis by a two-dimensional (2-D) Lagrangian continuum code such as DYNA,[3] and rebar-cutter experiments are especially pertinent to this effort from the standpoint of providing data as well as filling a need.

Concrete Properties

To construct a continuum model, we need to know available properties of concrete. A great deal of attention has been devoted to the understanding of

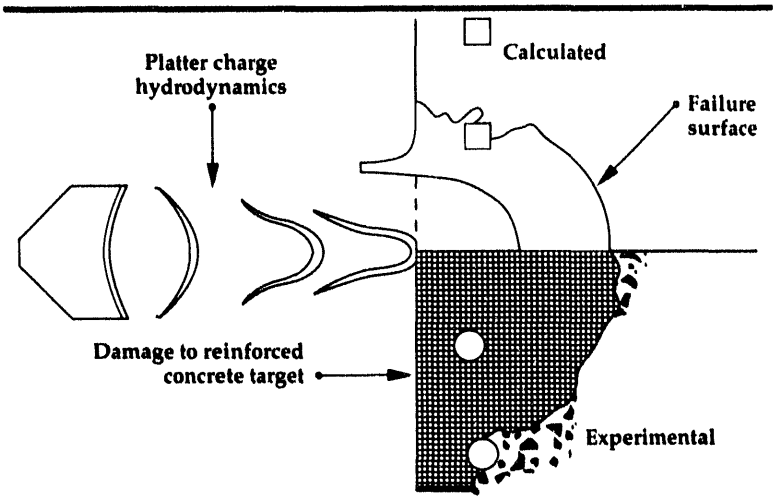


Figure 1. Overlay of calculation and actual hole profile. The material boundaries show the concrete surface after crush and the limit to which material is failed. All the original material is still in the problem and no excavation or 'throw-out' has occurred. The rebar is approximated as rings of square cross section with the central rebar omitted.

concrete in the low-pressure, quasi-static regime. However, surprisingly little quantitative information is available in the high-pressure and high-strain-rate regime. The most important data available to cover the high-pressure and high-strain-rate regime are twenty years old.[4,5] These data have been rehashed by several authors, but until recently little has been done to validate or to extend the old data. Exceptions include Hugoniot data[6] and static strength data.[7] These data cover only a small subdomain of the old data, but it is reassuring that the new results generally support the old data. An inspection of the data of Chinn and Zimmerman and Neely, *et al.* indicates that 'strength' of concrete does not appreciably affect yield strengths when the mean normal stress (which in part results from confinement) reaches about 1 kb. Tang has investigated dynamic behavior of concrete at strain rates up to 280/s.[8] Tang's result of most interest to us is a measured value of Poisson's ratio, which he found to be independent of strain rate. It should be noted that our problem develops strain rates up to 10^4/s.

Another important property of concrete is porosity. It has been measured and estimated to range from 8 to 20% by volume.[9]

Progress

We have found that the main premises of this work have been validated. A continuum model for concrete to describe its behavior and damage when attacked by a PAM rebar-cutter charge is applicable. Pore crush and shear failure appear to be sufficient to explain the depth of the resulting borehole. The effects of tensile failure are discussed below.

The presence of rebar does not seem to be important to the damage. The mechanism for forming the hole is not so much that of excavating material but rather failing material without a lot of hydrodynamic motion. The majority of the damage by volume results from shear failure. The strain rates that develop in the concrete when attacked by demolition munition reach about 10^4/s with pressures of about 200 kb. The calculations were done in KDYNA,[10] which is a minor variant of the DYNA code[3] in that it includes a material model from TENSOR.[11] The problem was much simplified by replacing the actual impacting metal from the rebar-cutter charge as a time- and space-dependent pressure profile on the surface of the concrete. The pressure history was obtained using CALE when the charge was acting against aluminum.[12] Comparisons were made using aluminum in DYNA to assure that resulting pressure distributions in the aluminum were properly calculated in comparison to CALE. The charge was in fact well represented by the pressure-time history on the surface. The explicit absence of the charge greatly simplified our calculations and made them more suitable to treatment by a Lagrangian code.

Using our model for concrete (discussed below), we calculated the effect of a rebar-cutting charge against plain, non-reinforced concrete. **Figure 1** does include an annular approximation to the rebar as an illustration of the results. All the original material still resides in the problem. The surface 'dent' results from pore crush and displacing some material into the crater lip. Another boundary is the limit of material failure that results from shear failure. This is considered to be the damage envelope, and failed material is assumed to either fall out of the hole or to vacate the region by other mechanisms such as interaction with explosive gases. A comparison is shown with an experimentally obtained hole. It is apparent that the hole depth as defined in the calculation is in excellent agreement with the experiment.

We took some pains to try to explain why the calculated hole width is much less than the experiments indicate (**Fig. 1**). Accuracy in the definition of the applied pressure profile in time and space was eliminated as a possible cause. Currently, we identify three possible causes. The first is related to the zoning that the Lagrangian code DYNA generates as the crater begins to form. Zones that are long in the direction parallel to the crater wall cannot accurately transmit a signal that emanates

from the crater bottom and travels along the crater wall. Furthermore, it appears this signal travels though material that is already dynamically failed and therefore is not reliably characterized. There is some indication that if the signal were well resolved by the zoning, that tensile failure would matter somewhat in this region. We should note, however, that the calculated bore hole is large enough to admit follow-through charges of current design so that the resolution of these questions, while highly desirable, may possibly not be central to our first objective. It is most interesting that tensile stresses were not found to matter anywhere in the problem except possibly right on the crater wall and then parallel to it. Reflections from the experimental sample block's boundary did not contribute appreciable reflections and tensile waves for two apparent reasons. The first reason is that the porosity in the concrete is a good shock attenuator and only low-level signals reach the boundary. The second reason is that a release from the target surface follows the main shock and contributes to its decay from behind.

Since we are limited by the constraint that DYNA is a 2-D code, two different attempts were made to estimate the importance of the rebar. In the first attempt, the rebar directly under the impact area was ignored, and the remaining rebars were represented as rings with radii of the rebar spacing (**Fig. 1**). The rebar was hardly displaced and had virtually no effect on the failure envelope. This result stems from the porosity-induced attenuation in the concrete and is consistent with experimental observations where the rebar is linear as opposed to circular. In the second estimate, the rebar was represented as a solid slab of steel, three centimeters thick, which was backed by concrete and covered with 6 cm of concrete (**Fig. 2**). Plastic strains of 10% in the steel were observed to a radius of about 3 cm. This result makes it plausible that a three-dimensional calculation would give some amount of gap in the rebar using our current model and its parameters. This, of course, is the objective of the rebar cutter.

Another interesting phenomenon was observed in this latter calculation. The concrete cover of the rebar absorbed sufficient momentum in the radial direction to continue to 'peel' off the rebar (the solid slab of steel). The occurrence of this phenomenon distinguished this calculation from those without any steel. We assumed zero bonding strength between the steel and the concrete. It seems plausible that this general phenomenology of the concrete cover peeling at a plane of weakened bonding may explain why all the concrete cover seems to blow or crack off when a PAM is tested against a sample block of reinforced concrete.

Model

We now give a brief review of the model we use. The model combines the effects of pore crush, shear failure, and tensile failure. The aspect of pore crush is represented by a hysteresis model with a tensorial model for shear failure. This basic model has been used in TENSOR[4] for a number of years to calculate the behavior of earth materials when subjected to high shock pressures. The model, with some improvements,[13] has been carried forward to KDYNA.[10] Numerical data for the model are obtained from Gregson's Hugoniot, which also accounts for the pore crush in loading. The pressure at which total crush of the porosity occurs is assumed to be 100 kb. There is a perhaps fortuitous match between Gregson's Hugoniot and the Hugoniot for fused silica (SiO_2 or Dynasil) at the high-pressure end. We follow this suggestion and assume the Hugoniot for fused quartz to be applicable to dynamically fully crushed (pulverized) concrete. This assumption seems reasonable, since concrete is mostly quartz.

The pore crush model works by letting a piece of the material load along the Hugoniot. Upon release, the unloading does not simply reverse the loading path; instead, hysteresis is approximated by interpolating a release path from the Hugoniot between the elastic portion of the loading curve and the Hugoniot for fully crushed material. The

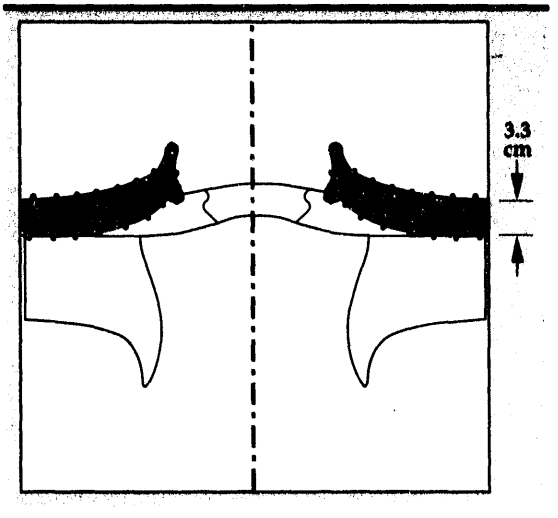


Figure 2. Overlay of rebar cut by the rebar-cutter charge and a calculation. The scales are only approximately the same. In the calculation, the rebar is approximated as a solid slab of steel. The lines of 10% plastic strain in the steel are shown as a suggestion where the limit of calculated failure in the steel might be.

hysteresis is of course a representation of the pore crush that is fully present and accounted for in the Hugoniot. At any point in this space, the pertinent bulk modulus is inferred. The resulting bulk modulus is thus a function of pressure that is related to strain rates in shock regions. Poisson's ratio is obtained as an extrapolation of Tang's work[8] so that the shear modulus is also defined.

To complete the model, we need values for the shear strength of concrete. The most extensive set of data at this time still appears to be that of Chinn and Zimmerman, who give values of shear strengths of cylindrical concrete samples for mean normal stresses up to 7.5 kb.[5] More recent data obtained by the Waterways Experiment Station (WES) validated the older data but ranges only up to 3.5 kb.[7] To complete the tensorial model, a guess is made for the yield strength of failed concrete of about a tenth of the virgin yield strength. According to WES, concrete "...is capable of a surprising amount of plastic deformation ... ."[7] Consequently, we choose 10% plastic strain as the criterion for maximum failure. This completes the rudimentary model we use. More sophisticated models exist, but the paucity of data does not warrant the introduction of any more 'adjustable' parameters. One such variable might be to introduce rate dependencies for yield strengths.

Future Work

Our near-term plans are to insert the above concrete model into a code with an Eulerian capability such as CALE. This would permit us to calculate the effect of a munition that penetrates the concrete more deeply and is probably much more complicated hydrodynamically than the rebar cutter. The ease of performing calculations would be greatly improved, and parameter studies such as a study of the possible importance of rate effects would be greatly facilitated. A recalculation of the rebar approximations should make it possible to demonstrate major shear displacements at failure surfaces. Furthermore, the details of the surface crater formation and its resulting width could be reexamined, with the question of zoning definition of the crater wall removed from consideration. The result might provide additional insight into the spall of the front surface of the concrete in the vicinity of the bore hole. If our results continue to be encouraging, this will constitute a first version of a design tool that can calculate complete systems of munitions and targets self-consistently.

Acknowledgements

We thank Russell W. Rosinsky for providing analytic design data and Robert M. Kuklo for providing experimental details on the rebar-cutting charge for the PAM.

1. National Concrete Ready-Mix Association, Private Communication, Silver Spring, Maryland (October 1992).
2. R.W. Rosinsky, "An Annular Rebar Cutting Charge for the Penetration Augmented Munition," *Munitions Technology Development 1990*, Sandia National Laboratories, Albuquerque, New Mexico, SAND90-1104 (1991).
3. J.O. Hallquist, *User's Manual for DYNA2D—an Explicit Two-dimensional Hydrodynamics Code with Interactive Rezoning and Graphical Display*, Lawrence Livermore National Laboratory, Livermore, California, UCID-18756, Rev. 3 (1988).
4. V.G. Gregson, Jr., *A Shock Wave Study of Fondu-Fyre WA-1 and a Concrete*, General Motors Technical Center, DNA 2797 F (February 1972).
5. J. Chinn and R.M. Zimmerman, *Behavior of Plain Concrete Under Various High Triaxial Compression Loading Conditions*, University of Colorado, WL TR 64-163 (August 1965).
6. N.C. Holmes, Private communication, Lawrence Livermore National Laboratory, Livermore, California (1992).
7. B.D. Neeley, M.I. Hammons, and D.M. Smith, *The Development and Characterization of Conventional-Strength and High-Strength Portland Cement Concrete Mixtures for Projectile Penetration Studies*, Waterways Experiment Station, Technical Report SL-91-15 (1991).
8. T. Tang, *Behavior of Concrete Under Dynamic Loading*, Ph.D. Dissertation, University of Florida (1990).
9. Waterways Experiment Station, Private communication, Vicksburg, Mississippi (March 1992).
10. J.L. Levatin, A.V. Attia, and J.O. Hallquist, *KDYNA User's Manual*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-ID-106104 (1990).
11. D.E. Burton, L.A. Lettis, Jr., J.B. Bryan, and N. Frary, *Physics and Numerics of the TENSOR Code*, Lawrence Livermore National Laboratory, Livermore, California, UCID-19428 (1982).
12. R.E. Tipton, *CALE User's Manual*, Version 920721, Lawrence Livermore National Laboratory, Livermore, California (1992).
13. K.H. Sinz, "A Consistent Tensile Failure Treatment for Hydrocodes", unpublished (1987).

Benchmarking of the Criticality Evaluation Code COG

William R. Lloyd and H. Peter Alesso
Fission Energy and Systems Safety Program

John S. Pearson
Health and Safety Division Hazards Control Department

The purpose of our technology transfer project is to benchmark the Lawrence Livermore National Laboratory computer code COG for nuclear criticality evaluations. COG is potentially the most accurate computational tool available for these evaluations.

Introduction

Assurance of subcriticality is the most important element in any nuclear facility operation involving special nuclear materials. A good understanding of the detailed nuclear fission process is the only way to assure subcriticality. Today, this assurance is provided by using an analytical computational tool to evaluate and analyze all possible scenarios and geometries. The reliability of the evaluation results depends upon the accuracy of the computational tool in representing the realistic condition of the operation in question. Proof of the accuracy of the computational tool in turn depends upon the proper benchmarking of the code against actual nuclear fission process experiments[1] (criticality experiments) similar to the operation being evaluated. The applicability of a code to a specific geometry and condition depends on whether a benchmark has been done for a similar type of experiment and how accurately the code predicts the result of the experiment.

The criticality evaluation code commonly used in both government and industry today is the KENO-Va code with the four cross-section sets available to it on the SCALE system,[2] all developed at Oak Ridge National Laboratory (ORNL). Development of KENO began in the Mathematics Division of ORNL in 1958. In the 1970's and 1980's, the Nuclear Regulatory Commission funded the development of the SCALE system, a modular code system for performing standardized computer analyses for licensing evaluation. KENO, written more than 20 years ago, used methods as exact as was possible at that time. Today, much better physical data are available, but these data do not fit the forms used by KENO.

Development of COG[3] began in 1983 at Lawrence Livermore National Laboratory (LLNL) as a shielding code. The principal consideration in developing the code was that the resulting calculation was to be as accurate as the input data provided to the code. Cross-section data were presented by evaluators in the 1980's as point-wise data; i.e., as a series of cross-section points as a function of neutron energy, for example, with the understanding that interpolation between adjacent points produces results as good as the data. COG was written to use this form of the cross-section data directly. The angular scattering data are likewise presented and used as the evaluators present them. No approximation has been made that would compromise the accuracy of these data. The geometric description of a problem for input into a criticality code should be as exact as possible. COG permits specification of a surface defined by input analytic equations containing terms up to the fourth degree.

Objectives

COG was developed on LLNL Cray Computers using the New Livermore Time Sharing System. COG is being moved from Crays to workstations that include the Hewlett Packard (HP) 9000/730 and a SUN computer using the UNIX operating system for greater availability.

COG and its cross-section set are being benchmarked against at least 250 criticality experiments to understand the bias of COG in a range of criticality situations.

The transfer of COG from LLNL to universities, industry, and other Department of Energy laboratories will be accomplished by placing it in the Nuclear Systems Safety Center (NSSC). Criticality and shielding services using COG can be offered from this system.

COG geometry input preparation can be tedious for complicated geometrical systems. The three-dimensional computer-aided-design (CAD) software Pro/ENGINEER and the LLNL code Pro/COG will be used to generate geometry input for COG.

Progress

COG has 160 subroutines that include 47 geometry subroutines and 86 cross-section subroutines. In May 1992, COG was compiled, assembled, and run on an HP 9000/730 computer. Further detailed checks of capabilities in COG, such as Russian roulette, path stretching, and importance weighting, were initiated on the HP 9000/730.

Preparation of 100 critical experiment models as input to COG were completed and run on LLNL Cray computers using LLNL Evaluated Neutron Data Library (ENDL) cross sections.

The Evaluated Nuclear Data File/B-V (ENDF/B-V) was converted from its parameterized format to a point-wise format suitable for use with COG. Techniques developed in this work will also permit conversion of other evaluated neutron libraries, including ENDF/B-VI, the Japanese Evaluated Neutron Data File-3 (JENDL-3), the Joint European File-1 (JEF-1), and BROND-2 (a Russian file). Benchmarking activities will provide criteria for unifying these evaluations into a single nuclear data library.

Output from the CAD software Pro/ENGINEER was combined with LLNL's Pro/COG to produce some COG geometry input for arrays of nuclear fuel rods.

Future Work

COG currently runs on an HP 9000/730 workstation with a UNIX operating system. Conversion of COG to run on a SUN Microsystems model S10MX is planned. Completed testing of the deep-penetration and code-optimization features is planned on both computers.

The running of models from at least 250 benchmark critical experiments is planned for the SUN and HP workstations.

Continued use of Pro/ENGINEER and development of the LLNL code Pro/COG to produce COG geometry input for arrays of fuel rods, the torus, and the sphere are planned.

Establishment of the NSSC is planned to provide a mechanism for exporting COG to other laboratories, universities, and industry.

Acknowledgements

The authors wish to recognize the individual contributions of C. Annese, R. Buck, D. Cullen, P. Giles, S. Hadjimarkos, D. Heinrichs, R. Howerton, D. Lappa, E. Lent, D. Resler, T. Wilcox, and R. White.

1. B.L. Koponen, T.P. Wilcox, Jr., and V.E. Hampel, *Nuclear Criticality Experiments from 1943 to 1978, An Annotated Bibliography, Vols. 1–3*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-52769 (1979).

2. *SCALE: A Modular Code System for Performing Standardized Computer Analyses for Licensing Evaluation*, NUREG/CR-0200, Revision 4 (ORNL/NUREG/CSD-2/R4), Vols. I, II, and III (1991); available from Radiation Shielding Information Center as CCC-545.

3. T.P. Wilcox, Jr. and E.M. Lent, *COG—A Particle Transport Code Designed To Solve the Boltzmann Equation for Deep-Penetration (Shielding) Problems, Vol. 1*-User manual, Lawrence Livermore National Laboratory, Livermore, California, M-221-1 (1989).

Fast Algorithm for Large-Scale Consensus DNA Sequence Assembly

Shin-yee Lu
Engineering Research Division
Electronics Engineering

Michael E. Colvin and
Richard S. Judson
Center for Computational Engineering
Sandia National Laboratory
Livermore, California

Elbert W. Branscomb
Biomedical Sciences Division

A major current objective of the Human Genome Center at Lawrence Livermore National Laboratory is completion of the physical map for chromosome 19. In the coming years, more emphasis will be given to completely sequencing stretches of DNA and to analyzing these sequences. The goal of our effort is to develop algorithmic and computational tools needed to meet new challenges that will arise from this shift of emphasis.

This article describes our approach, called 'key-search,' to DNA sequence assembly. The computational complexity of the key-search algorithm is nearly directly proportional to the number of DNA bases to be assembled. We have completed the implementation of the algorithm. We are now testing our assembly program, using a data set provided by the National Institute of Health.

Introduction

The process of DNA sequencing is typically accomplished by using a so-called 'shotgun sequencing' approach. The method involves sequencing randomly overlapping small fragments (200 to 500 base pairs) taken from a much larger piece (e.g., 40,000 base pairs), to a 5- to 10-fold redundancy; i.e., the total number of base pairs sequenced is 5 to 10 times the size of the original piece. The original large sequence can be recovered in principle by pasting together fragments that share common subsequences.

Sequence assembly is computationally difficult for two reasons. First, there are a large number of fragments (1000 or greater); a direct comparison of every pair to determine which pairs contain common subsequences is very slow. The second problem arises because of errors in data collecting, imperfections in the fragmentation processes, and the statistical nature of fragment selection. Existing methods use sequence alignment programs to determine overlap between fragments, and optimization methods to paste overlapping fragments together. These methods generally do not perform well for assembling large sequences.

We have developed a new approach to the assembly problems, using a 'key-search' method based on the computer science idea of hash tables. We first encode every 15-base segment of every fragment into an integer called a key. These keys are sorted and stored in a table, together with pointers to the fragments in which they were found and the locations along those fragments. Starting from any key, one can determine the adjacent key (e.g., to the right) in the original sequence by examining all of the fragments that contain the current key (i.e., bases 1 to 15), generating a consensus for the next base (i.e., base 16). The next key is generated from bases 2 to 16. The process is repeated key by key until an end condition is detected. Similarly, we can reconstruct the sequence to the left of the starting key. **Figure 1** illustrates the basic concept of this approach. Extensions such as using consensus calculation at each base for automatic error correction, and methods for resolving the confusion that may be caused by motifs, periodic patterns, and long repeats are added to increase the robustness of the assembly program.

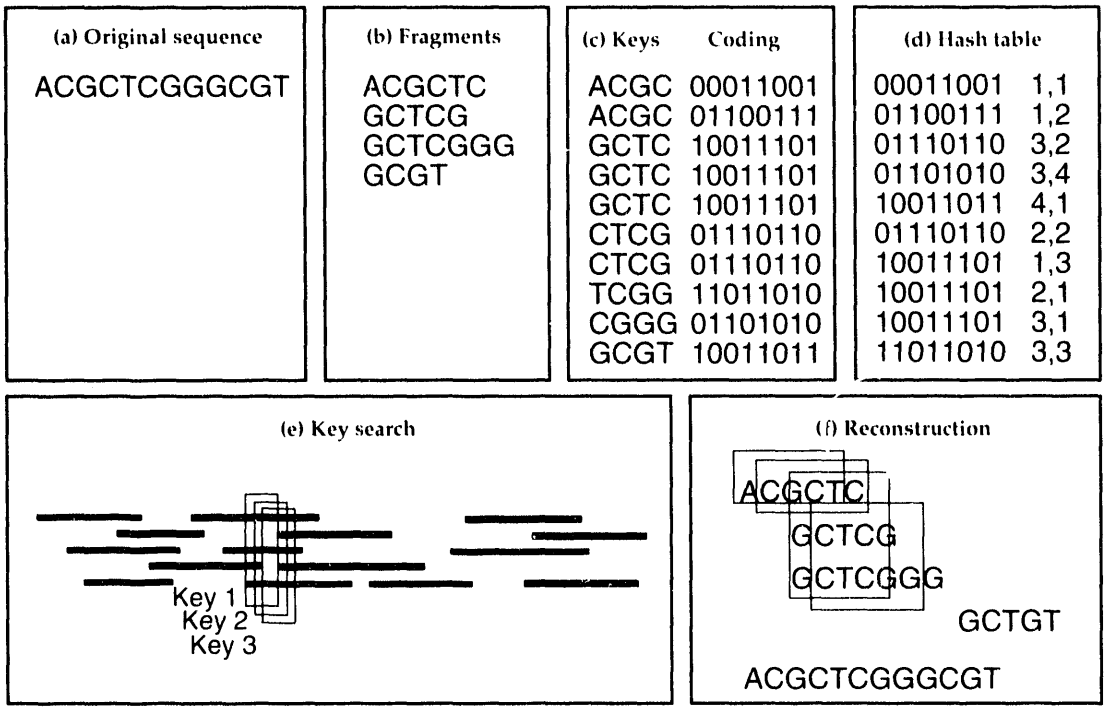


Figure 1. The basics of the key-search approach to DNA sequence assembly: (a) a DNA sequence, (b) fragments from the original sequence, (c) four-base keys and the encoding of these keys from all the fragments, (d) a sorted table of all the keys, (e) a base-by-base reconstruction procedure, and (f) final reconstruction.

The major advantages of our approach are its computational efficiency and its potential for generating more reliable reconstructions. Our method can generate a complete consensus sequence, the exact location at which each fragment resides along the consensus sequence, and locations where errors occur. This information provides necessary data for a statistical estimation of confidence in the reassembled sequence. Biologists consider such estimation critical for their applications, and often fault current assembly methods for lacking the ability to provide confidence estimation.

Progress

We have completed the implementation of the key-search algorithm and tested it on a known DNA sequence of approximately 33,000 base pairs to validate the approach. A fragmentation program that simulates shotgun sequencing and generates a synthetic fragment database was implemented. We typically assume 10-fold coverage, 200- to 400-base-pair fragment length, and random cutting sites. We then corrupt the sequence with errors at a realistic rate based on published data. Errors increase exponentially along the length of the fragment, from 1% for lengths below 200 base pairs to 7% at 500 base pairs. Our assembly program successfully reconstructs the complete sequence, except for a residual error rate of about 2%. The reconstruction uses less then 40 min on a SUN Sparc Station II. We also tested the program with fragment data sets corresponding to less coverage, at 7-, 8-, and 9-fold. As the coverage decreases, the reconstruction can only generate islands, because some regions of the original sequence are not covered in the fragment database. Based on Monte Carlo simulation, 10-fold coverage is the minimum required to cover the entire sequence.

Our current task is to reconstruct a large sequence from an actual fragment data set provided to us by the National Institute of Health (NIH). The NIH database has a relatively low coverage of 5- to 6-fold, so we anticipate finding gaps in the original sequence. However, the reconstruction code is able to generate several islands in the 6000- to 10,000-base-pair size range, for a total sequence of approximately 34,000 base pairs. The reconstruction result will be evaluated by biologists at the Human Genome Center at Lawrence Livermore National Laboratory (LLNL).

Future Work

Real fragment data sets are currently being generated at the LLNL Human Genome Center. We will test the assembly program on these data. We will also compare our reconstruction results with results generated by other available software packages such as the sequence analysis software from Intelligenetics, Inc., and the Staden package from Cambridge University.

Using Electrical Heating To Enhance the Extraction of Volatile Organic Compounds from Soil

H. Michael Buettner and
William D. Daily
Engineering Research Division
Electronics Engineering

We have developed a method of using electrical soil heating in combination with vacuum venting to enhance the removal of volatile organic compounds from contaminated soil. The results of two engineering-scale tests show that this technology has great potential for environmental remediation at both government and private facilities.

Introduction

The problem of contamination of ground water and soil by volatile organic compounds (VOC's) is widespread in this country. Vacuum venting has long been used as a remediation method in such cases. We proposed that electrical soil heating (joule heating) could be used in combination with vacuum venting to enhance the removal of VOC's. We describe here the results of work to demonstrate electrical heating of the ground in engineering-scale tests for use as an adjunct to vacuum venting or cyclic steam injection for the removal of VOC's from soil.

Progress

We performed two engineering-scale tests. The first of these, in September, 1991 at an uncontaminated site, Sandia National Laboratories, Livermore, California, proved that soils can be effectively heated using powerline-frequency energy. The second test, from May to July, 1992 at Lawrence Livermore National Laboratory's (LLNL) Site 300, which is contaminated with trichloroethylene (TCE), proved that electrical heating can enhance VOC removal from soils.

Our purposes were (1) to learn about the practical aspects of electrical heating such as selection and sizing of electrode materials and wires, and maintenance of low contact resistance between the electrodes and the ground; (2) to compare actual heating rates with those based on simple calculations; (3) to provide data for scaling the experiment up for application at other contaminated sites; (4) to provide data for estimates relating to the economics of electrical heating; and (5) to demonstrate that electrical heating enhances extraction of VOC's and to quantify the effect.

For our first engineering-scale field test, at an uncontaminated site, we used a pattern of six heating wells equally spaced on the circumference of a circle with a diameter of 6.1 m (20 ft). The electrodes were made of stainless steel tubing sections, and the contact resistance was maintained at a low value by saturating the sand pack around the electrode with water via a feed tube. The heating wells were powered with 3-phase, 400-V, 60-Hz power supplied by a 125-kVA generator.

Fixed thermocouples were used to monitor the temperature as a function of time during the test. We ran the test around the clock for 10.76 days, then during the day only for four additional days. The currents to the electrodes, and thermocouple temperatures were monitored on a regular basis. At the end of the 24-h/day heating period, the temperature in the center of the pattern (the coldest point) at a depth of 4.88 m (16 ft) rose from a starting value of 19°C to 38°C. During the daytime-only heating, the temperature rose to 44°C. At this point, the power was turned off, and the temperature continued to rise to 54°C in a 10-day period, after which we stopped temperature monitoring. Other thermocouples nearer to the periphery read as high as 73°C. These experimental results agree closely with very simple calculations

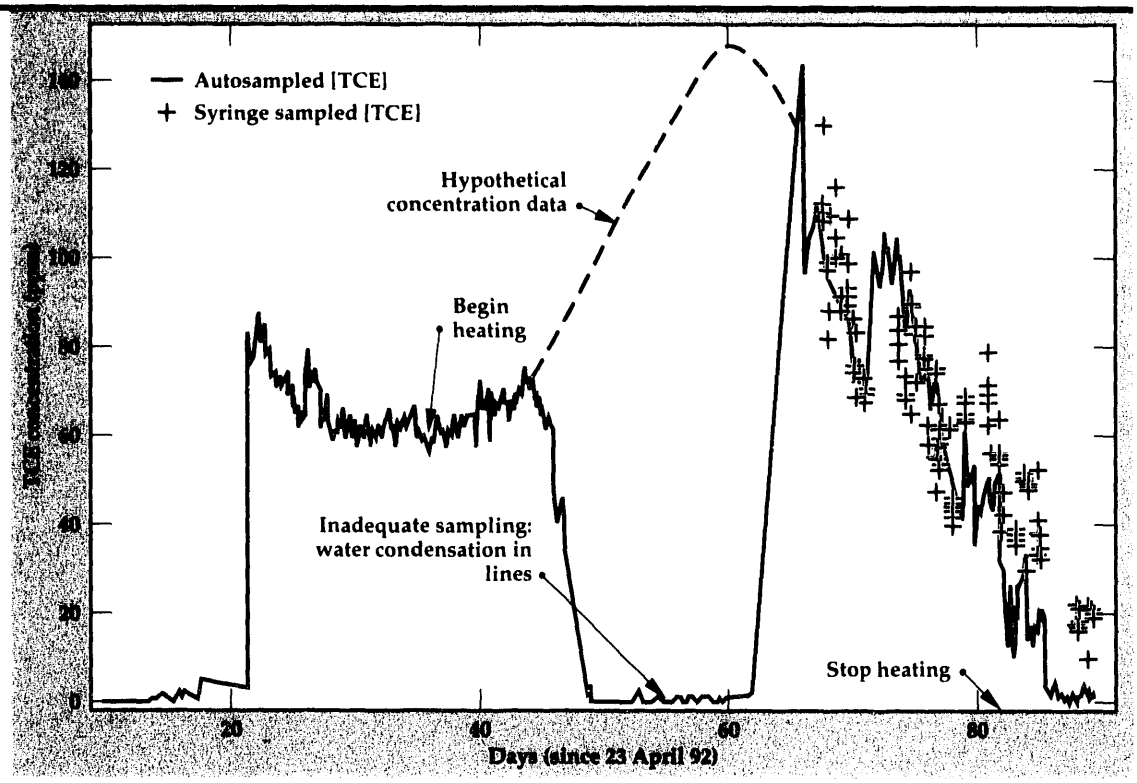


Figure 1. TCE concentration vs time before, during, and after electrical soil heating at LLNL's Site 300.

based on a two-dimensional model, assuming homogeneous electrical and thermal properties. Over the duration of the heating phase of the experiment, the total energy dissipated in the ground was about 15,000 kWh.

For our second test, at a site contaminated with TCE, the heating wells were also located on a 6.1-m circle. A vapor extraction well was located in the center of the pattern. The heating wells were powered with 3-phase, 480-V, 60-Hz power supplied by a 100-kVA generator. Again, the contact resistance of the electrodes was maintained at a low value by saturating the sand pack around the electrode with water via a feed tube.

We ran the test during the day only for 40 days. The currents to the electrodes and thermocouple temperatures were monitored regularly. During the heating phase of the experiment, the temperature of the vapors extracted from the central well rose from 16°C to 36°C, and continued rising thereafter to 39°C, when we stopped collecting data. The TCE concentration in the collected vapor decreased steadily as vacuum was applied to the central well during the period before heating, to a low of about 60 ppm. Once electrical heating began, the concentration increased. The concentration rose to over 140 ppm during the heating period, and then decayed to values of less than 20 ppm. Results are shown in **Fig. 1**. The amount of electrical energy deposited in the ground was about 9600 kWh.

The lessons learned as a result of this work have been applied to a much larger and highly visible project at LLNL, the Dynamic Underground Stripping Project.[1,2] This project seeks to clean up approximately 17,000 gallons of gasoline from the soil and ground water at an old gasoline station site. In this demonstration project, steam injection and vacuum extraction are used to clean gasoline from the sands and gravels, while electrical heating drives the gasoline from the clay layers.

Future Work

The technology we have developed has great potential for the Department of Energy (DOE) and for the private sector. LLNL is at the forefront of this work. We are actively seeking partners inside and outside DOE for further development and/or licensing of the technology. For example, the technology is being considered for remediation work at the DOE's Rocky Flats, and for facilities of British Petroleum of America.

1. R. Aines and R. Newmark, "Rapid Removal of Underground Hydrocarbon Spills," *Energy and Technology Review*, Lawrence Livermore National Laboratory, Livermore, California (July 1992).

2. "Ground water cleanup researchers make headway," *Newsline*, Lawrence Livermore National Laboratory, Livermore, California (October 2, 1992).

Fabrication Technology

The mission of the Fabrication Technology thrust area is to have an adequate base of manufacturing technology, not necessarily resident at Lawrence Livermore National Laboratory (LLNL), to conduct the future business of LLNL. Our specific goals continue to be to (1) develop an understanding of fundamental fabrication processes; (2) construct general purpose process models that will have wide applicability; (3) document findings and models in journals; (4) transfer technology to LLNL programs, industry, and colleagues; and (5) develop continuing relationships with the industrial and academic communities to advance our collective understanding of fabrication processes.

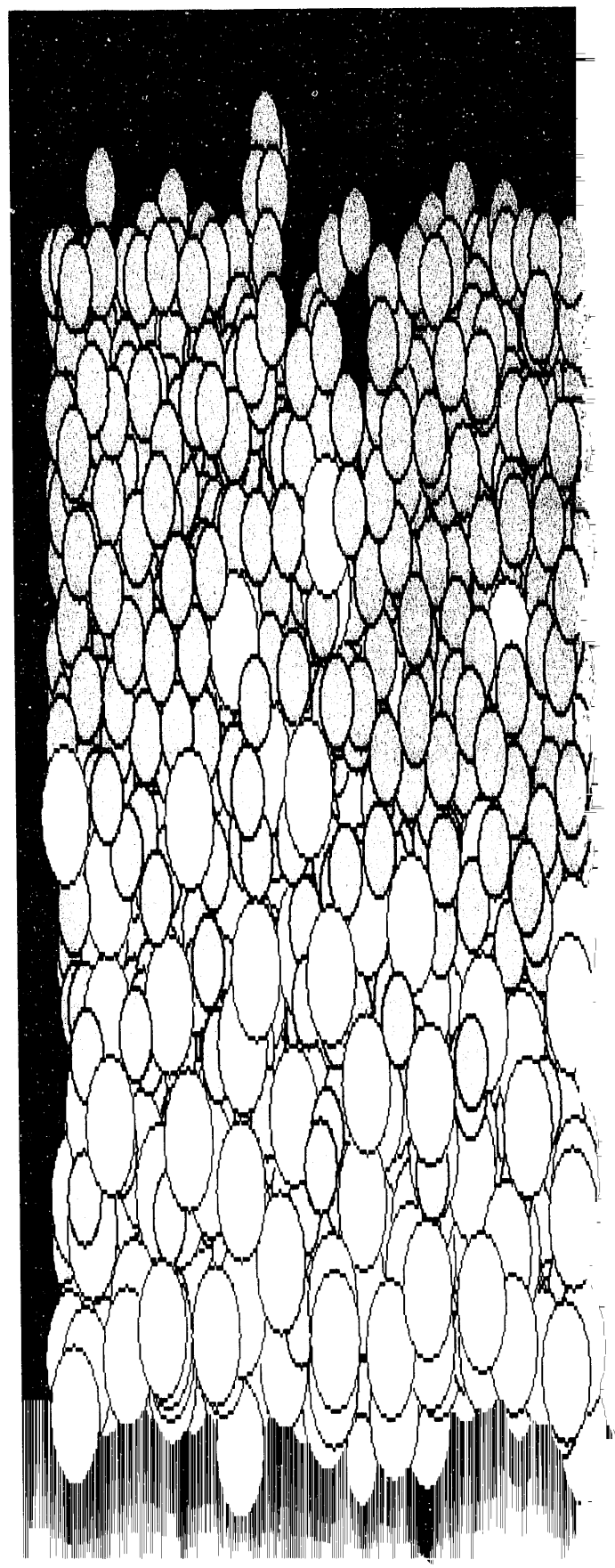
The strategy to ensure our success is changing. For technologies in which we are expert and which will continue to be of future importance to LLNL, we can often attract outside resources both to maintain our expertise by applying it to a specific problem and to help fund further development. A popular vehicle to fund such work is the Cooperative Research and Development Agreement with industry.

For technologies needing development because of their future critical importance and in which we are not expert, we use internal funding sources. These latter are the topics of the thrust area.

Three FY-92 funded projects are discussed in this section. Each project clearly moves the Fabrication Technology thrust area towards the goals outlined above. We have also continued our membership in the North Carolina State University Precision Engineering Center, a multidisciplinary research and graduate program established to provide the new technologies needed by high-technology institutions in the U.S. As members, we have access to and use of the results of their research projects, many of which parallel our own precision engineering efforts at LLNL.

Kenneth L. Blaedel
Thrust Area Leader

Section 5



5. Fabrication Technology

Overview

Kenneth L. Blaedel, Thrust Area Leader

Fabrication of Amorphous Diamond Coatings

Steven Falabella, David M. Sanders, and David B. Boercker 5-1

Laser-Assisted Self-Sputtering

Peter J. Biltoft, Steven Falabella, Steven R. Bryan, Jr., Ralph F. Pombo, and Barry L. Olsen 5-5

Simulation of Diamond Turning of Copper and Silicon Surfaces

David B. Boercker, James Belak, and Irving F. Stowers 5-7

Fabrication of Amorphous Diamond Coatings

Steven Falabella and David M. Sanders
Materials Fabrication Division
Mechanical Engineering

David B. Boercker
Condensed Matter Physics Division
Physics Department

Amorphous diamond is a hard, electrically insulating, inert and transparent form of carbon that has the sp^3 bond character of crystalline diamond, but lacks a long-range ordered structure. The potential applications of amorphous diamond (a:D) are many. This material has several important advantages over conventional chemical-vapor-deposition diamond coatings, making it a more attractive coating for applications such as cutting tools, tribological surfaces, spacecraft components, and medical implants. In FY-92, we produced carbon coatings with hardness rivaling that of natural diamond, and began to evaluate the use of this material in practical applications. We have produced amorphous diamond films on a routine basis, and have produced coatings up to 8 μm thick on carbide tool bits. The combination of extreme hardness, low atomic number, smoothness, low friction, and low deposition temperature make a:D unique in the world.

Introduction

The physical properties of diamond make it an ideal material for many critical applications. However, natural diamonds are rare, expensive, and too small for many applications. A substantial amount of work is being done to produce diamond coatings on less expensive substrates, to take advantage of the properties of diamond without the need for large diamond monoliths. There are four critical problems that need to be solved before diamond coatings will be practical:

(1) Temperature of deposition. High process temperatures eliminate aluminum, tool steels, glasses, and polymers as possible substrate materials, limiting the usefulness of the coating. Also, heating and cooling of substrates adds time, complexity, and expense to the coating process.

(2) Adhesion to substrate. Thin films rely on the substrate for much of their mechanical integrity, depending on adhesion to the substrate for support. Failure of adhesion usually means unpredictable and rapid failure of the coated part.

(3) Stress. Internal stress limits the permissible thickness of a coating when the stress in the coating causes delamination or deforms the substrate.

(4) Smoothness of coating. In tribological applications, smoothness is essential for low friction and long life. Also, for optical coatings, any coating roughness will degrade the performance of the optic.

Diamond films produced by chemical vapor deposition have difficulty in all four areas. The adhesion is poor; deposition temperature is generally above 800°C; thermally induced stress is often excessive; and the polycrystalline films produced have high surface roughness, requiring expensive polishing.

The situation is very different for a:D. Amorphous diamond coatings are produced by the condensation of carbon ions on cooled substrates (at room temperature or below). They also replicate the substrate surface finish, and can be very adherent. We feel that only adhesion and stress are still problems, and may exclude the use of some substrate materials. However, an adherent interface can be created in several ways: a thin layer of a binder material can be deposited before coating with a:D; or, since the process is ion-based, substrate biasing can form a diffuse, adherent interface. Stress can be lowered by several means:

increasing the incident ion energy by rf-biasing the substrate during deposition; increasing the substrate temperature; and incorporating impurity elements in the film. These methods to reduce stress and improve adhesion may also reduce certain qualities of the coatings, so tradeoffs will need to be made.

We have identified four areas where the extraordinary properties of a:D can have a large impact. The first is the coating of tool bits for use on diamond turning machines to exploit the toughness, adhesion, hardness, and wear resistance of a:D. If successful, this process will lead to cost savings where the surface finish and precision required is less than that produced by diamond turning, yet better than can be produced by conventional cutting bits. At present, the finish obtained on a part is limited by the edge quality of our coated carbide bits, which in turn is limited by current polishing methods. If a better method can be found to form the tip radius of a coated bit, geometries and precision not practical with natural diamond could be achieved.

Second, the surfaces of metrology blocks, caliper faces, and precision slides can be coated and polished to provide hard, smooth, and wear-resistant surfaces that will not change dimensions or scratch the parts under test. This will allow more confidence in the continued accuracy of the tools, save re-calibration time, and prolong the life of the equipment.

Third, there are applications that would benefit from the tribological properties of a:D, in air and in vacuum. The friction coefficient for a:D is measured to be 0.2 or less in all conditions. There are several important areas where a long-life solid lubricant could have prevented the failure of mechanical systems on spacecraft, and would enable new mechanisms to be practical in spacecraft. A representative example is the Galileo probe's main antenna that failed to deploy due to the failure of the MoS_2 lubricant on its opening mechanism. Coating both disk surface and heads will reduce the damage caused by 'head crashes' and may enable magnetic recording media of higher density by allowing smaller head-to-disk distance.

Finally, there are several applications of a:D in the medical field. Due to the wear resistance and biocompatibility of a:D, the potential is great for coating scalpels, replacement-joint wear surfaces, and other implanted parts. If a suitable technique is developed to allow a coated blade to achieve the sharpness of a natural diamond scalpel, the potential benefits would be tremendous. Incisions made by diamond scalpels heal up to five times faster than those made with steel blades, reducing recovery times (and hospital costs) for many surgical procedures. The high cost of diamond scalpels (several thousand dollars each) is now the main limit to their use.

Progress

In FY-92, we produced carbon films with our filtered cathodic-arc system, which was developed in previous years. The cathodic-arc source produces a carbon ion beam from a graphite target, in a high vacuum environment.[1] Our goals for the year were to investigate the conditions under which a:D is formed, to improve adhesion to various substrate materials, to model the deposition process using molecular dynamics (MD), and to reduce residual stress in the films, which is required to deposit greater thicknesses.

We installed a cooled and biasable holder to control the substrate temperature during deposition. Initially, the holder was cooled by liquid nitrogen, but we found that water cooling produced equivalent results. By using a high-voltage bias for the first few seconds of coating, we have produced coatings on cemented carbide tool bits with adhesion above 10 kpsi (limit of the Sebastian pin-pull tester). We are investigating methods that will measure adhesion to higher values.

We were able to achieve hard carbon coatings that are low in hydrogen content. The hardness of carbon films is inversely related to the hydrogen content; e.g., 10 to 20% hydrogen in a carbon film (known as diamond-like-carbon, or DLC) reduces the hardness by a factor of four. The hydrogen content of our coatings was measured to be less than 0.1%, using forward recoil scattering (FRS). We determined the density of our films from the areal density obtained, using Rutherford backscattering (RBS) and the film thickness. We measure the density of a:D to be 2.7 ± 0.3 g/cc, which is between graphite at 2.26 g/cc and diamond at 3.5 g/cc.

One of the most appealing properties of amorphous diamond is its extraordinary hardness. However, standard hardness tests made by indenting are generally difficult to interpret when the coating is thin and harder than the substrate material. To get a true measurement of the coating hardness, the indent depth must be less than 7 to 20% of the coating thickness.[2] Quantitative hardness tests are in progress with an ultra-microhardness tester, which uses such a small indent that the measurement is not influenced by the substrate. A standard Vickers indent of a tungsten carbide tool bit coated with 8 µm of a:D with loads up to 500 g gives a hardness of 10,000 ± 10%, the same as

natural diamond. At 500 g, the indent depth is 17%. To put the severity of this test in perspective: the stress put on the film at the 500-g load is over 17×10^6 psi. For thinner coatings, the Vickers test gives a hardness above 5000 at low indent loads (25 g) and then decreases, as the load increases to roughly the substrate hardness value (hardness of tungsten carbide is ~2000 H_V). To get another assessment of the hardness, we used an abrasion test. We assessed the hardness of the a:D coatings that we produced, by abrading various hard materials against a coated plate. We were able to polish facets in all materials attempted, including natural and synthetic diamond, indicating that the coating is approximately as hard as diamond. This may point to yet another application, i.e., the surfacing of ceramic, or even diamond tools.

The greatest difficulty with a:D films is their high intrinsic stress. Our filtered cathodic-arc source produces a fully ionized beam of carbon with a mean energy of 22 eV[3], and produces stress levels of 6 to 10 GPa. This can be reduced by increasing the incident ion energy impinging on the substrate. We used a 13.56-MHz rf supply to provide a bias during deposition. Since the films produced are non-conductive, rf bias is required to maintain the potential at the film surface during coating. We have reduced the intrinsic stress in a:D films by a factor of two using bias alone, and by a factor of five using a combination of bias and the incorporation of 7% nitrogen in the films. A plot of the residual stress vs bias voltage on the substrate (DC level) is shown in **Fig. 1**. Coatings with and without nitrogen are shown. Although residual stress is reduced by the addition of nitrogen, the measured hardness of the films is reduced to ~6000 H_V, as noted above. Residual stress was inferred from the bowing of two-inch silicon wafers. On carbide-forming materials, the adhesion is sufficient to produce thick coatings, without delamination of the coating caused by the compressive stress, as long as the bias voltage is kept above 150 V during deposition. The 8-μm-thick coating produced on a tungsten carbide tool bit was limited only by source material depletion.

The fine structure of a:D was characterized by TEM and electron diffraction. TEM showed no evidence of any ordered structure down to 10Å, indicating its amorphous nature. Unlike natural diamond or DLC, a:D has a flat transmission spectrum from 0.8 to >50 μm, which is due to its amorphous nature and the lack of hydrogen. The transmission of a free-standing film was measured using a FTIR spectrophotometer. From the interference between the front and back surfaces and

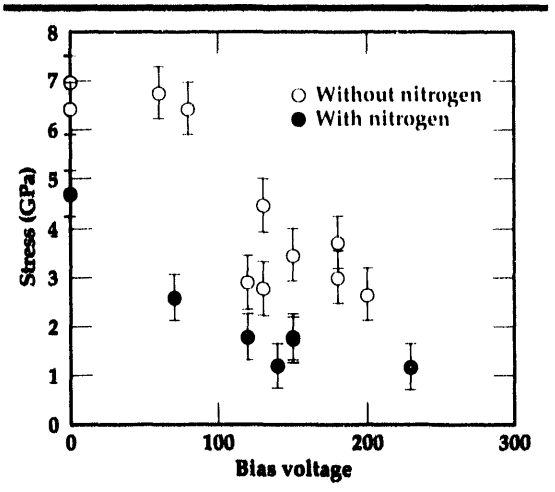


Figure 1. The intrinsic stress in amorphous diamond films vs bias voltage. Stress is reduced substantially by the addition of bias during deposition, and even further by the addition of nitrogen. Data taken with no nitrogen during deposition are in open circles; data taken with a nitrogen background are in solid circles. In both cases, the stress reaches its lowest value around 150 V and is roughly constant above that value.

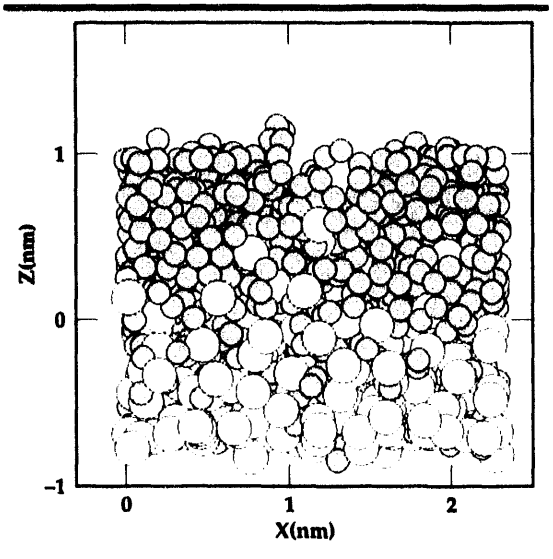


Figure 2. A molecular dynamics simulation of 20 eV carbon (880 atoms) impinging on a silicon surface. Substantial mixing occurs at the interface. The carbon atoms are shown as dark gray circles; the silicon atoms are light gray. The view is parallel to the original silicon surface.

the measured thickness of the films, we determined the index of refraction of our a:D to be in the range 2.47 to 2.57. This is close to 2.42, the refractive index of natural diamond.

Using MD simulations, we have modeled the condensation of carbon atoms onto a silicon substrate to see the effects of deposition energy on coating structure and stress. **Figure 2** shows carbon deposited on a silicon surface. Even at the deposition energy of 20 eV, there is substantial

mixing at the interface. We are now using the code results to interpret the electron diffraction measurements. By Fourier transforming the atom positions in the simulation, we were able to closely match the observed positions of diffraction rings.

Future Work

We have described only a few of the possible applications of a:D, with others to be realized as the material becomes better characterized. The combination of extreme hardness, low friction, smoothness, and low deposition temperature make amorphous diamond a unique and very promising material.

The next step in the development of this material would be to test our amorphous diamond films in practical applications. However, we have not yet obtained continued funding for this project.

Acknowledgements

We wish to thank R. Musket for providing the RBS and FRS measurements; R. Chow and G. Loomis for the optical measurements; M. Wall for the TEM and electron diffraction work; J. Ferriera for the hardness tests; J.H dePruneda for insight into medical applications; and the Vacuum Processes Laboratory staff for their technical support.

1. S. Falabella and D.M. Sanders, *J. Vac. Sci. and Technol. A* **10** (2), 394 (1992).
2. C. Feldman, F. Ordway, and J. Bernstein, *J. Vac. Sci. and Technol. A* **8** (1), 117 (1990).
3. P.J. Martin, S.W. Filipczuk, R.P. Netterfield, J.S. Field, D.F. Whitnall, D.R. McKenzie, *J. Mater. Sci. Lett.* **7**, 410 (1988).

Laser-Assisted Self-Sputtering

Peter J. Biltoft,
Steven Falabella,
Steven R. Bryan, Jr., and
Ralph F. Pombo
Materials Fabrication Division
Mechanical Engineering

Barry L. Olsen
Materials Division
Chemistry and Materials Science Department

Our goal for FY-92 was to demonstrate laser-assisted self-sputtering as a method for sputter deposition of thin film coatings in a high vacuum environment.

Introduction

Our experimental program was designed to investigate merging the technology of magnetron sputtering[1] and laser ablation[2] to create a well-controlled deposition process free of the need for a process gas. Self-sputtering of copper, using a conventional magnetron sputter gun, has been reported.[3] In this process, a glow discharge plasma was initiated by operating a magnetron in the conventional manner, with argon as the process gas at a pressure in the range of from 5 to 20 mTorr. After the plasma was well established, the process gas pressure was slowly reduced. As this was done, sputtering was maintained by ionization of sputtered copper atoms in close proximity of the cathode. We hope to demonstrate self-sputtering initiated by a laser-induced plasma in the absence of any process gas.

Progress

Our first goal was to design and build a fixture that would accommodate installation of a

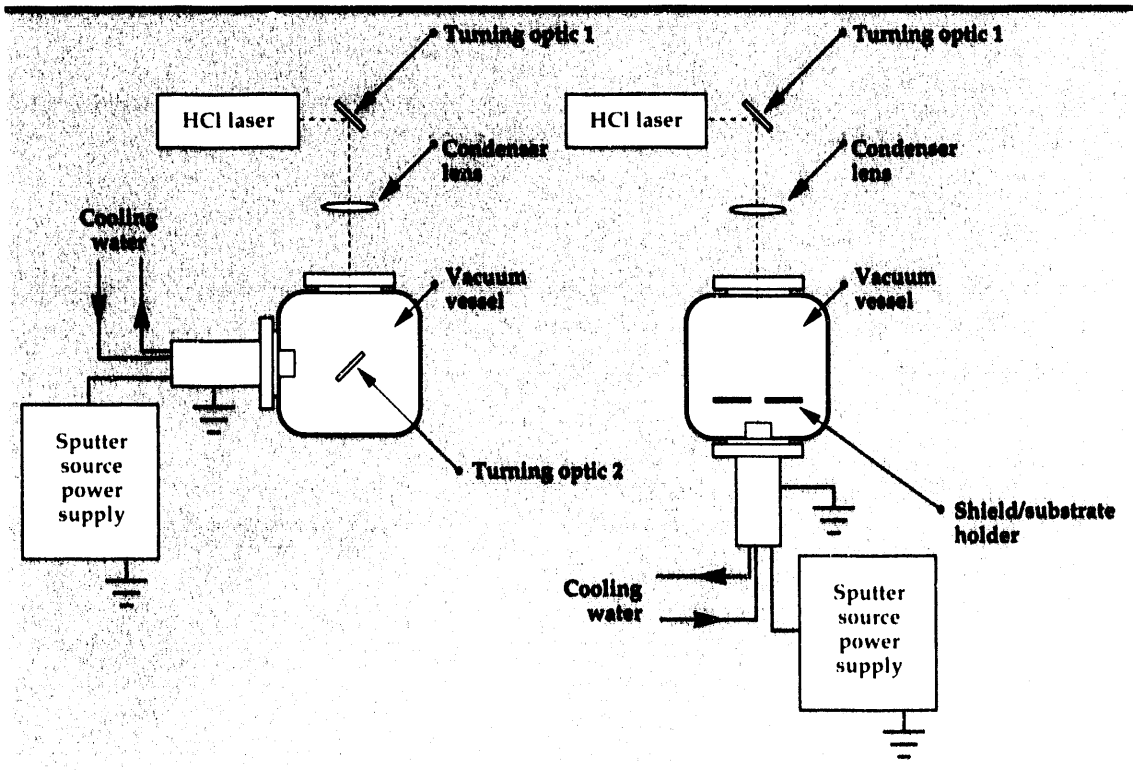


Figure 1. Initial (left) and final (right) configuration of the laser-assisted self-sputtering apparatus.

conventional magnetron sputter gun into an existing vacuum vessel designed for thin film growth by laser ablation. We selected a small, commercial, sputter-deposition source (2-in. US gun) for our first evaluation. A schematic of the experimental apparatus is presented in **Fig. 1** (left). Initial deposition runs were conducted at pressures below 5×10^{-5} Torr, as measured using a hot cathode ionization gauge on the vacuum vessel. We used a pulsed-output HCl laser operating at a wavelength of 308 nm to initiate the plasma. Typical operating parameters for the laser were 1- to 10-Hz repetition rate and 160- to 320-mJ pulse power. The laser beam was de-magnified using a 500-mm focal length, plano-convex lens external to the vacuum vessel. Power density at the sputter source cathode was between 21 and 42 J/cm^2. A high-output power supply designed for magnetron sputtering was used to bias the cathode to -5000 V. While we were able to briefly maintain a plasma at the sputter source, we discovered that the laser damaged turning optic 2 rapidly, reducing the power density we were able to deliver to the cathode. To rectify this problem, we reconfigured the apparatus as shown in **Fig. 1** (right). In the second configuration, we were able to initiate and maintain indefinitely a toroidal plasma at the sputter target. The color of the plasma for the copper target was bright green, indicating the presence of high concentrations of copper species in the plasma.[4] Using this setup, we deposited several thin films of copper. During 10-minute deposition runs, the magnetron power supply outputs indicated that the peak voltage was 5000 V, and average current was 0.1 A.

Using a storage oscilloscope, we observed that the voltage output of the magnetron power supply was reduced almost to zero following every laser pulse. In an effort to deliver higher current to the sputter source, we installed a 0.1 μF capacitor capable of operating at > 5 kV, between the magnetron sputter supply and the sputter source. No appreciable benefit was realized through this modification.

Results

We have deposited thin films of copper, aluminum, and tantalum by laser-assisted self-sputtering in a high vacuum environment. Deposition rates for the copper films were observed to be greater than 0.1 nm/s. This represents an increase in deposition rate of greater than a factor of 50 compared to pulsed-laser deposition of copper under identical circumstances.

Acknowledgements

We would like to thank Neil Lund, who designed and built the hardware used in this work, and Bob Tench for his assistance with the HCl laser.

1. J.L. Vossen and W. Kern, *Thin Film Processes*, Academic Press, Inc. (New York, New York), 1978.
2. D.B. Chrisey and A. Inam, *Mater. Res. Bull.*, 37 (February 1992).
3. R. Kukla, T. Krug, R. Ludwig, and K. Wilmes, *Vacuum R.* **41** (7–9), 1968 (1990).
4. *CRC Handbook of Chemistry and Physics*, CRC Press (Boca Raton, Florida), 1990.

Simulation of Diamond Turning of Copper and Silicon Surfaces

David B. Boercker and James Belak
Condensed Matter Physics Division
Physics Department

Irving F. Stowers
Precision Engineering Program
Engineering Directorate

We have applied molecular dynamics modeling to the diamond turning of a ductile metal (copper) and a covalent material (silicon). On the nanometer-length scale, both materials show ductile behavior, but the atomistic mechanisms that allow the behavior are significantly different in the two cases. In addition, we studied the wear of small diamond asperities while they machined a silicon surface.

Introduction

Diamond turning is, by now, a well established technique for machining high-quality surfaces with dimensional tolerances of a few tens of nanometers. This technique is particularly successful when applied to non-reactive, ductile metals such as copper. It is less useful when applied to carbide-formers, like iron, or to brittle materials. Tribochemical reactions can cause excessive tool wear, while brittle fracture produces surface damage. Recently, there has been interest in diamond turning silicon to obtain precisely shaped optical substrates. In this case, both problems occur. Silicon is a strong carbide former, and it is a covalently bonded, hard material that is prone to fracture.

To gain insight into the atomistic mechanisms of importance to diamond turning and to diamond tool wear, we have performed molecular dynamics (MD) simulations of the machining of both copper and silicon surfaces with diamond tools. The basic MD method is the same as that used previously[1] to simulate orthogonal cutting and nano-indentation. The simulations are performed in the rest frame of the cutting tool and follow the detailed, microscopic motions of the atoms, both in the tool and in the work piece, as it moves under the tool. Such simulations give good qualitative descriptions of chip formation and dislocation propagation.

The central input to the simulations is an appropriate interatomic force law. In the copper simulations, we use the embedded atom potential[2] for the interaction between two copper atoms, while the atoms in the diamond tool are assumed to interact with the metal atoms through a Lennard–Jones potential. For the silicon simulations, we have implemented interatomic potentials for silicon and carbon,[3] which include angular-dependent forces that are very important in covalent materials with low coordination. Interactions between like and unlike atoms are included in this model.

Progress

We have performed two types of simulations, each designed to look at a different aspect of the problem. One class is designed to simulate orthogonal cutting and to focus on chip formation and mechanisms of plastic flow. The other looks in detail at possible wear mechanisms, such as graphitization and carbide formation, for the tool.

In both types of simulation, the work piece is a large slab containing tens of thousands of atoms oriented with a specific crystal direction face up. Most of the atoms in the work piece move freely according to Newton's laws. Relatively few atoms near the upstream boundary and the lower boundary have additional constraint forces that maintain their temperature at a constant value,[4] allowing heat generated at the tool tip to flow out of the system. Finally, a constant velocity boundary condition is imposed on the lowest atoms in the slab. Atoms leaving the simulation cell at the 'downstream' end are destroyed, and new ones are periodically produced at the 'upstream' boundary. Performing the calculation in the rest frame of the

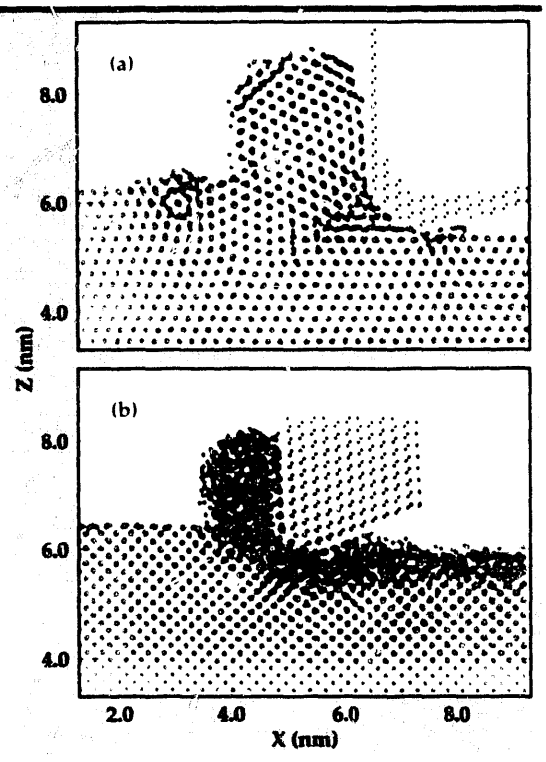


Figure 1. Contrasting behavior of copper and silicon under orthogonal cutting: (a) copper chip remains crystalline but reorients to slip along the easy (111) plane; (b) silicon amorphizes and then 'flows.'

tool allows the simulation of cutting over lengths that are many times the computational cell dimension, without having to follow the motion of a prohibitively large number of atoms.

Orthogonal Cutting

We simulated diamond turning in the orthogonal cutting geometry by creating a wedge-shaped tool with a close-packed (111) cutting face, and imposing periodic boundary conditions in the direction parallel to the surface and normal to the cutting direction. In the case of copper, the diamond tool comprised a rigid array of atoms with about a 2 nm radius of curvature. The work piece contained 36,000 copper atoms with the (111) face up, and moved under the tool at a speed of about 100 m/s from left to right. A cross-sectional 'snapshot' of the simulation is shown in **Fig. 1a**. From this picture, we notice that the chip has remained crystalline, but it has been reoriented to form a (111) slip plane in the primary shear zone in front of the tool.

In contrast to the copper simulation, the silicon calculation allows the lower atoms in the tool to move according to their force laws, and only the uppermost atoms are held rigid. Atoms just below the rigid layers are maintained at constant temperature. The work piece consisted of 20,160 atoms with the (001) plane face up, moving at about 540 m/s, left to right. A cross-sectional 'snapshot' of this simulation is shown in **Fig. 1b**. The first thing to notice is that both the chip and the cut surface are amorphous. In addition, there appears to be a boundary layer of silicon clinging quite tightly to both the rake and clearance faces of the tool.

Tool Wear

Tool wear was simulated by suspending two small carbon asperities from a flat diamond surface and observing their interaction with the silicon work piece. The asperities differed in size, but were both shaped as square pyramids with the four triangular faces being (111) surfaces. The square base of the larger pyramid contained 64(= 8×8) atoms, while the base of the smaller one contained 36(= 6×6) atoms. All of the atoms in both asperities were free to move as Newton's equations dictate. The bases of the pyramids were (001) planes attached to the bottom (001) plane of rectangular diamond slab, four atomic layers thick. The atoms in the bottom two layers of the slab also moved according to Newton's equations, but their temperature was controlled. The atoms in the top two layers were kept in a rigid lattice that initially moved downward at a constant velocity, but stopped after the desired penetration was obtained. After that time, these atoms were held fixed in space. Soon after the asperities made contact with the silicon, the atoms in their tips began to break away, and some were replaced by silicon.[5] Later in the simulation, a graphitic cluster of six carbon atoms appeared at the surface on the downstream side of each asperity. No other damage to the asperities, except for a build-up of silicon on the pyramid faces, was visible during the simulation time of about 10 ps.

The central result of this work is the contrasting behavior of our prototype materials, copper and silicon, under orthogonal cutting. Copper forms a face-centered-cubic (fcc) crystal with a single-atom basis. As a result, slip along the close-packed (111) planes is analogous to sliding stacks of marbles over each other, and as seen in **Fig. 1a**, the copper chip remains crystalline, but reorients and slips along the easy plane. In contrast, silicon forms a diamond lattice that is also fcc, but contains a two-atom basis. Consequently, sliding along the (111) plane is hindered by the strong angular forces, and **Fig. 1b** shows that silicon amorphizes and then 'flows.' This suggests that the surface selects the state that minimizes the work done by the tool.

Our simulation of the wear of small diamond asperities while cutting silicon showed evidence

of both 'graphitization' and carbide formation. Six carbon atoms broke off the asperities and formed hexagonal rings, while silicon atoms filled the resulting vacancies by bonding strongly to the diamond.

Future Work

The ability to understand and control the ductile-brittle transition in glass is critical to improving the economic viability of the state-of-the-art machining capabilities being developed at Lawrence Livermore National Laboratory (LLNL). Our next objective is to define the mechanisms of microplasticity and damage initiation in fused silica by using MD techniques, to follow changes in the structural properties and the dynamic interactions of the atomistic glass network. We hope that an explicit demonstration of the ability to model these processes will greatly enhance the competitiveness of LLNL's materials fabrication efforts within the Department of Energy and elsewhere.

1. J. Belak and I.F. Stowers, "Molecular Dynamics Modeling of Surface Indentation and Metal Cutting," *Engineering Research, Development, and Technology*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-91, 4-3 (1992).
2. D.J. Oh and R.A. Johnson, "Embedded Atom Method for Close-Packed Metals," *Atomistic Simulation of Materials: Beyond Pair Potentials*, V. Vitek and D.J. Srolovitz (Eds.), Plenum Press (New York), 233 (1989).
3. J. Tersoff, *Phys. Rev. B* **39**, 5566 (1989).
4. W.G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).
5. D.B. Boercker, J. Belak, I.F. Stowers, R.R. Donaldson, and W.J. Siekhaus, "Simulation of Diamond Turning of Silicon Surfaces," *Proc. ASPE 1992 Annual Meeting* (Grenelefe, Florida), 45 (October 18–23, 1992).

Materials Science and Engineering

The objective of the Materials Science and Engineering thrust area is to enhance our understanding of the physical and mechanical behavior and the processing/structure/property correlations for structural materials that are of interest to Lawrence Livermore National Laboratory (LLNL) programs and U.S. industry. We also seek to enhance our ability to model the processing of these materials using LLNL's finite element codes. Our activities are currently focused on composite materials, superplasticity, and process modeling.

Composite Materials

Our work in composite materials is directed toward polymer matrix composites and metal matrix composites. LLNL has a long history of achievements in the investigation of polymer matrix composites. These materials have received considerable research attention and product application at LLNL and in industry because of their unique properties, including high specific strength, high specific stiffness, composition of low Z atoms, corrosion resistance, and the possibility for a low coefficient of thermal expansion. These properties can also be tailored to specific applications. During FY-92, we have focused on studying the three-dimensional mechanical response of continuous fiber, polymer matrix composites. These studies have significantly enhanced our understanding of the response of these materials and our ability to test and model this behavior using finite element codes. During FY-92 we have also studied the use of laser Raman spectroscopy to monitor *in situ* the state of cure in polymer matrix composites.

We also work on metal matrix composites (MMC's), which are materials of choice in applications requiring high specific strength and stiffness. These materials can have excellent thermal and electrical conductivity and, depending on the alloy matrix, excellent high-temperature behavior. During FY-92, we studied the processing/structure/property correlations in a unique form of MMC, called a laminated metal composite, in which alternating metallic layers are press-bonded together.

Superplastic Materials

Superplastic materials are crystalline solids that can be deformed in tension to such an extent that large strains will be attained at very low flow stresses. These materials, which deform like hot glass, permit components to be formed into shapes, the dimensions of which are very close to those desired in the final product ('net shape processing'). Thus, machining and machining-related operations can be reduced or eliminated. Our work in this technology has been stimulated by U.S. industry, which has demonstrated a strong interest in superplasticity for net shape processing. Currently, LLNL is engaged in two collaborative research and development projects with industry in the area of superplasticity. One project, with three partners, is developing the technology for commercial production of superplastic, ultra-high-carbon steels. Another project is developing a superplastic aluminum alloy with a faster forming rate and the capability for diffusion bonding. Su-

perplastic forming can also reduce environmental, safety, and health problems in the Department of Energy nuclear weapons complex through the reduction of toxic and radioactive scrap produced during the fabrication of components. This year the thrust area has been studying the microstructural changes that take place during superplastic deformation. A model is being developed for use in LLNL's finite element codes that will account for the influence of material microstructure and its evolution on the stress-strain-strain rate behavior of superplastic materials.

Process Modeling

Our work in process modeling is inspired by the enormous impact of this technology on economic manufacturing competitiveness and by the unique opportunities for LLNL to assist industry, given its extensive experience with modeling problems and its extensive computational resources and codes. Researchers within this thrust area are enhancing the capability of LLNL codes to model the casting process. Casting is a common industrial manufacturing process that is also very complex and influenced by many process and component variables. For these reasons, finite element modeling is a very powerful tool for understanding and predicting the success or failure of industrial casting operations. Work is continuing on a unique fluid-thermal-stress finite element code that will predict the final shape and stress state of precision cast parts.

Donald R. Lesuer
Thrust Area Leader

Section 6

6. Materials Science and Engineering

Overview

Donald R. Lesuer, Thrust Area Leader

Processing and Characterization of Laminated Metal Composites

Chol K. Syn, Donald R. Lesuer, and O.D. Sherby **6-1**

Casting Process Modeling

Arthur B. Shapiro **6-7**

Characterizing the Failure of Composite Materials

Scott E. Groves, Roberto J. Sanchez, William W. Feng, Albert E. Brown, Steven J. DeTeresa, and Richard E. Lyon **6-11**

Fiber-Optic Raman Spectroscopy for Cure Monitoring of Advanced Polymer Composites

Richard E. Lyon, Thomas M. Vess, S. Michael Angel, and M.L. Myrick **6-17**

Modeling Superplastic Materials

Donald R. Lesuer, Chol K. Syn, Charles S. Preuss, and Peter J. Raboin **6-23**

Processing and Characterization of Laminated Metal Composites

Chol K. Syn and Donald R. Lesuer
Engineering Sciences
Mechanical Engineering

O.D. Sherby
Department of Materials Science and Engineering
Stanford University
Palo Alto, California

We have made laminated metal composites of (1) ultrahigh carbon steel (1.8% C) and brass (70% Cu-30% Zn), and (2) Al 5182 and Al 6061-25 vol % SiC_p. The laminates were prepared by hot pressing alternating layers of the component materials in an argon gas atmosphere. Tensile and fracture toughness were measured for different processing conditions of surface oxide descaling, layer thickness, and heat treatment. Descaling of the surface oxide prior to the press-bonding was found to eliminate premature delamination along interfaces, resulting in an increased yield strength and tensile ductility. Reduction in the layer thickness brought a large increase in tensile ductility, and a small decrease in yield strength and fracture toughness. T6 heat treatment on the Al laminates induced a substantial increase in the yield and tensile strength, but a decrease in tensile ductility. Fracture toughness measured both in the crack-arrester and crack-divider orientations showed a large enhancement over that of the component materials. Damping capacity measurements also showed rather remarkable increases over that of the component materials.

Introduction

The idea of laminating different metals and alloys to form a composite material that exploits the good properties of the constituent materials has been known from antiquity: *The Illiad* of Homer, e.g., describes Achilles' shield, made of two outer layers each of bronze and tin and one middle layer of gold.[1] The idea has also been used in many industrial applications.[2] However, most of the current industrial metal-based laminates contain only two or three layers, and are used to save material cost while maintaining required wear or corrosion resistance. Recent studies[3,4,5] show that multilayer laminated metal composites (LMC's) can have superior damage-critical properties such as fracture toughness and fatigue resistance, over that of the component materials. Damage crack propagation in a laminate of dissimilar materials is inherently difficult, since an interface can act as a barrier to the crack propagation, especially when such an interface delaminates at the crack tip and blunts the crack.[6] Studies[3,4,5] also show that it is possible to design a LMC with given performance character-

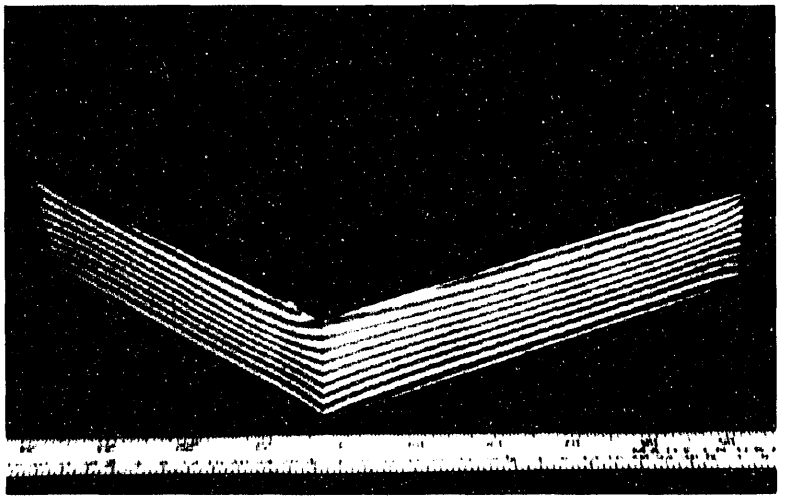


Figure 1. An example of as-pressed Al/Al-SiC_p laminate after the edges were trimmed.

Table 1. Tensile properties of Al 5182/AL 6061-25 vol% SiC_p

Processing	Layer Thickness, μm	Yield Strength, MPa(ksi)	Ultimate Tensile Strength, MPa(ksi)	Elongation, %
Scale Not Removed	750	138(20.0)	262(38.0)	10.0
Descaled	750	162(23.5)	266(38.5)	16.9
Descaled, T6	750	232(33.6)	333(48.2)	7.2
Descaled, Re-pressed, T6	100	201(29.2)	324(47.0)	12.0

istics, through the choice of component materials, number of layers, thickness of the layers, and interfacial bond strength. A good example is the LMC formed by press-bonding alternate layers of ultra-high carbon steel (UHCS) and mild steel. In this LMC, the dynamic fracture toughness is far higher than that of either of the component materials. LMC's can also have damping capacity superior to that of the component materials, which can be very useful in structures requiring high acoustic damping.

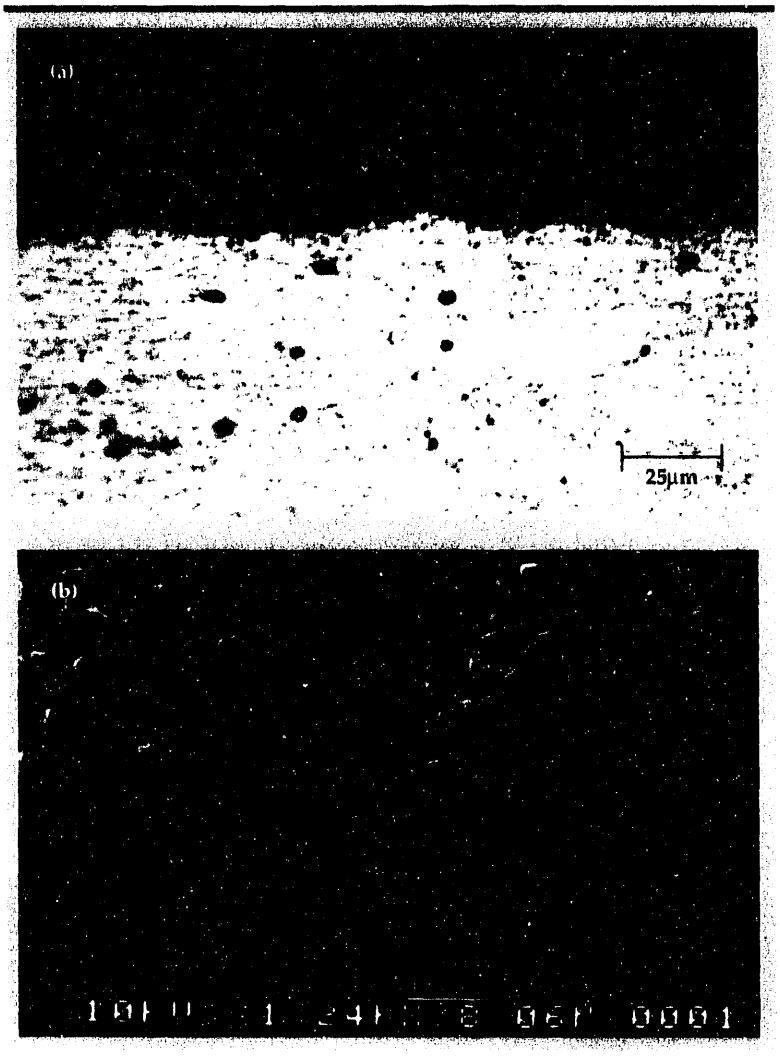


Figure 2. (a) Optical and (b) scanning electron microscope microstructure in the vicinity of an interface in Al/Al-SiC_p laminate.

Progress

We initiated the present research in FY-91[7] to investigate the influence of processing and structural variables on the mechanical properties of multilayer LMC's made of two constituent materials, one ductile but tough and the other brittle but strong. We chose in FY-91 to study two LMC systems, UHCS/brass and Al/Al-SiC_p, and continued this study in FY-92. These two systems were chosen to show that the toughness at ambient temperature of hard and brittle UHCS and Al-SiC_p can be enhanced substantially by lamination with ductile but tough counterparts. The main thrust for FY-92 was to study the influence of the surface preparation of the component materials and the layer thickness of the laminates on their interfacial microstructure and mechanical properties.

Experimental Procedure

Materials and Processing. UHCS of a nominal composition of Fe-1.8% C-1.65% Al-1.5% Cr-0.5% Mn was preprocessed[8] to have a fine-grained ferrite matrix of about 0.5-μm grain size and spheroidized iron carbides on grain boundaries. Brass (70% Cu-30% Zn), Al 5182 (Al-4.5% Mg-0.35% Mn), and Al-SiC_p [Al 6061 (Al-1.0% Mg-0.6% Si-0.28% Cu-0.2% Cr) matrix with 25 vol.% SiC particulate] were obtained from commercial sources.

All materials were sliced to 50 mm-x-50 mm squares. Al 5182 and Al 6061-SiC_p squares were descaled using an acid solution. UHCS and brass were surface-machined and degreased. A lami-

nate containing an equal volume fraction of the two component materials was prepared by hot-pressing a stack of alternate layers of the components. Each stack was pressed to one third or one fourth of its original height. Such a large reduction ensured good bonding at interfaces. UHCS laminates were press-bonded at 750°C, and Al laminates at 450°C. Some laminates were sliced into four equal-sized pieces, re-stacked, and re-pressed to obtain laminates with reduced layer thickness. Average layer thickness was about 750 μm for both laminates after initial pressing; about 200 μm for UHCS and 100 μm for Al laminates after re-pressing. Some Al laminates were given the T6 heat treatment for the Al 6061 matrix of the Al-SiC_p component.

Testing. Tensile tests were performed with flat specimens cut with the tensile axis parallel to the layers. Fracture toughness was measured with chevron-notched short bar or three-point bend bar specimens in which the notch was cut either in the crack arrester or crack divider orientation. In the crack arrester orientation, the crack front propagates in the thickness direction, cutting the layers sequentially. In the crack divider orientation, the crack front propagates through the laminate, cutting all the layers simultaneously. Damping capacity along the thickness direction of the laminates was evaluated by a pulse-echo method for the ultrasonic frequency range, and by a torsion bar technique for the 0.1 to 100 Hz range.

Experimental Results

Interfacial Bonding and Microstructure. The Al laminates that were chemically descaled and the UHCS laminates whose layers were surface-machined prior to the press-bonding were well bonded and did not show any sign of interfacial delamination during machining of test specimens. A typical well-bonded as-pressed Al laminate, after its edges were trimmed, is shown in **Fig. 1**. **Figure 2** shows an interface in the Al laminate shown in **Fig. 1**, both in (a) optical and (b) scanning electron microscope photomicrographs. No interfacial pores or unbonded areas, and no secondary phases are visible, indicating that no reaction between the component materials occurred.

Tensile Properties. Tensile properties of UHCS laminates were included in our FY-91 report.[7] Summarized in **Table 1** and **Fig. 3** are the tensile properties for the Al laminates in the as-pressed, and T6 heat-treated conditions. For the T6 heat-treated condition, laminates with two different average layer thicknesses (750 μm and 100 μm) are compared: the 750-μm-layer thickness material was obtained from the initial pressings, and the 100-μm-layer thickness material was obtained from re-pressings as described earlier.

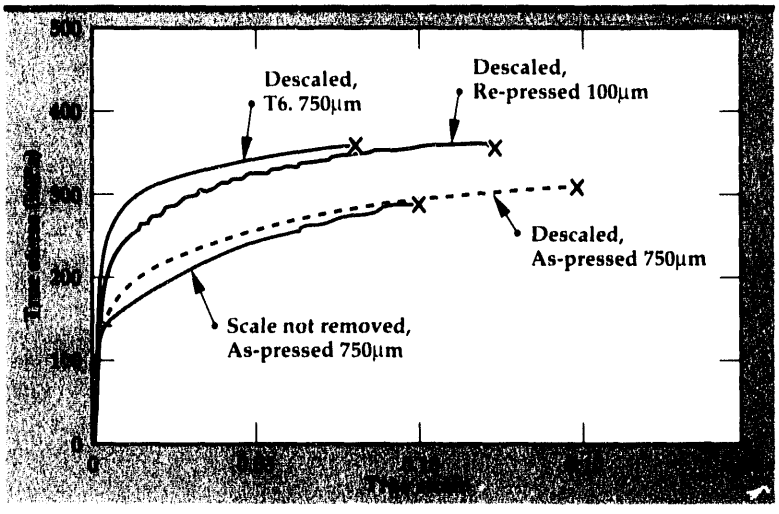


Figure 3. Tensile stress-strain behavior of Al/Al-SiC_p laminate. Results shows the influence of surface descaling, heat treatment, and layer thickness.

Table 1 clearly shows the effect of surface oxide removal for two 750-μm-layer laminates. Descaling of the constituent materials led to a noticeable increase in the yield strength, from 138 MPa (20 ksi) to 162 MPa (23.5 ksi), and to a very substantial increase in ductility (by almost 7%), from 10% to 16.9%. No significant change in the ultimate tensile strength was observed. **Figure 3** shows that the descaling treatment increases the flow stress over the entire strain range, most likely as a result of good bonding between the constituent layers, which prevents premature delamination. The importance of preventing delamination can probably be traced to the fact that in these materials, flow localization precedes fracture. Good bonding inhibits flow localization in the less ductile layers, which in turn results in greater elongation (and higher strength) before fracture.

Heat treatment considerably influences the mechanical properties, as shown by the results in **Table 1** for 750 μm-layer laminates. The T6 treatment increased the yield and ultimate tensile strength by about 70 MPa (10 ksi), but reduced the total elongation drastically, from about 17% to 7%. **Figure 3** shows that the flow stress was also increased.

Reduction of the layer thickness affects the tensile properties significantly. When the layer thickness was reduced from 750 μm to about 100 μm under the T6 heat-treated condition, the yield strength was decreased slightly, but the total elongation was increased rather remarkably, from 7.2% to 12%. No significant change in the ultimate tensile strength was observed. A similar strong corre-

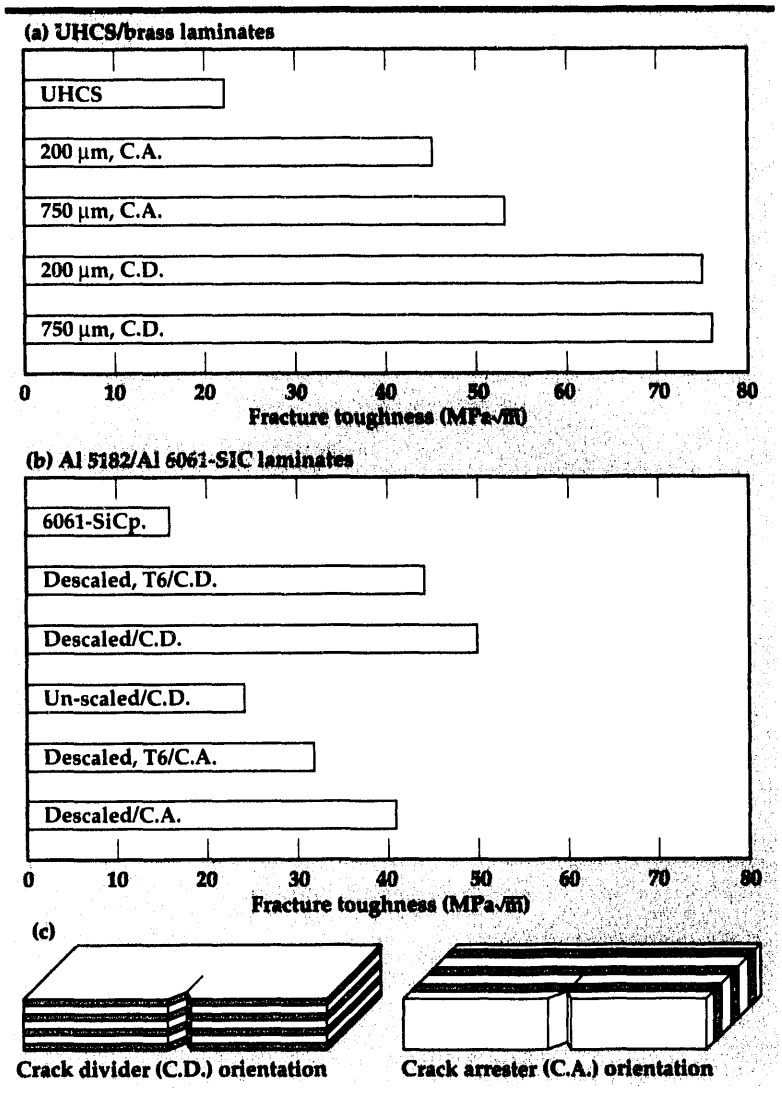


Figure 4. Fracture toughness measured (a) for the different layer thicknesses in UHCS/brass laminates and (b) for different processing conditions in Al/Al-SiC$_P$ laminates. The inset drawings (c) define the crack divider (C.D.) and crack arrester (C.A.) orientations used in the fracture toughness tests.

lation between the layer thickness and ductility has been observed in the UHCS laminates, as reported in FY-91.[7]

Fracture Toughness. Results of fracture toughness tests are summarized in **Fig. 4** for both UHCS and Al laminates. For laminates of both systems, it is clearly demonstrated in **Fig. 4** that the lamination of a hard material (UHCS or Al 6061-SiC$_P$) with a ductile material (brass or Al 5182) results in substantial enhancement of toughness.

For UHCS/brass, the laminates with thin (200 μm) layers show slightly lower toughness than the laminates with thick (750 μm) layers regardless of the specimen orientation, i.e., crack arrester or crack divider, relative to the layers. This trend could be due to the influence of the interfacial delamination on crack growth. Studies have shown that the delamination tendency decreases with decreasing layer thickness.[9] Thus, it is likely that the higher toughness in the thick-layer laminates could be due to the blunting of an advancing crack by delamination. It is interesting to note, however, that delamination reduces the tensile ductility and strength, as observed for the tensile properties.

Interfacial delamination was observed also in the Al laminates regardless of the descaling or T6 treatment. The descaling treatment, however, led to an increased fracture toughness (measured in the crack divider orientation), as shown in **Fig. 4**, while the T6 treatment led to a reduced toughness (as measured both in the crack arrester and divider orientations). The beneficial effect of the descaling treatment results from the controlled and timely delamination of an interface as a crack approaches the interface. In the laminates made without the descaling treatment, delamination was extensive and occurred rather prematurely. The increased yield strength and reduced tensile ductility upon T6 treatment, as shown in **Fig. 3**, were reflected in the reduced fracture toughness, a trend observed similarly in most monolithic materials.

Damping Capacity. Damping capacity was measured only in the as-pressed condition, where the layer thickness was 750 μm for both UHCS/brass and Al laminates. At low frequencies, damping in the UHCS/brass laminate was two to three times the damping typically observed in brass or steel, and was lowest at 2 Hz. Ultrasonic attenuation measurements of longitudinal waves showed that at 2.25 MHz, the UHCS/brass laminate had an attenuation coefficient of 160 dB/m, over 12 times the attenuation coefficient for the steel component and over four times the attenuation coefficient of the brass component.[10] The ultrasonic attenuation coefficient for the Al/Al-SiC$_P$ laminate (266 dB/m) was greater than that for the UHCS/brass laminate. These results clearly show that LMC's can be more effective damping materials than their components.

Future Work

We are continuing to characterize the UHCS/brass and Al/Al-SiC laminates regarding their (1) damping capacity in the audible frequency range, (2) fatigue behavior, (3) response to ballistic impact, and (4) deformation behavior at elevated temperature. We are planning to make LMC's of other light materials such as Mg alloys, (a) containing a high damping capacity material as a component and (b) containing an intermetallic or superalloy as a component. These new lami-

nates will be tested for their strength, ductility, toughness, and other characteristics reported here for the UHCS/brass and Al laminates.

Acknowledgements

We sincerely appreciate assistance provided by Chris Steffani for descaling Al alloys; Ralph Otto and Bill Stutler for pressing laminates; Dick Sites for preparing test samples; Al Shields for conducting mechanical property tests; and Jim Ferreira for metallography.

1. *The Iliad of Homer*, translated by R. Lattimore, University of Chicago Press (Chicago, Illinois), 411 (lines 270–272), 1951.

2. E.S. Wright and A.P. Levitt, "Laminated Metal Composites," *Metallic Matrix Composites*, K.G. Kreider (Ed.), Academic Press (New York, New York), 37, 1974.

3. C.K. Syn, D.R. Lesuer, K.L. Cadwell, O.D. Sherby, and K. Brown, "Laminated Metal Composites of Ultrahigh Carbon Steel/Brass and Al/Al-SiC: Processing and Properties," *Proc. Conf. Developments in Ceramic and Metal-Matrix Composites*, K. Upadhya (Ed.), TMS, 311, 1991.

4. D.W. Kum, T. Oyama, J. Wadsworth, and O.D. Sherby, *J. Mech. Phys. Solids*, **31**, 173 (1983).

5. O.D. Sherby, S. Lee, R. Koch, T. Sumi, and J. Wolfenstine, *Materials and Manufacturing Processes*, **5**, 363 (1990).

6. J. Cook and J.E. Gordon, *Proc. Royal Soc. London*, **282**, 508 (1964).

7. C.K. Syn, D.R. Lesuer, K.L. Cadwell, K.R. Brown, and O.D Sherby, "Processing and Testing of Metal Composites of Ultrahigh Carbon Steel/Brass Laminates and Aluminum Laminates," *Engineering Research, Development, and Technology*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-91 (1992).

8. O.D. Sherby, T. Oyama, D.W. Kum, B. Walser, and J. Wadsworth, *J. Metals*, **37**, 50 (1985).

9. C.K. Syn, D.R. Lesuer, J. Wolfenstine, and O.D. Sherby, *"Layer Thickness Effect on Ductile Tensile Fracture of Ultrahigh Carbon Steel–Brass Laminates,"* Livermore National Laboratory, Livermore, California, UCRL-JC-110413 (1992), accepted for publication in *Metall. Trans.*, TMS.

10. B.P. Bonner, D.R. Lesuer, C.K. Syn, and O.D. Sherby, "Damping Measurements for Ultrahigh Carbon Steel/Brass Laminates," *Proc. Symp. Damping of Multiphase Inorganic Materials* (Chicago, Illinois), R. Bhagat (Ed.) (November 1–5, 1992); to be published by ASM International.

Casting Process Modeling

Arthur B. Shapiro
Nuclear Test Engineering
Mechanical Engineering

In predicting the quality of a cast part, two important factors are (1) correct modeling of the fluid flow and heat transfer during the filling of a mold with a molten metal, and (2) the thermal-mechanical physics of solidification and cool-down. Determining the dynamics of the flow and the free surface shape during filling are essential in establishing the temperature gradients in the melt and in the mold. Correctly modeling the physics of volume change on solidification, shrinkage on cooling, and contact resistance across the part-mold interface directly affects the cooling rate and, ultimately, the final cast shape and stress state of the cast part. This year our efforts were focused on modeling fluid fill and on the physics of solidification.

Introduction

Casting manufacturing covers a broad range, from the large tonnage of continuously cast steel products, through the intermediate-weight output of superalloy precision die castings, to the relatively small quantity of high-purity crystals. Although this project benefits modeling efforts in each of these three casting areas, we have focused on modeling precision die castings of superalloy parts.

Our approach to casting process modeling is to use the computational fluid dynamics code ProCAST[1] to model elements of the mold-filling process, including tracking of (1) the free surface of the molten metal as it rapidly fills the mold; (2) solidification on the walls of the mold; (3) temperature transients in the mold; and (4) temperature transients in the liquid and solidifying metal. When the mold is completely filled with liquid metal, the existing temperature field at that instant in time is re-mapped (using REMAP[2]) onto a new mesh for a CAST2D[3] analysis to predict the final cast shape, stress state, and defects.

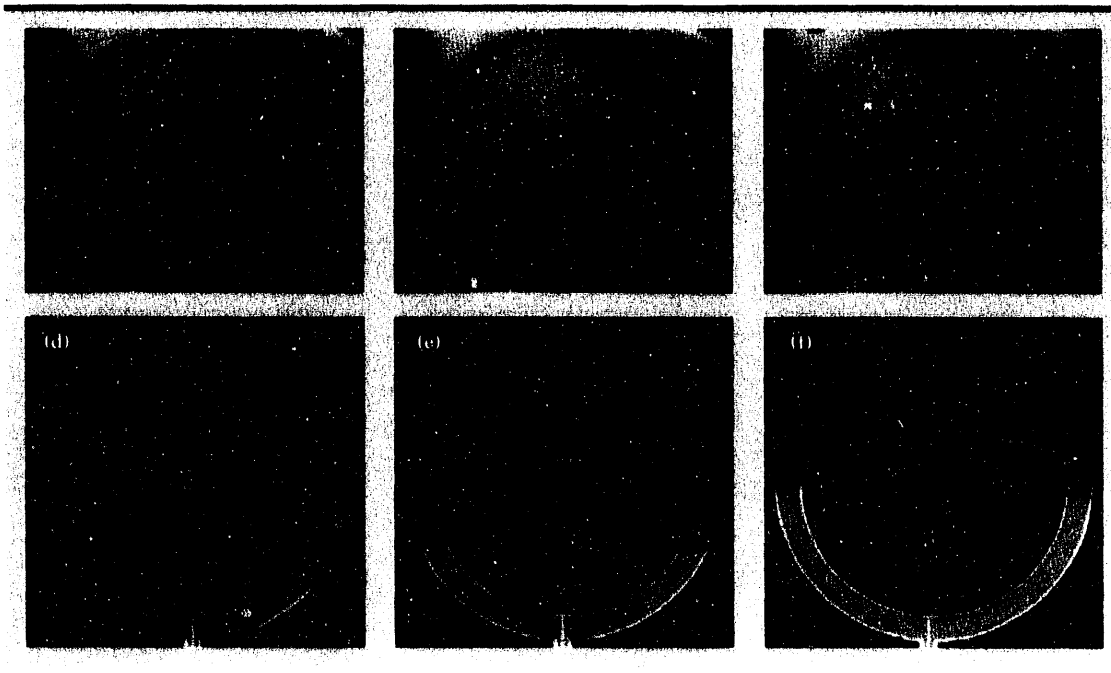


Figure 1. Experimental (a, b, c) and numerical (d, e, f) results for the filling of a spherical annulus mold, with a liquid metal at 0.6, 0.9, and 1.5 s, respectively.

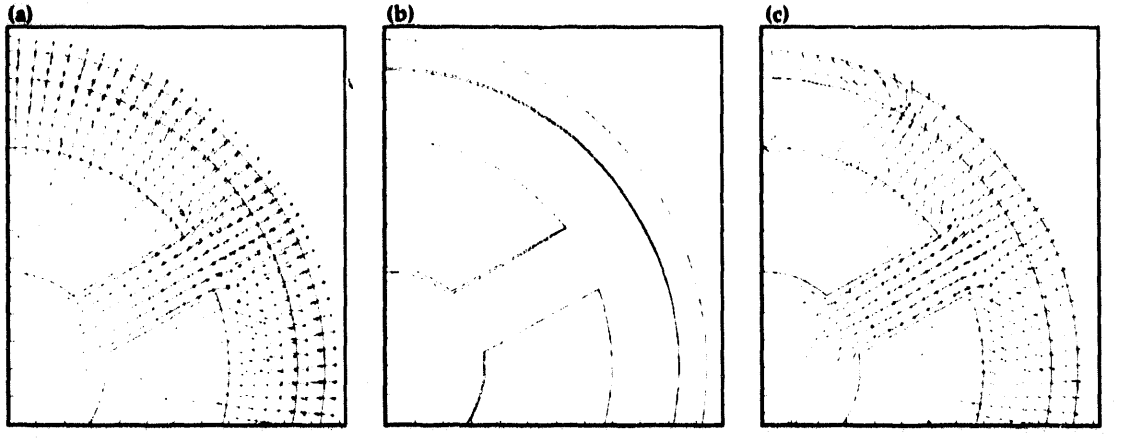


Figure 2. (a) CAST2D calculations at an early time during cool-down, showing the heat flow paths to be radially outward. (b) At later times: gaps between the casting and mold appear due to the -6.6% volume change of the aluminum casting on solidification and shrinkage on cool-down. (c) The direction of the heat-flux vectors, changed due to the part shrinking away from the mold on solidification. The heat-flux vectors are seeking the path of least resistance to heat flow.

CAST2D models the thermal-mechanical response during cooling to room temperature including volume change on phase transformation. CAST2D also calculates thermal contact resistance across the part/mold interface.

Progress

An experimental and numerical analysis was performed to investigate the filling of a spherical mold with a liquid metal. The experiment provided visual data of the filling of the shell for comparison with the numerical calculations. The die-casting process consisted of pressurizing a pool of molten metal in a crucible and forcing the melt up a small tube to be injected into the bottom of a spherical annulus-shaped cavity. The experiment was conducted in a vacuum. The experimental data consists of a 16-mm motion picture (24 frames per second) of radiographs of the filling of the spherical annulus. The spherical annulus filled in approximately four seconds. **Figure 1** shows radiographs at three different times during the filling process. The radiographs show that initially the melt splashes fairly high up the annulus (**Fig. 1a**) in a nonsymmetric fashion. Although still showing considerable wave motion at later times (**Fig. 1b**), the melt is seen to be filling the annulus in a more symmetric fashion. At still later times, the fill level advances evenly and symmetrically up the annulus (**Fig. 1c**).

ProCAST was used to numerically model the mold filling process. The spherical annulus was modeled as a two-dimensional, axisymmetric problem. Fluid properties, i.e., density, specific heat, and viscosity, were allowed to be functions of temperature. Results of the analysis are presented in **Figs. 1d, 1e,** and **1f**. These three figures show the free surface and level of fill at the same times as the experimental results. The numerical analysis does not show the wave motion and non-symmetry at early times as observed in the experiment. However, it does show the wall jet effect as the melt enters the annulus (**Fig. 1d**) and a fill level that is near the average level of the experiment (**Fig. 1e**). At later times, when the flow is more even and symmetric, the analysis compares favorably with experiment (**Fig. 1f**).

We performed a fluid-thermal-mechanical analysis of the casting of a three-spoke, 38-cm-dia aluminum wheel. The liquid aluminum at 780°C was injected into the steel mold, which is heated to 730°C, at a rate that fills the mold in two seconds. By forced convection, the outer surface of the mold lost heat to the environment. ProCAST was used to model the fluid-filling process. When the mold was completely filled with liquid metal, the existing temperature field at that instant in time was re-mapped onto a new mesh for a CAST2D analysis. CAST2D was used to model the thermal-mechanical response during cooling to room temperature. The aluminum undergoes phase change at 660°C with a volume change of -6.6%. The aluminum part is in good contact with the mold at early time. The heat flux vectors are seen to go radially outward (**Fig. 2a**) through the aluminum part and mold to the environment. At a later time, the aluminum has shrunk away from the mold due to volume shrinkage on solidification, and gaps have opened up (**Fig. 2b**). The heat-flux paths

shown in **Fig. 2c** are seen to be considerably changed from the pattern of **Fig. 2a**. The direction of the heat-flux vectors has changed while seeking the path of least resistance to heat flow.

Future Work

In the future, we plan to develop a closely coupled fluid-thermal-mechanical code to be used for analysis of casting problems. Numerical modeling in the areas of fluid fill, solidification physics, and material constitutive development must be refined for such a code to be useful in casting process modeling. We also plan to conduct experiments for code validation.

1. *ProCAST™ User's Manual Version 2.0*, UES, Inc., 4401 Dayton-Xenia Road, Dayton, Ohio 45432-1894.
2. A.B. Shapiro, *REMAP—A Computer Code That Transfers Node Information Between Dissimilar Grids*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-ID-104090 (1990).
3. A.B. Shapiro, *CAST2D—A Finite Element Computer Code for Casting Process Modeling*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-MA-108598 (1991).

Characterizing the Failure of Composite Materials

Scott E. Groves,
Roberto J. Sanchez,
William W. Feng, and
Albert E. Brown
Engineering Sciences
Mechanical Engineering

Steven J. DeTeresa
Richard E. Lyon
Materials Division
Chemistry and
Materials Science Department

Our goal for this project has been to characterize the three-dimensional (3-D) performance of continuous-fiber polymer composite materials, by developing new experimental and theoretical methods. This report highlights our major accomplishments: (1) multiaxial testing of composites; (2) the development of a new composite-failure criterion; (3) the development of ORTHO3D, a 3-D orthotropic finite element code; (4) the dynamic testing of composites; and (5) a helical compression study of filament-wound composite tubes.

Introduction

This project has helped to develop carbon-fiber composite materials for use in penetrating warhead cases, gun barrels, advanced munition components, projectiles, nuclear weapons, SDI, and high-energy-density flywheels.

We have had the opportunity to jointly pursue some of this work with various research and development centers of the U.S. Army and Navy, as well as with various contractors in private industry. The primary reason for selecting carbon-fiber composite materials is their high specific strength and stiffness. Other factors influencing design include low 'Z' material composition, low coefficient of thermal expansion, impact resistance, and fire safety.

These materials are generally limited not by their performance capabilities but by our lack of understanding and ability to model their complex 3-D response. Our research effort has made great strides in providing the necessary tools and information to optimize these designs.

Progress

Our major accomplishments have been (1) multiaxial testing of composites; (2) the development of a new composite-failure criterion; (3) the development of ORTHO3D, a 3-D, orthotropic finite element code; (4) the dynamic testing of composites; and (5) a helical compression study of filament-wound composite tubes.

Multiaxial Testing of Composites

Our most important accomplishment has been the development of a unique multiaxial test system

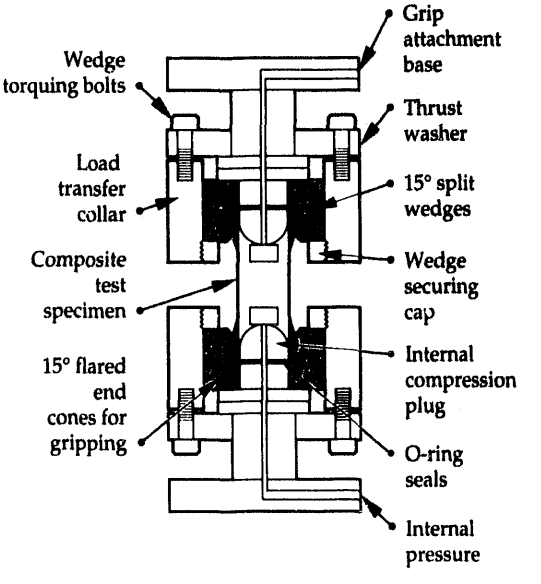


Figure 1. Multiaxial test system for composite tubes.

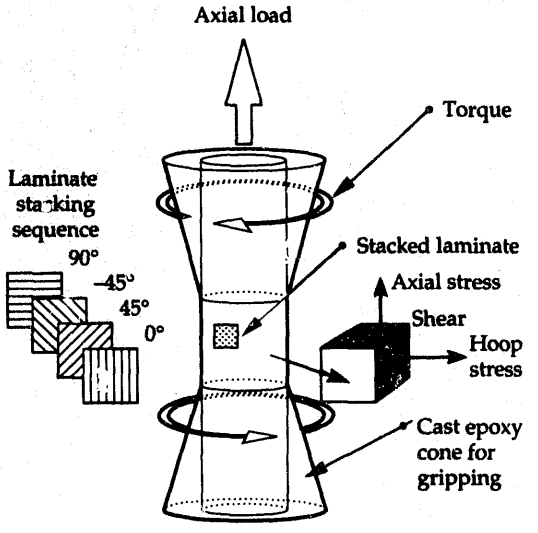


Figure 2. Load description for multiaxial test specimen.

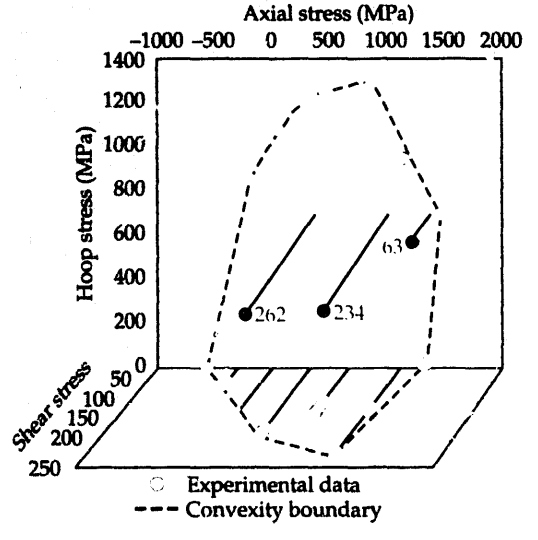


Figure 3. Multiaxial failure surface for a filament-wound Toray 1000/DER332-T403 [±1.5, ±45, ±89] carbon/epoxy laminate.

for polymer composite materials.[1] This technique allows testing of 2-in.-dia composite tubes under a combination of axial, torsion, and internal pressurization. The unique advancement with this system is the simple but effective gripping mechanism that incorporates a 15°-potted epoxy cone for providing a smooth transition in load between the grip and the test specimen. The test specimen itself is a straight-walled composite tube. Diagrams of the multiaxial grip and test specimen are shown in **Figs. 1** and **2**. This test system has provided Lawrence Livermore National Laboratory (LLNL) with an unrivaled capability to generate multiaxial failure data for polymer composite materials.

The most extensive multiaxial failure surface that was generated with this system is shown in **Fig. 3**. The material system used in this study consists of the Toray 1000G carbon fiber having a 950-ksi tensile strength, impregnated with DER332-T403 epoxy having a tensile strength of 10 ksi. This material system has been used in the majority of the composite structures we have designed and fabricated.

The multiaxial gripping concept has lead to the development of an efficient high-strength shear joint for split composite pressure vessels and for modular composite gun barrels. We have also successfully scaled up the gripping system to test 9-in.-dia composite tubes under axial load. The biggest use of this system, however, has been focused on the optimization of the compression performance of filament-wound tubular structures, such as those used for composite penetrators, projectiles, and support structures.

New Composite-Failure Criterion

In this project, we have developed a new failure criterion for composite materials, the 'Feng failure criterion.'[2] The failure criterion is written in terms of the strain invariants in finite elasticity. These invariants are written as functions of the Cauchy strains and the deformation gradients. Among these strain invariants, two are functions of the fiber orientation, and three are not. Therefore, the failure criterion can be further divided into two modes, the fiber-dominated failure mode and the matrix-dominated failure mode. The criterion contains five failure material constants for infinitesimal, general, 3-D strain states.

In the criterion, there are three quantities governing the failure surface in composites: the distortional energy, the dilational energy, and the difference between compressive and tensile strengths. The minimum number of constants required is three for each failure mode. Therefore, this failure criterion represents the minimum number of constants required for determining the failure surface of composites for the second-order strain-failure criterion.

We have previously obtained the unidirectional lamina failure surface for Toray 1000/DER332-T403 carbon/epoxy fiber composites.[3] In this project, we have obtained experimentally the fundamental material properties for this unidirectional composite, both elastic constants and strength. The corresponding strengths obtained by the Feng failure criterion for symmetrically balanced angle-ply laminates, is shown in **Fig. 4**. The results show that the Feng criterion predicts the fiber- and matrix-dominated failure modes. Furthermore, for Toray 1000/DER332 symmetrically balanced angle-ply laminates subjected to uniaxial load, the initial failure consistently initiates in the matrix.

Development of ORTHO3D

During the second year of this program, we began an effort to develop a simple, 3-D, orthotropic finite element program for the evaluation of composite failure criteria. This algorithm has become known as ORTHO3D and has been developed under a university contract with Texas A&M University.[4] In all, five failure criteria have been implemented: maximum stress, maximum strain, Tsai-Wu, Hashin, and the Feng failure criterion.

To perform accurate failure modeling of a laminated composite structure, ORTHO3D was written to permit detailed sub-lamina (single-ply) analysis of generic structures such as cylinders and cubes. These two generic shapes represent characteristic local volumes of larger structures such as penetrator missile cases or thick laminated plates. Performing this level of analysis explicitly with NIKE3D or DYNA3D is too cumbersome to be cost effective. Furthermore, large structural analysis codes like NIKE cannot simultaneously model the global behavior of a structure and model the characteristic local volume in sufficient detail to perform accurate local-failure analysis (local/global modeling).

ORTHO3D, which is written in Fortran, is operational on a variety of computer platforms, including Macintosh, SUN, Vax, and IBM. The size of problems (local volume) that one can solve is limited only by computer memory. Even small local volumes require an astonishingly large amount of memory. Typically, a single ply is 0.005 in. thick, and the local volume is composed of many of these layers. Generally, we recommend a minimum of three elements (8-noded brick elements) through the thickness of each ply to capture representative local stress/strain behavior. Maintaining a respective aspect ratio (< 10) for each element, the number of elements required to model a local volume can be very large. At 8 nodes per element and 3 degrees of freedom per node, the memory requirements can quickly exceed most small computers. Typical local volumes that we have solved require ~ 20 Mb of ram. Efforts over the last year have focused on minimization of memory requirements via more efficient equations solvers, nodal number schemes, and array sharing.

To facilitate the 3-D modeling of composite structures, ORTHO3D was adapted to generate the effective, 3-D, homogenized properties for the characteristic local volumes required by NIKE for analysis of large (global) composite structures. The homogenized 3-D properties are considered more representative of the actual structural behavior of the local volume than those properties predicted by available, 3-D, micromechanical constitutive solutions. Once NIKE solves the structural problem, the local traction set can be passed to ORTHO for detailed failure analysis.

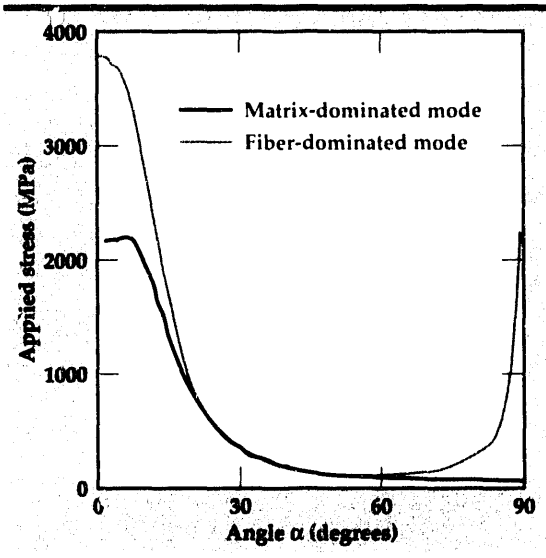


Figure 4. Failure strengths predicted by the finite-strain-invariant failure criterion for Toray 1000/DER332 symmetrically balanced angle-ply laminates.

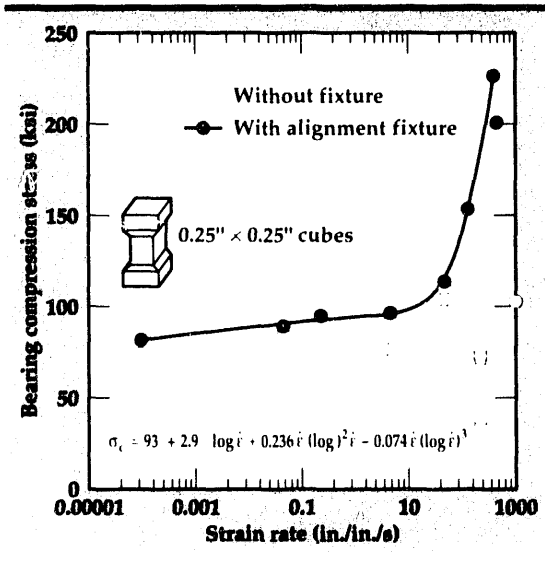


Figure 5. High strain-rate-bearing compression of IM6/974 [0,90] tapered cubes.

Dynamic Testing of Composites

To provide design support of the composite penetrators, gun barrels, and projectiles being developed at LLNL, we have evaluated and developed a variety of new high-strain-rate testing techniques for polymer composite materials.[5] Prior to this investigation, an extremely limited data set was available on the high-strain-rate response of polymer composites. We have successfully generated strain rate data from 0/s to 3000/s in compression and 0/s to 100/s in tension, using a variety of test machines. Our ef-

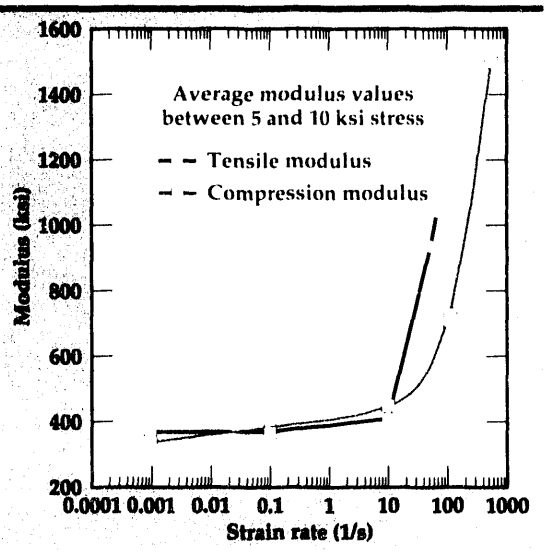


Figure 6. Strain rate sensitivity of the average modulus for DER332-T403 epoxy.

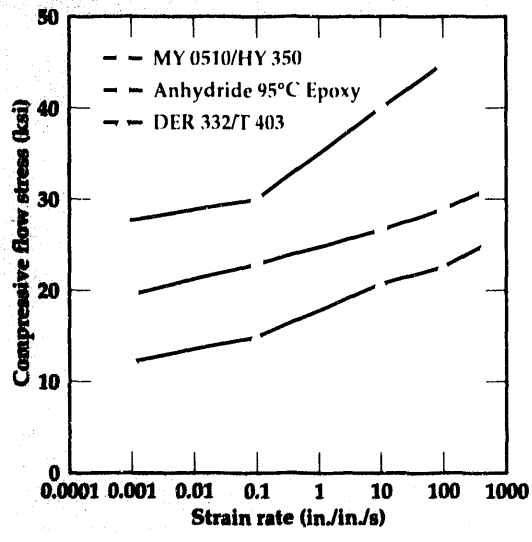


Figure 7. Strain rate sensitivity effects on the compressive flow stress for three epoxy resins.

forts have yielded some very interesting and encouraging material responses.

Most of our efforts focused on developing an acoustically damped, high-energy drop tower for evaluating the high-strain-rate compressive performance of composite materials. Three major advances occurred to the drop tower system that created a highly capable material-evaluation system. First, an acoustically damped base system was installed that eliminated spurious shock waves that were convolved with the quartz load cell data. Second, a high-speed data-acquisition system was developed that significantly automated data gathering and reduction. Third, a precision alignment fixture was developed along with precision-machined test specimens, which resulted in a significant improvement in the measured mean strength of the composite materials as well as minimized data scatter.

Three interesting results from our efforts are presented. All of our experimental results indicate an increase in both strength and modulus with increasing strain rate. First, the bearing compression strength of a [0,90] laminated composite material is presented in **Fig. 4**. This result reveals a significant strengthening occurring at strain rates above 10/s. An examination of neat resin behavior revealed similar trends. **Figure 5** shows the change in neat resin modulus as a function of strain rate. Finally, **Fig. 6** shows strain-rate sensitivity effects on the compressive flow stress of three different epoxy resin systems being evaluated at LLNL.

Helical Compression Study of Filament-Wound Composite Tubes

During the last year of this program, an evaluation of the compressive performance of helically wrapped carbon/epoxy tubes[6] was performed using the three different epoxy resins shown in **Fig. 6**. The objective of this study was to optimize the compression performance of filament-wound composite structures. **Table 1** lists the basic properties of these systems that influence the compression strength of composites. The last two columns in **Table 1** are processing parameters.

The compressive strength of a unidirectional composite material is controlled by the properties of the matrix surrounding the fibers. It has been argued that compression of unidirectional composites is a micro-buckling controlled event in the fibers, and thus dependent on factors such as the local shear modulus of the matrix. What we were hoping to find was an improvement in compression strength of helical composite tubes fabricated with the MY0510-HY350 epoxy system.

Figure 8 shows the variation in axial compression strength for Toray 700 [89, $\pm\theta_2$, -89] helically wrapped composites tubes as a function of helical angle, θ. These tests were conducted using the multiaxial gripping system. The MY0510 system might be considered to be stronger, but the evidence is not conclusive. A problem with polymer composite materials is that it is impossible to isolate single variables. The viscosity of the epoxy

Table 1. Basic epoxy properties.

Epoxy System	G_m (ksi)	$s_{c,flow}$ (ksi)	t_y	T_g (°C)	Viscosity (P)
DER332-T403	160	12	6.5	90	7.8
Anhydride	180	18	8.7	140	3.4
MY0510-HY350	205	28	10.7	180	14

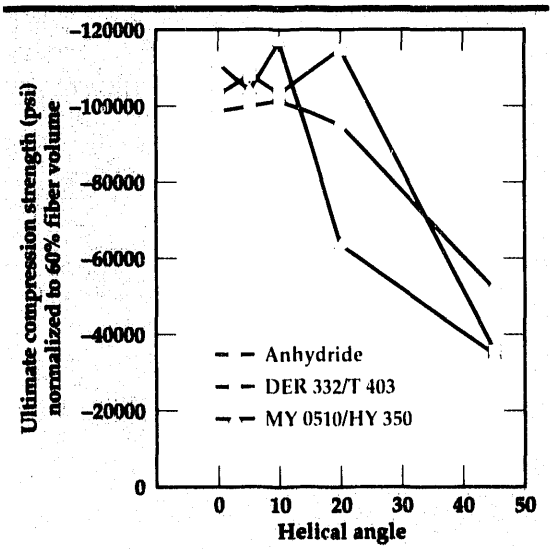


Figure 8. Normalized axial compression strength for T700 [89, ±θ, -89] helically wrapped carbon/epoxy tubes with three different matrices.

turned out to be a major factor as well. The MY0510 system has a much higher viscosity than either the Anhydride or DER332 epoxies, which makes it very difficult to process. This resulted in a composite material with higher void contents and resin-rich areas. In contrast, the Anhydride epoxy system produces very high quality composite materials.

A somewhat surprising result was the small variation in compression strength at helical angle between 0° and 10°. Again, processing influences these results; it is very difficult to achieve uniform part quality for helical winding angles less than 10°. Furthermore, helical angles greater than 10° are much faster (cheaper) to fabricate.

Future Work

We have secured a long-term Cooperative Research and Development Agreement with Boeing Commercial Airplane Group to study the "Strength and Durability of Continuous Fiber Polymer Composites." Natural extensions of our past efforts in "Characterizing the Failure of Composites" will be to modify ORTHO3D to model long-term thermal-viscoelastic effects for these materials. The multiaxial test specimen developed in this program has been selected as a prime test specimen for long-term fatigue testing because of the lack of a free edge associated with the cylindrical test specimen. Finally, we will continue to support the LLNL weapons-related activities in composites, especially in the area of compression optimization, that will make significant use of the multiaxial test specimen.

1. S.E. Groves, R. Sanchez, and W.W. Feng, "Multiaxial Failure Characterization of Composites," *Composites: Design, Manufacture, and Application*, S.W. Tsai and G.S. Springer (Eds.), *Proc. 8th Int. Conf. Composite Materials* (published by SAMPE), (July 1991).
2. W.W. Feng and S.E. Groves, *On the Finite Strain Invariant Failure Criterion for Composites*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-104825 (1990); accepted for publication in *J. Compos. Mater.*
3. W.W. Feng and S.E. Groves, *J. Advanced Composites Letters* **1** (1), 6 (1992).
4. M.A. Zocher, D.H. Allen, and S.E. Groves, "Predicted Stiffness Loss Due to Delamination in Filament Wound Composite Cylinders," *Composites: Design, Manufacture, and Application*, S.W. Tsai and G.S. Springer (Eds.), *Proc. 8th Int. Conf. Compos. Mater.* (published by SAMPE), (July 1991).
5. S.E. Groves, R.J. Sanchez, R.E. Lyon, and A.E. Brown, "*High Strain Rate Effects for Composite Materials*," Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-107836 (1992); accepted for publication in *ASTM Composite Materials: Testing and Design* (1992).
6. S.E. Groves, R.J. Sanchez, and S.J. DeTeresa, "Evaluation of the Compressive Performance of Helically Wrapped Carbon/Epoxy Tubes with Three Different Epoxy Matrices," presented at the ASTM Symposium: Compression Response of Composite Structures, Miami, Florida (November 16–17, 1992).

Fiber-Optic Raman Spectroscopy for Cure Monitoring of Advanced Polymer Composites

Richard E. Lyon
Materials Division
Chemistry and Materials Science Department

Thomas M. Vess and S. Michael Angel
Environmental Sciences Division

M.L. Myrick
Department of Chemistry
University of South Carolina
Columbia, South Carolina

The curing reaction of an epoxy resin matrix that is used for wet-filament-wound composites was monitored using Raman spectroscopy measured over fiber optics. The resin system consists of the diglycidyl ether of bisphenol-A in combination with a polyethertriamine hardener in a 1:1 stoichiometric ratio. The extent of chemical reaction of the epoxy as a function of time was measurable through changes in peak heights of several vibrational modes. A Raman peak associated with a phenyl ring vibration in the epoxide component was used as an internal reference to correct for density changes and instrumental variations. The feasibility of simultaneous temperature measurements was successfully demonstrated with the same fiber optics used to obtain the cure chemistry data, by measuring the intensity of anti-Stokes Raman scattering from the epoxy.

Introduction

Although significant improvements in the performance of fiber reinforcements and polymer matrix materials have been achieved in the past decade, composite processing technology has not kept pace with these advances. Consequently, high-performance material properties are not realized in composite parts fabricated using current processing methods, and manufacturing costs are high. 'Smart' processing of thermoset matrix composites could dramatically reduce manufacturing costs by reducing the rejection rate and improving part quality, through cure cycle optimization and 'on the fly' process adjustments to account for variations in the chemical composition of the starting materials. Unfortunately, commercially available cure sensors for automated control are currently limited to dielectric[1,2,3] or ultrasonic[4] measurements, which sense only mechanical property changes in the resin and cannot provide a direct measure of the cure chemistry in the composite. Furthermore, recently proposed fiber-optic spectroscopic sensors using mid-infrared[5,6,7] or ultraviolet-visible[8,9] wavelengths are either prohibitively expensive and yield little or no additional information when compared with commercially available cure sensors, or contain a large number of spectral interferences that make data interpretation difficult, if not impossible.[10]

Raman spectroscopy is an established technique for the analysis of polymers,[11,12,13] chemical reactions,[14] and thermosetting polymer composite cure reactions.[15,16] It has many advantages over mid-

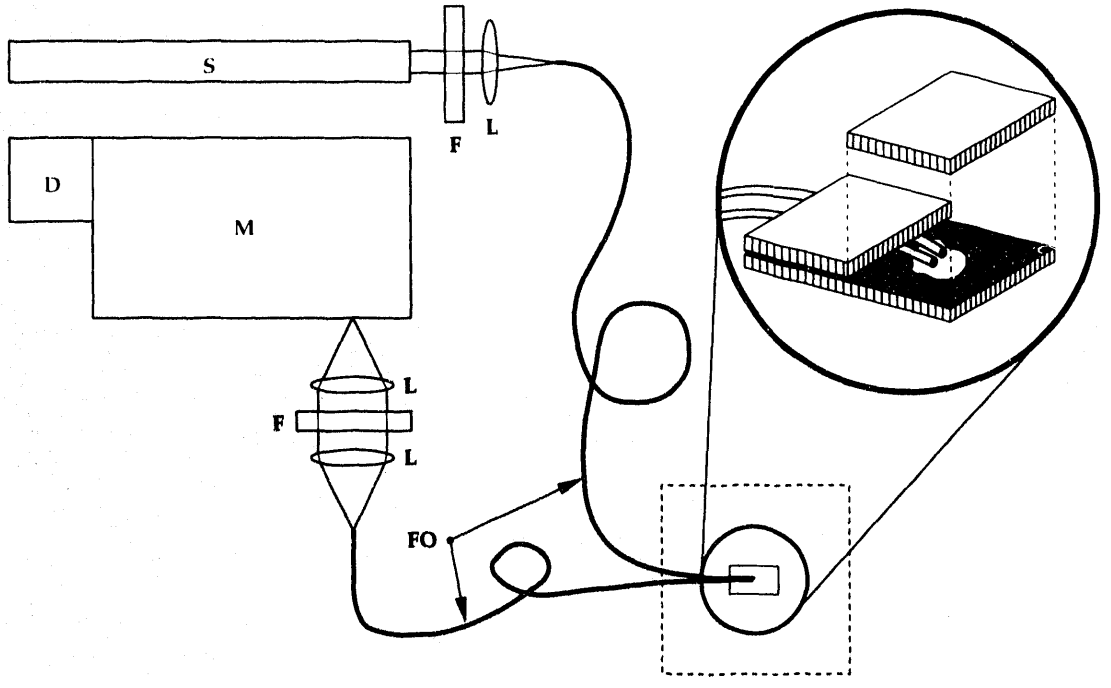


Figure 1. Experimental arrangement for dual-fiber probe geometry. Inset emphasizes sample location. M: spectrograph, D: detector, L: lens, S: laser source, FO: fiber optic, F: filter.

infrared absorption and UV-visible fluorescence spectroscopies for polymer composite cure monitoring, including broader applicability, potentially higher sensitivity and selectivity, as well as the freedom from large background corrections caused by fiber absorption. Moreover, Raman spectroscopic measurements can be conducted remotely and *in situ* using rugged, inexpensive, fused silica optical fibers (available from domestic suppliers) and economical, diode laser-excitation sources commonly used in communications and electronic equipment.[17] Subsequent to monitoring the cure reaction of the composite matrix, the quartz optical fibers could be used as embedded strain or damage sensors, or used for monitoring chemical degradation or moisture absorption of the resin.

Progress

During FY-92, the first year of the project, Raman-active vibrational bands in epoxy resins were identified, and tentative band assignments were made for the epoxide functional group and the phenyl ring backbone, from model compound studies. Cure monitoring of standard epoxy resins was demonstrated using several meter lengths of 200-μm-dia quartz optical fiber using either single-fiber or dual-fiber probes.[18] The single-fiber probe experiments were discontinued due to problems with high fiber background and the inability to accurately correct the data. The utility of economical near-infrared (NIR) diode lasers as excitation sources for cure monitoring was demonstrated in comparison studies with a conventional Ti-sapphire laser operating at 819-nm wavelength. We also demonstrated for the first time the proportionality between Raman peak ratios and epoxide group concentration in full density epoxide resins, validating the Raman scattering technique for thermoset resin cure monitoring.

Experimental

The epoxy system studied was a 1:1 stoichiometric ratio of diglycidyl ether of bisphenol-A resin (DER 332, Dow Chemical) and a polyoxypropylene triamine hardener (Jeffamine T-403; Texaco Chemical) in a weight ratio of 100/45 resin/hardener. A detailed characterization of this epoxy has been reported elsewhere.[19,20]

The dual-fiber probe experiment is shown in **Fig. 1** and has been described in detail elsewhere.[18] Dual-fiber Raman probes consist of two fibers, side-by-side, cemented in place between two microscope slides so as to meet at an angle of approximately 15°, with the polished probe ends of the fibers extended several millimeters beyond the top slide, as shown in the inset of **Fig. 1**. One fiber-optic was used to transmit the laser to the epoxy sample, while the other was used to collect the Raman scattering from the sample and transmit it to the detection system. Dual-fiber probe cure measurements were made by thoroughly mixing the liquid epoxy components, and then adding a drop

of the epoxy mixture to the probe tip so as to completely cover the sensor. No effort was made to de-gas the epoxy before injection. Some air bubbles appeared during processing, but these did not appear to affect the signal quality. A cover slide was then placed over the probe area, using spacers to provide a 2-mm gap filled with epoxy, and the entire assembly was placed in a temperature-controlled oven.

Results and Discussion

Separate Raman measurements of the DER 332 resin and Jeffamine T-403 hardener components revealed that the resin component scattering was approximately three to five times as intense as that of the hardener, so that in the 100/45 w/w ratio used for the epoxy cure studies, the resin scattering is dominant. **Figure 2** shows a series of *in situ* spectra obtained during a typical epoxy curing. These spectra are averages of spectra accumulated over several minute intervals. Measurements at one-second intervals yielded comparable results and confirm the potential of this technique for real-time cure monitoring. One major peak near 1112 cm^{-1} remains relatively constant during the curing process and can be used as an internal standard to correct for fluctuations in the sample density, clarity, refractive index, and instrumental factors during measurement. This peak has been associated with an asymmetric breathing vibration of the aromatic rings in the diglycidyl ether of bisphenol-A epoxide resin. Several peaks disappear or shift during the curing process, while several new peaks appear. The most obvious change is in the peak located near 1240 cm^{-1}. This peak loses much of its intensity during the cure and can be assigned to a stretching vibration of the epoxide ring. Although all of the peaks appeared to change in absolute intensity, probably in response to the changing density of the epoxy as it cures, the ratios of the 1240 cm^{-1} peak to the 1112 cm^{-1} peak decrease smoothly as a function of time.

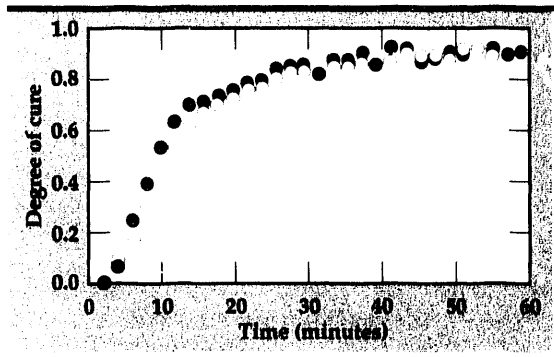


Figure 3: Isothermal degree of cure (α) for epoxy at 90°C vs time. Solid and open circles are data from diode-laser and Ti:Sapphire-laser Raman experiments, respectively.

Figure 3 is a plot of two sets of Raman data for degree of cure, α(t), vs time for an isothermal cure at 90°C. Both experiments used the 2-mm sample thickness, dual-fiber probe arrangement; however, one experiment used a Ti:sapphire laser while the other used a diode laser. The agreement between the duplicate cure experiments is seen to be excellent, indicating a high degree of reliability for the technique. Moreover, the trend of α vs time closely approximates previously published degree-of-cure data for this same DER 332/T-403 epoxy,[19] using NIR absorption spectroscopy[10] at slightly different stoichiometry, temperature, and sample thickness.

In addition to the peak height changes, the peaks were seen to shift to lower energy as the cure progressed. The total shift was small (approximately 5 to 20 cm^{-1}), but was reproducible. A study performed on a previously cured sample showed no change in vibrational frequency with temperature. Consequently, the peak shift to lower energy during cure is not simply a thermal phenomenon, but may be caused by a change in

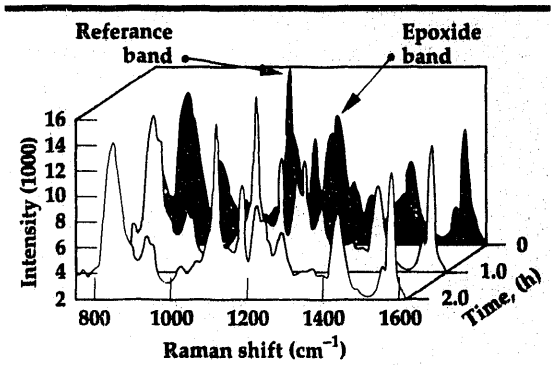


Figure 2. Series of corrected, dual-fiber probe, Raman spectra for the epoxy, taken immediately after mixing (black), 1.2 h at 75°C (gray), and 2.2 h at 75°C followed by a 1-h post-cure at 90°C (white).

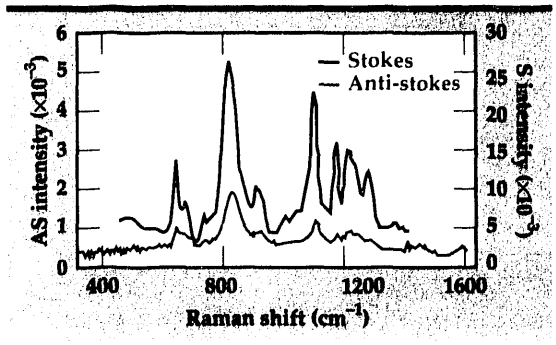


Figure 4: Anti-Stokes Raman spectrum of the cured epoxy, superimposed on Stokes Raman spectrum for a temperature of 100°C.

the local environment as the curing process occurs. For example, the increase in viscosity during the curing process may stiffen the local environment and mix the vibrations of the individual monomers more strongly with low frequency bulk vibrations. If this is the case, it is conceivable that this phenomenon could be used as a measure of local viscosity, but this possibility has not yet been tested.

The measurement of sample temperature by comparison of Stokes and anti-Stokes Raman scattering (see **Fig. 4**) is straightforward and readily accomplished. The theoretical ratio of the intensities of the anti-Stokes (I_{AS}) and Stokes (I_S) scattering is

$$\frac{I_{AS}}{I_S} = \left[\frac{\nu(0) + \nu(i)}{\nu(0) - \nu(i)}\right]^4 \exp\left\{\frac{-hc\nu(i)}{kT}\right\}, \quad (1)$$

where ν(0) is the laser frequency; ν(i) is the vibrational energy of the i-th mode; and h, c, and k are Planck's constant, the speed of light, and Boltzmann's constant, respectively. For the epoxy system studies here, a plot of the natural logarithm of I_{AS}/I_S (using the 829±3 cm^{-1} vibrational mode) against the inverse temperature, 1/T, for temperatures ranging from T = 294 to 455 K yielded a straight line with a slope of 1200 K, an intercept of 0.517, and a correlation coefficient of 0.997. This calibration curve can now be used to determine the *in situ* temperature of the resin system at any given time during a cure cycle. The precision of temperature measurements is limited by the signal-to-noise ratio of the much weaker anti-Stokes peak (S/N ≈ 20 in **Fig. 4**). An uncertainty in temperature, ΔT ≈ ±5 K, at 373 K is estimated from the relationship, $\Delta T = T^2/|B(S/N)|$, obtained by differentiation of **Eq. 1**, where B is the slope of the line. However, the accuracy of temperature measurements can be improved over the results described here, by using high-performance optical filters and also by using longer integration times.

In summary, fiber-optic Raman spectroscopy can be used for remote, *in situ* monitoring of the reaction chemistry and temperature of epoxies used as matrix materials in fiber composites during the cure cycle. While single-fiber probes were found to suffer from fiber-background effects, dual-fiber probes having dimensions on the order of 200 μm have been demonstrated successfully. Moreover, the quality of the spectra obtained for the epoxy is sufficient to warrant use of the dual-fiber Raman probe as an *in situ* quality control technique prior to cure monitoring.

Future Work

Future work will focus on (1) refining and miniaturizing the sensor; (2) evaluating the Raman fiber-optic technique for monitoring other thermoset polymer cure chemistries; (3) performing measurements in thermoset matrix fiber composites; and (4) developing simultaneous, multi-point sampling capability. Based on the preliminary results presented in this paper, we feel that in combination with compact new instrumentation and economical diode-laser excitation sources, fiber-optic Raman spectroscopy can be used to configure a rugged process monitoring and control system for an automated, polymer composite production environment.

Acknowledgements

The authors would like to express thanks to Gerald Goldstein of the Office of Health and Environmental Research (RPIS No. 003906) for supporting a part of this research, and to Katherine Chike of the University of South Carolina for experimental work in validating this application of Raman spectroscopy.

1. P.R. Ciriscioli and G.S. Springer, *SAMPE J.* **25** (3), 35 (1989).
2. W. Sichina and D. Shepard, *Mater. Eng.*, 49 (July 1989).
3. D.R. Day, D.D. Shepard, and A.S. Wall, "Thermoset Process Control Utilizing Microdielectric Sensors," *Proc. ASME Conf. Advanced Composites and Processing Technology* (Chicago, Illinois), 1 (November 27–December 2, 1988).
4. S.S. Saliba, T.E. Saliba, and J.F. Lanzafame, "Acoustic Monitoring of Composite Materials During the Cure Cycle," *Proc. 34th Int. SAMPE Symposium* **34** (1), 397 (1989).
5. R.E. Schirmer and A.G. Gargus, *Am. Laboratory* **37**, (November 1988).
6. P.R. Young, M.A. Druy, W.A. Stevenson, and D.A.C. Compton, *SAMPE J.* **25** (2), (1989).
7. M.A. Druy, L. Elandjian, W.A. Stevenson, R.D. Driver, G.M. Leskowitz, and L.E. Curtiss, "Fourier-Transform Infrared (FTIR) Fiber Optic Monitoring of Composites During Cure in an Autoclave," *SPIE Proc. Vol. 1170, Fiber Optic Smart Structures and Skins II* (Boston, Massachusetts), 150 (September 5–8, 1989).

8. R.L. Levy and S.D. Schwab, *Polym. Compos.* **12** (2), 96 (1991).

9. N.H. Sung, W. Dang, and H.J. Paik, "*In situ* Monitoring of Epoxy Cure by Fiber-Optic Molecular Sensors," *Proc. 36th Int. SAMPE Symposium* **36** (2), 1461 (1991).

10. H. Dannenberg, *SPE Trans.* **3** (1), 78 (1963).

11. D.L. Gerrard and W.F. Maddams, *Appl. Spectrosc. Rev.* **22**, 251 (1986).

12. W.F. Maddams, *American Laboratory* (March 1986).

13. C.E. Miller, D.D. Archibald, M.L. Myrick, and S.M. Angel, *Appl. Spectrosc.* **44**, 1297 (1990).

14. W. Doyle and N.A. Jennings, *Spectroscopy Int. J.* **5** (1), 34 (1990).

15. C. Johnson and S.L. Wunder, *SAMPE J.* **26** (2), 19 (1990).

16. J.C. Johnson, *FT-Raman Investigation of Curing Reactions in Polyimides*, Ph.D. Dissertation, Temple University, Philadelphia, Pennsylvania (1990).

17. S.M. Angel, M.L. Myrick, and T.M. Vess, "Remote Raman Spectroscopy Using Diode Lasers and Fiber-Optic Probes," *Proc. SPIE '91, Optical Methods for Ultrasensitive Detection and Analysis: Techniques and Applications* (Los Angeles, California), **1435**, 72 (1991).

18. M.L. Myrick, S.M. Angel, R.E. Lyon, and T.M. Vess, *SAMPE J.* **28** (4), 37 (1992).

19. T.T. Chiao and R.L. Moore, "A Room-Temperature Curable Epoxy for Advanced Composites," *Proc. 29th Annual Technical Conf. Reinforced Plastics/Composites Inst., SPI*, Section 16-B, 1 (1974).

20. F.M. Kong, C.M. Walkup, and R.J. Morgan, "Structure-Property Relationships of Polyethertiamine-Cured Bisphenol-A-diglycidyl Ether Epoxies," *Epoxy Resin Chemistry II*, R.S. Bauer (Ed.), ACS Symposium Series **221**, 211, 1983.

Modeling Superplastic Materials

Donald R. Lesuer,
Choi K. Syn, and
Charles S. Preuss
Engineering Sciences
Mechanical Engineering

Peter J. Raboin
Nuclear Explosives Engineering
Mechanical Engineering

We have developed a model that accounts for grain growth during superplastic flow, and its subsequent influence on stress/strain/strain rate behavior. Our studies are experimentally based and have involved two different types of superplastic materials: a quasi-single phase metal, Coronze 638, and a microduplex metal, ultrahigh-carbon steel. We have studied the kinetics of strain-enhanced grain growth in both materials as a function of strain, strain rate, and temperature. An equation for the rate of grain growth has been developed that incorporates the influence of temperature. Our model integrates grain growth laws derived from these studies, with two mechanism-based, rate-dependent constitutive laws to predict the stress/strain/strain rate behavior of materials during superplastic deformation. The material model has been added to the NIKE2D code through an enhancement of the Deformation Mechanism Model. The predictions of the model have been compared with data from several experiments.

Introduction

Superplasticity is the capability to deform crystalline solids in tension to unusually large plastic strains, often well in excess of 1000%. This phenomenon results from the ability of the material to resist localized deformation much the same as hot glass. The material also deforms with very low flow stress. Thus, materials with superplastic properties provide the opportunity to form complex components into shapes very near final dimension. This greatly reduces machining and material costs and minimizes the amount of scrap produced.

Superplastic materials exhibit high elongations, because a deformation mechanism known as grain boundary sliding (GBS) is active. This deformation behavior occurs within a relatively narrow range of temperature and strain rate. If the strain rate is too high, then a different mechanism called diffusion-controlled dislocation creep (slip creep) is activated, and ductilities are substantially reduced. On the other hand, if the strain rate is too low, then a deformation mechanism known as diffusional flow prevails, and the ductility is also reduced relative to GBS. From a commercial standpoint, forming components at high strain rates is attractive, because operations can be done with less time and cost. Often this means superplastically forming at strain rates close to the slip creep regime. Thus, our work is concentrating on the two higher strain rate regimes, GBS and slip creep.

The active deformation mechanisms also depend strongly on the microstructure of the material, such as an ultra-fine grain structure. Unfortunately, these ultra-fine grains can grow during deformation, resulting in the loss of superplasticity. Thus, it is important to gain a quantitative understanding of this process and its influence on material forming.

For these reasons, material models for the constitutive behavior of materials during superplastic flow should account for microstructure, its evolution, and changes in deformation mechanism throughout the deformation history. The objective of this project is to develop a model of these structural changes and their influence on stress/strain/strain rate behavior, using mechanism-based constitutive laws.

Progress

In our work during FY-91, we (a) established the kinetics of strain-enhanced grain growth for isothermal conditions and (b) developed a model that integrates these grain growth laws with mech-

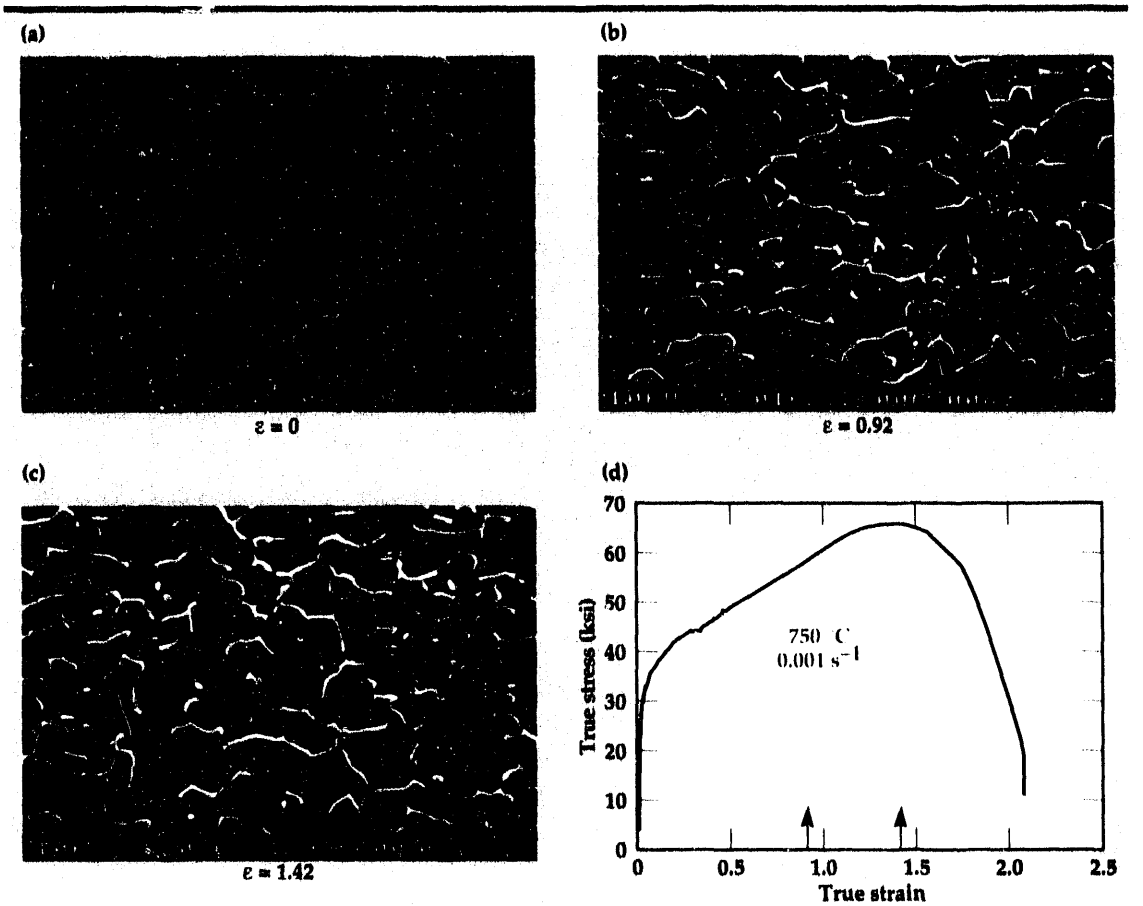


Figure 1. Microstructure of ultrahigh-carbon steel before deformation (a) and after superplastic deformation to a true strain of .92 (b) and 1.42 (c), respectively. Microstructure consists of ferrite grains and iron carbide particles. Superplastic deformation has caused growth in both the ferrite grains and the carbide particles. The true stress/true strain behavior for this material is shown in (d). Strains at which the photomicrographs were taken are indicated on the plot.

anism-based constitutive equations. During FY-92, we expanded the strain-enhanced grain growth equations to account for the influence of temperature. The material model was also put into the NIKE2D code, and the performance of the model was evaluated with laboratory experiments.

Microstructural Evolution

Our studies have involved two superplastic materials with significantly different microstructures: ultrahigh-carbon steel (UHCS), which has a microduplex structure, and Coronze 638, which is quasi-single phase. In both of these materials, the dominant microstructural change during superplastic deformation is grain growth, and no dynamic recrystallization was expected at the temperatures and strain rates of interest. The UHCS had the composition Fe-1.8C-1.6Al-1.5Cr-.5Mn and had been thermomechanically processed to produce a matrix of ultrafine, equiaxed ferrite grains (mean linear intercept was approximately 0.74 μm) containing spheroidized carbides. The Coronze 638 (Cu-2.8Al-1.8Si-0.4Co) is a commercially available alloy that consists of essentially pure copper containing a sub-micron-size dispersion of CoSi and $CoSi_2$ particles. The material was received as a repeatedly rolled and annealed sheet with mean linear intercept grain size equal to 1.9 microns. Further details of the experimental plan are reported elsewhere.[1]

Strain-Enhanced Grain Growth. **Figure 1** shows the influence of superplastic deformation at 750°C, and a true strain rate of .001/s on the microstructure of UHCS. **Figure 1a** was taken in the grip section of the sample, and was thus exposed to the testing temperature without plastic deformation. **Figures 1b** and **1c** were taken in the gage section of samples deformed to true strains of 0.92 and 1.42, respectively. A comparison of the figures shows strain-enhanced grain growth in that the ferrite grains have grown. The carbide particles have also coarsened, and the size of the ferrite grains appears to be determined by the intercarbide spac-

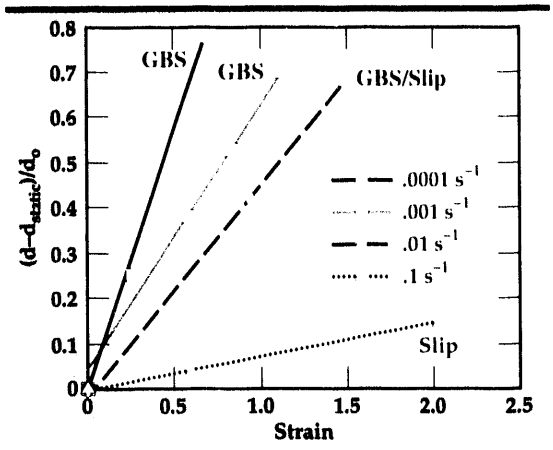


Figure 2. Normalized strain-enhanced grain growth vs strain for Coronze superplastically deformed at four true strain rates. Dominant deformation mechanism is noted for each strain rate.

ing. This suggests that the kinetics of grain growth are determined by the kinetics of carbide coarsening. The stress/strain curve in **Fig. 1d** shows the importance of this grain growth on the deformation behavior of UHCS: increasing the grain size from its initial size (0.74 μm) to the size at a strain of 1.42 (1.48 μm) has raised the flow stress from 5 ksi to over 9 ksi. Thus, grain growth has produced significant hardening, and the grain size is an important parameter for characterizing the current mechanical state of the material.

In these studies, static annealing grain growth (normal grain growth) and strain-enhanced grain growth are assumed to be additive. Thus, the kinetics of grain growth can be expressed as

$$\frac{\dot{d}}{d_0} = \frac{\dot{d}_a}{d_0} + \frac{\dot{d}_{se}}{d_0}, \qquad (1)$$

where $\dot{d}$ is the total rate of grain growth; $\dot{d}_a$ is the grain growth rate due to static annealing; $\dot{d}_{se}$ is the grain growth rate due to strain; and d_0 is the initial grain size prior to deformation or exposure to elevated temperature.

The grain structure in the gage section of samples is the result of both static and strain-enhanced grain growth. On the other hand, the grain structure in the grip is the result of static grain growth only. The strain-enhanced grain growth was calculated as the difference in mean-linear-intercept grain size between measurements taken in the gage and grip sections of the sample. We used this procedure to determine the normalized strain-enhanced grain growth response for Coronze. Wilkinson and Caceres[2] have obtained data for this

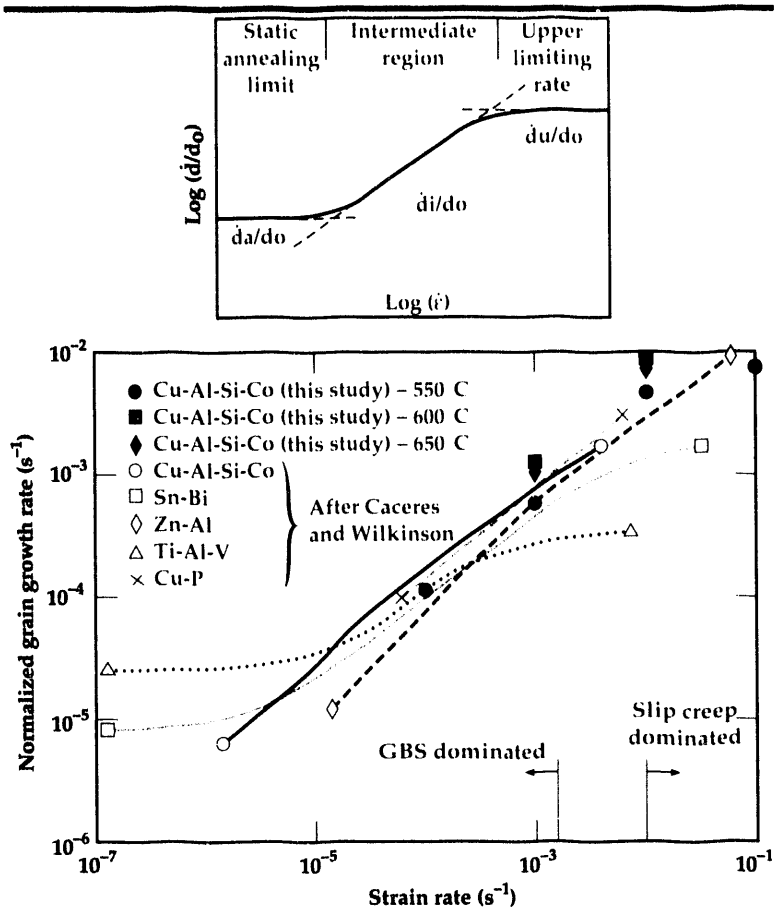


Figure 3. Normalized grain growth rate vs strain rate for a number of quasi-single phase and microduplex superplastic materials. Plot is from the work of Caceres and Wilkinson.[2] Data from our study of Coronze at 550 C, 600 C, and 650 C have been added to the plot. The strain rates over which there is a transition in deformation mechanism from GBS to slip creep, are indicated. The inset shows the three different regions for the curve.

material. The present studies have obtained data at higher strain rates. The strain-enhanced grain growth for Coronze is plotted as a function of true strain in **Fig. 2**, for tests conducted at 550°C and four strain rates. Results have been normalized by the initial grain size. The tests at the three slowest strain rates were in the region in which GBS is the dominant deformation mechanism. The test at the highest strain rate was in the region where the dominant deformation mechanism was slip creep. For the three slowest strain rates, the normalized strain-enhanced grain growth was found to have a linear dependence on strain and a power-law dependence on strain rate. These results are consistent with the observations of Caceres and Wilkinson on the Coronze alloy.[2] For the highest strain rate, the grain growth data in **Fig. 2** had a much smaller slope. The reason for this will be discussed in the following paragraphs.

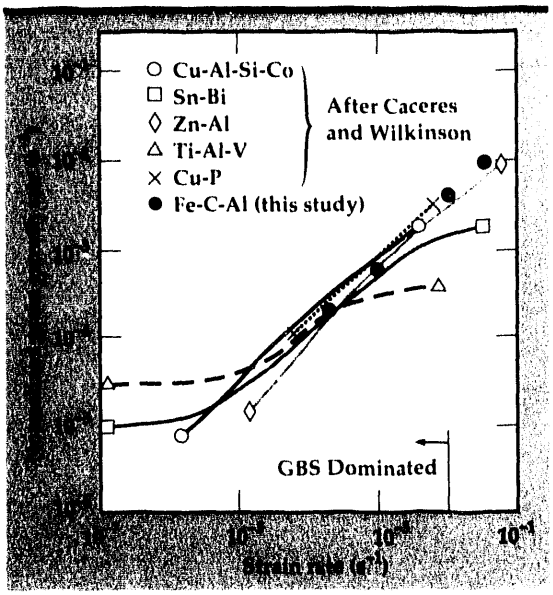


Figure 4. Normalized grain growth rate vs strain rate for a number of quasi-single phase and microduplex superplastic materials. Plot is from the work of Caceres and Wilkinson.[2] Data from our study of the superplastic deformation of UHCS at 750°C has been added to the plot. Strain rate below which GBS is the primary deformation mechanism, is indicated.

The normalized, strain-enhanced grain growth rate (with respect to time) can be calculated from the data in **Fig. 2**, by multiplying the slopes of individual lines by the strain rate for that test. These grain growth rates have been calculated and added to a figure previously reported by Wilkinson and Caceres,[2] which shows a log-log plot of normalized grain growth rate vs strain rate. The results are shown in **Fig. 3**. The plot is quite significant, since it shows data for both quasi-single phase and microduplex materials, for different homologous temperatures and for a range of starting grain sizes. It is clear that a common equation can describe the grain growth behavior of a number of different materials, including Coronze. The curve has three distinct regions. In one region, the normalized grain growth rate is a power law function of strain rate; at higher or lower strain rates, the normalized grain growth rate reaches a limiting value that is independent of strain rate. These regions are shown schematically in the inset for **Fig. 3**. The region at the lowest strain rate is the result of static grain growth, whereas the regions at the intermediate and high strain rates represent the result of strain-enhanced grain growth. It is reasonable to assume that the highest grain growth rate represents a limiting rate determined by the kinetics of grain boundary migration. The curve shown in **Fig. 3** can be described by

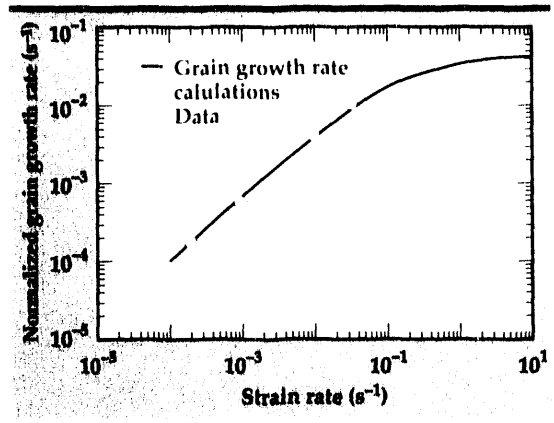


Figure 6. Grain growth rates for UHCS at 750°C, predicted by Eq. 5. Calculations are based on the parameters in Table 1. Experimental data is provided.

$$\frac{\dot{d}}{d_0} = \frac{\dot{d}_a}{d_0} + \frac{1}{d_0}\left(\frac{\dot{d}_i \dot{d}_u}{\dot{d}_i + \dot{d}_u}\right), \qquad (2)$$

where $\dot{d}_i$ is the grain growth rate at intermediate strain rates, and $\dot{d}_u$ is the upper limiting grain growth rate. The intermediate region has a power-law dependence on strain rate, which produces

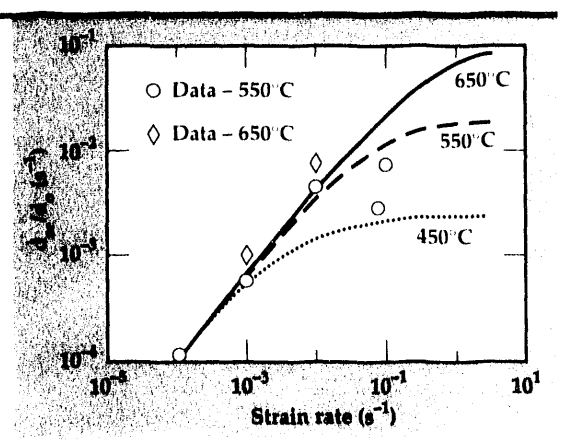


Figure 5. Grain growth rates for Coronze at 450°C, 550°C, and 650°C, predicted by Eq. 5. Calculations are based on the parameters given in Table 1. Experimental data at 550°C and 650°C are given.

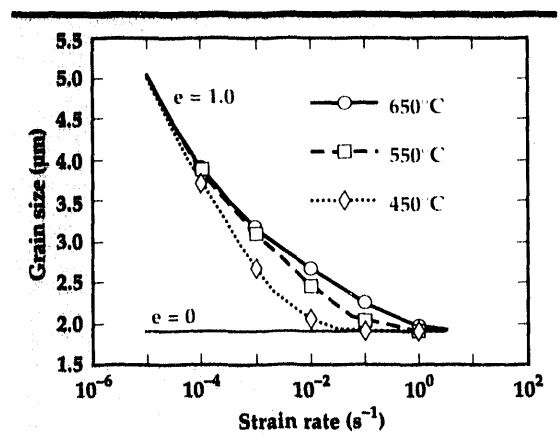


Figure 7. Calculated grain sizes for Coronze after constant strain rate testing to a true strain of 1 at the indicated strain rates.

the following empirical expression for the normalized rate of grain growth:

$$\frac{\dot{d}}{d_0} = \frac{\dot{d}_u}{d_0} + \frac{1}{d_0}\left(\frac{\lambda_0 \dot{\varepsilon}^n \dot{d}_u}{\lambda_0 \dot{\varepsilon}^n + \dot{d}_u}\right), \qquad (3)$$

where λ_0 and n are constants.

The strain rates at which there is a shift in the operating deformation mechanisms (from GBS to slip creep) are also shown in **Fig. 3**. It is important to note that the grain growth rates for the three slowest strain rates appear to have a power-law dependence on strain rate. The grain growth rate for the highest strain rate, however, shows a substantially smaller increase with increasing strain rate than the grain growth rates at the lower strain rates. This transition occurs at about the same strain rate as the transition to slip-creep-dominated deformation. The obvious implication is that the loss of GBS as a deformation mechanism has reduced the contributions of strain-enhanced grain growth to the total grain growth rate. Several mechanisms have been proposed to explain strain-enhanced grain growth.[3-6] All of these mechanisms result from grain boundary sliding or grain switching events. It is reasonable to assume that strain-enhanced grain growth will exist only if these mechanisms provide significant contributions to the total strain. Thus, contributions from strain-enhanced grain growth can be limited by a loss of superplastic flow or by the limiting grain growth rates defined by the rates of grain boundary migration.

Identical procedures were used to determine the normalized strain-enhanced grain growth rate for UHCS during superplastic deformation at 750°C. Results are presented in **Fig. 4** and fall within the range of grain growth rates for other materials represented on the plot. For UHCS, no upper limit is reached on grain growth rate over the strain rates studied.

Temperature Dependence of Strain-Enhanced Grain Growth. A general extension of **Eq. 3** that accounts for the temperature dependence of grain growth can be developed assuming different temperature dependencies for the three processes in **Eq. 2**. The temperature dependence of normal grain growth kinetics has been studied and equations developed (see, for example, **Ref. 7**). The primary interest in this study is strain-enhanced grain growth, and thus the intermediate and upper rate processes represented in **Eq. 2**. The temperature dependence of these processes can be represented as

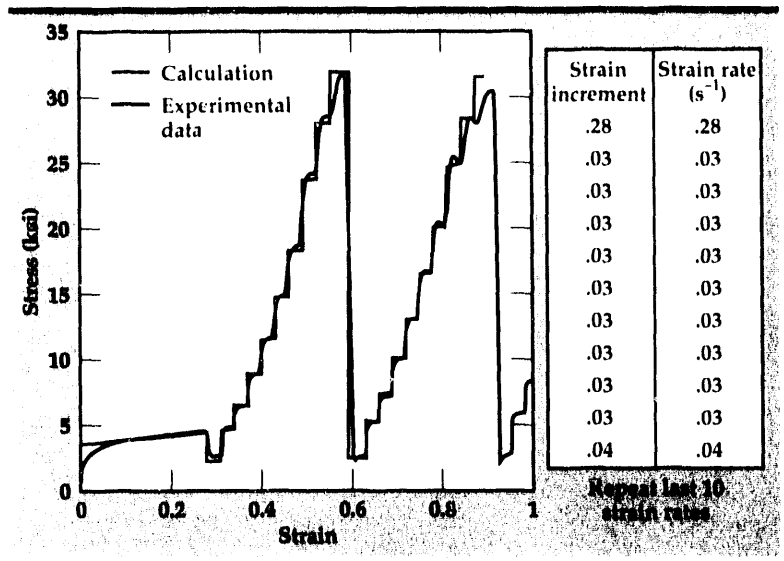


Figure 8. Stress/strain response for UHCS deformed at 750°C through a predetermined strain rate/strain history. The strain rate/strain history is shown in the inset. The calculations are based on the material model described in this report.

$$\dot{d}_i = \lambda_0 \dot{\varepsilon}^n \exp\left(\frac{-Q_i}{RT}\right) \qquad (4a)$$

$$\dot{d}_u = \left(\dot{d}_u\right)_0 \exp\left(\frac{-Q_u}{RT}\right), \qquad (4b)$$

where Q_i and Q_u are activation energies for the intermediate and upper regions, respectively; λ_0 and $(\dot{d}_u)_0$ are constants; R is the gas constant, and T is the absolute temperature. Combining these expressions yields a general equation for the temperature dependence of strain-enhanced grain growth:

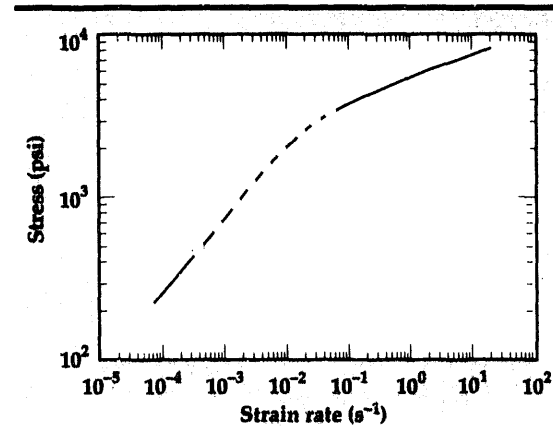


Figure 9. Stress/strain rate response for UHCS deformed at 750°C.

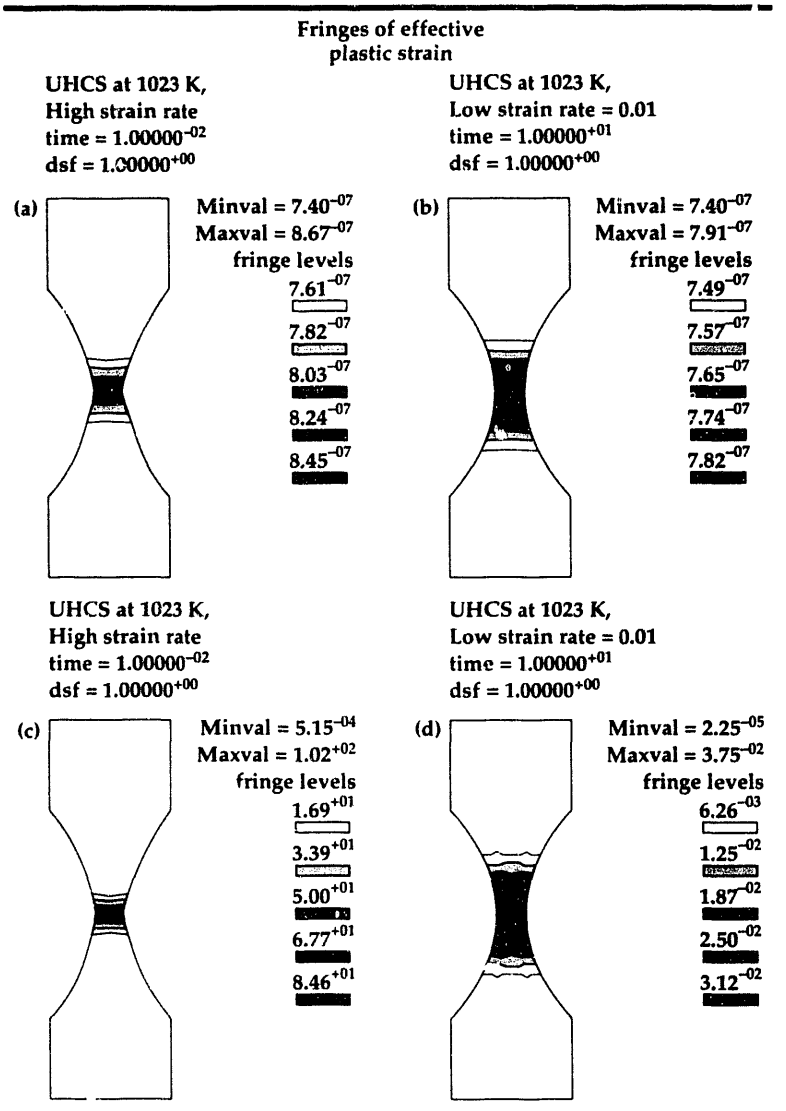


Figure 10. Hourglass-shaped sample of UHCS deformed at 750 C for two different extension rates: one in which GBS is the dominant deformation mechanism, and one in which slip creep is the dominant deformation mechanism. The figure shows contours of constant grain size and strain rate for the sample.

Table 1. Parameters used in Eq. 5 for temperature dependence of strain-enhanced grain growth.

	d_0 (μm)	λ_0 [(μm/s)s^n]	$(\dot{d}_u)_0$ (μm/s)	n	Q_i	Q_u (kJ/mole)
Coronze	1.5	.178	7.98×10^4	.806	0	104*
UHCS	.74	.136	1.92×10^5	.799	0	170†

* Activation energy for grain boundary diffusion in pure copper[5]

† Activation energy for grain boundary diffusion in pure iron[9]

$$\frac{\dot{d}_{se}}{d_0} = \frac{1}{d_0}\left[\frac{\lambda_0 \dot{\varepsilon}^n (\dot{d}_u)_0 \exp\left[\frac{-(Q_i + Q_u)}{RT}\right]}{\lambda_0 \dot{\varepsilon}^n \exp\left(\frac{-Q_i}{RT}\right) + (\dot{d}_u)_0 \exp\left(\frac{-Q_u}{RT}\right)}\right]. \quad (5)$$

The temperature dependence of strain-enhanced grain growth for the Coronze alloy was experimentally evaluated at 600°C and 650°C. The resulting grain growth rates have been added to the plot in **Fig. 3**, and appear to fall within the range of strain-enhanced grain growth rates for other materials. These results suggest that strain-enhanced grain growth for the Coronze alloy is independent of temperature and that Q_i for this material is zero. As mentioned in the previous section, the limiting grain growth rate at high strain rate is probably controlled by the rate of grain boundary migration. We therefore assume that a reasonable value for Q_u is the activation energy for grain boundary diffusion. The calculated grain growth rates that are predicted by **Eq. 5** are plotted as a function of strain rate in **Fig. 5**. Calculations are shown for three temperatures, 450°C, 550°C, and 650°C; the parameters are given in **Table 1**. The calculated grain growth rates show good agreement with rates derived from experimental data. Because the strain-enhanced grain growth is independent of temperature in the intermediate region, at very high temperatures (higher than the temperature studied here), the contribution of static annealing to the total grain growth rate could be significantly higher than the contribution from strain-enhanced grain growth. In **Table 1**, the parameters for UHCS are also given. Both Coronze and UHCS have very similar strain rate exponents (n) and values for the constant λ. The calculated strain-enhanced grain growth rates for UHCS deformed at 750°C are shown in **Fig. 6**. Good agreement was obtained with experimental data.

The final grain sizes that would be obtained for Coronze after tensile testing (to a true strain equal to 1) at a constant strain rate are shown in **Fig. 7**. Calculations, which were done for three temperatures (450°C, 550°C, and 650°C), are based on **Eq. 5**, using the parameters given in **Table 1**, and, thus, do not include the effects of static annealing. Results in **Fig. 7** show a decreasing final grain size with increasing strain rate and very little grain growth above .1/s for all testing temperatures. Grain growth decreases with increasing strain rate (**Fig. 7**) despite the increasing grain growth rate

with increasing strain rate (**Fig. 5**). This occurs because the strain rate exponent (n) in **Eq. 5** is less than one. The total amount of grain growth is very sensitive to the value of n.

Mechanical Response

Two rate-dependent constitutive equations have been used for GBS and slip creep,

$$\dot{\varepsilon}_{gbs} = A_{gbs} \exp\left(\frac{-Q_{gb}}{RT}\right) \sigma^{n_{gbs}} d^{-p_{gbs}} \qquad (6)$$

$$\dot{\varepsilon}_{slip} = A_{slip} \exp\left(\frac{-Q_l}{RT}\right) \sigma^{n_{slip}} \lambda^{p_{slip}}, \qquad (7)$$

where $\dot{\varepsilon}_{gbs}$ and $\dot{\varepsilon}_{slip}$ are the strain rates for grain boundary sliding and slip creep, respectively; σ is the stress; A_{gbs}, A_{slip}, n_{gbs}, n_{slip}, p_{gbs}, and p_{slip} are constants; λ is the minimum barrier spacing governing slip creep (typically, the interparticle spacing or the grain size); d is the grain size; Q_{gb} is the activation energy for grain boundary diffusion; and Q_l is the activation energy for lattice diffusion. Since the deformation mechanisms represented by these equations are additive, the total strain rate can be represented by

$$\dot{\varepsilon}_{total} = \dot{\varepsilon}_{slip} + \dot{\varepsilon}_{gbs}. \qquad (8)$$

The mean, linear, intercept grain size is typically used for the grain size term in **Eqs. 6** and **7**. For fine grain materials deforming in or near the region of GBS, the minimum barrier spacing is the grain size, and thus for these studies, we have assumed that λ equals d. The grain size was obtained from a time integration of **Eq. 1**.

Model Implementation and Evaluation

The grain growth kinetics, expressed by **Eqs. 1** and **5**, and the constitutive laws, expressed by **Eqs. 6**, **7**, and **8**, were integrated into an existing material model in the NIKE2D code, called the Deformation Mechanism Model (DMM).[10] This material model solves the constitutive equations, with an implicit solution procedure. The evolution of grain size is also solved with an implicit procedure. The numerical methods used in this model emphasize accuracy, but all of the algorithms are vectorized for the Cray computer. Generally, the speed performance of the DMM is within a factor of three of NIKE2D model 19, a rate-dependent, power-law plasticity model.

Table 2 Parameters used in Eqs. 6 and 7 for UHCS.

A_{gbs}	Q_{gb}*	n_{gbs}	p_{gbs}
$[s^{-1}(psi)^{-n}(\mu m)^{p}]$	(kJ/mole)		
1.84×10^{-3}	177	2.27	3.0

A_{slip}	Q_l†	n_{slip}	p_{slip}
$[s^{-1}(psi)^{-n}(\mu m)^{-p}]$	(kJ/mole)		
1.41×10^{-21}	252	7.14	3.0

σ in psi
d and λ in μm

* determined experimentally
† obtained from **Ref. 9**

We have evaluated the performance of the material model, using a series of experiments of increasing complexity. The first experiments were simple tensile tests conducted at constant true strain rates, and excellent agreement was obtained between model predictions and experimental data. The results were reported in our FY-91[11] report. The second set of experiments involved deforming tensile samples through a predetermined strain rate/strain history. This applied strain/strain rate history and the resulting stress/strain response for UHCS deformed at 750°C is shown in **Fig. 8**. The parameters used in **Eqs. 6** and **7** are shown in **Table 2**. In **Fig. 9**, we show the stress/strain rate behavior of UHCS. In both cases, excellent agreement was obtained between model predictions and experimental data. A third set of calculations was done to evaluate the material model on a sample containing a non-uniform stress state. The sample had an hourglass shape and was deformed at two different constant extension rates. At one rate, GBS was the dominant deformation mechanism, and at the other rate, slip creep was the dominant deformation mechanism. The extension rates are indicated in **Fig. 10**, which shows contours of constant grain size (**Figs. 10a** and **10b**) and constant strain rate (**Figs. 10c** and **10d**) after an extension of x in. The sample deformed in the slip creep region has started to neck, and the contours of strain rate are highly localized. The sample deformed in the region of GBS has avoided necking (exhibited characteristics leading to superplastic behavior) by distributing the strain rates throughout the hourglass region.

Conclusions

We draw four conclusions from our work:

1. The dependence of grain growth rate on strain rate for superplastic UHCS and coronze falls on a master curve, as originally proposed by Wilkinson and Caceres.[2] In the copper alloy, the transition in grain growth rate from a power-law dependence on strain rate to an upper limiting rate occurs at the transition from GBS-dominated behavior to slip-creep-dominated behavior. The transition to an upper limiting rate ($\dot{d}_u$ in **Fig. 3**) can occur because of a loss of superplastic flow or a limiting grain growth rate defined by grain boundary migration. For UHCS, within the strain rates studied, no upper limit was found to the grain growth rate.
2. An equation describing the temperature dependence of the strain-enhanced grain growth rate has been developed. The equation predicts grain growth rates that agree well with experimental data. For Coronze, strain-enhanced grain growth appears to be independent of temperature in the intermediate region. In the high strain rate region, the strain-enhanced grain growth rate appears to have an activation energy equal to the activation energy for grain boundary diffusion.
3. A material model has been developed that combines the temperature-dependent grain growth law described above and mechanism-based constitutive equations.
4. This model was incorporated into the NIKE2D code, and validation experiments show excellent agreement between model calculations and experimental data.

Acknowledgements

We are indebted to Oleg Sherby (Stanford University) and Amiya Mukherjee (University of California Davis) for helpful discussions on superplasticity. We are also indebted to Jack Crane (Olin Corporation) for providing the Coronze 638 and to Oleg Sherby for providing the superplastic ultrahigh-carbon steel.

1. D.R. Lesuer, C.K. Syn, K.L. Cadwell, and S.C. Mance, "Microstructural Change and Its Influence on Stress-Strain Behavior of Superplastic Materials," *Superplasticity in Advanced Materials*, S. Hori, M. Tokizane, and N. Furushiro (Eds.), (Osaka, Japan), 139, 1991.
2. D.S. Wilkinson and C.H. Caceres, *J. Mater. Sci. Lett.* **3**, 395 (1984).
3. M.A. Clark and T.H. Alden, *Acta Metall.* **21**, 1195 (1973).
4. D.S. Wilkinson and C.H. Caceres, *Acta Metall.* **32** (9), 1335 (1984).
5. K. Holm, J.D. Embury, and G.R. Purdy, *Acta Metall.* **25**, 1191 (1977).
6. D.J. Sherwood and C.H. Hamilton, *Scripta Metall.* **25**, 2873 (1991).
7. P. Cotterill and P.R. Mould, *Recrystallization and Grain Growth in Metals*, John Wiley and Sons (New York, New York), 279, 1976.
8. M.F. Ashby, *Acta Metall.* **20**, 887 (1972).
9. B. Walser and O.D. Sherby, *Met. Trans. A* **10A**, 1461 (1979).
10. P. Raboin, *A Deformation-Mechanism Material Model for NIKE2D*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-ID-112906 (1993).
11. D.R. Lesuer, D.K. Syn, K.L. Cadwell, and C.S. Preuss, "Modeling Superplastic Materials," *Engineering Research, Development, and Technology*, Lawrence Livermore National Laboratory Livermore, California, UCRL-53868-91 (1992).

Microwave and Pulsed Power

The goals of the Microwave and Pulsed Power thrust area are to identify realizable research and development efforts and to conduct high-quality research in those pulse power and microwave technologies that support existing and emerging programmatic requirements at Lawrence Livermore National Laboratory (LLNL). Our main objective is to work on nationally important problems while enhancing our basic understanding of enabling technologies such as component design and testing, compact systems packaging, exploratory physics experiments, and advanced systems integration and performance. During FY-92, we concentrated our research efforts on the six project areas described in this report.

1. We are investigating the superior electronic and thermal properties of diamond that may make it an ideal material for a high-power, solid-state switch.

2. We are studying the feasibility of using advanced Ground Penetrating Imaging Radar technology for reliable non-destructive evaluation of bridges and other high-value concrete structures. These studies include conceptual designs, modeling, experimental verifications, and image reconstruction of simulated radar data.

3. We are exploring the efficiency of pulsed plasma processing techniques used for the removal of NOx from various effluent sources.

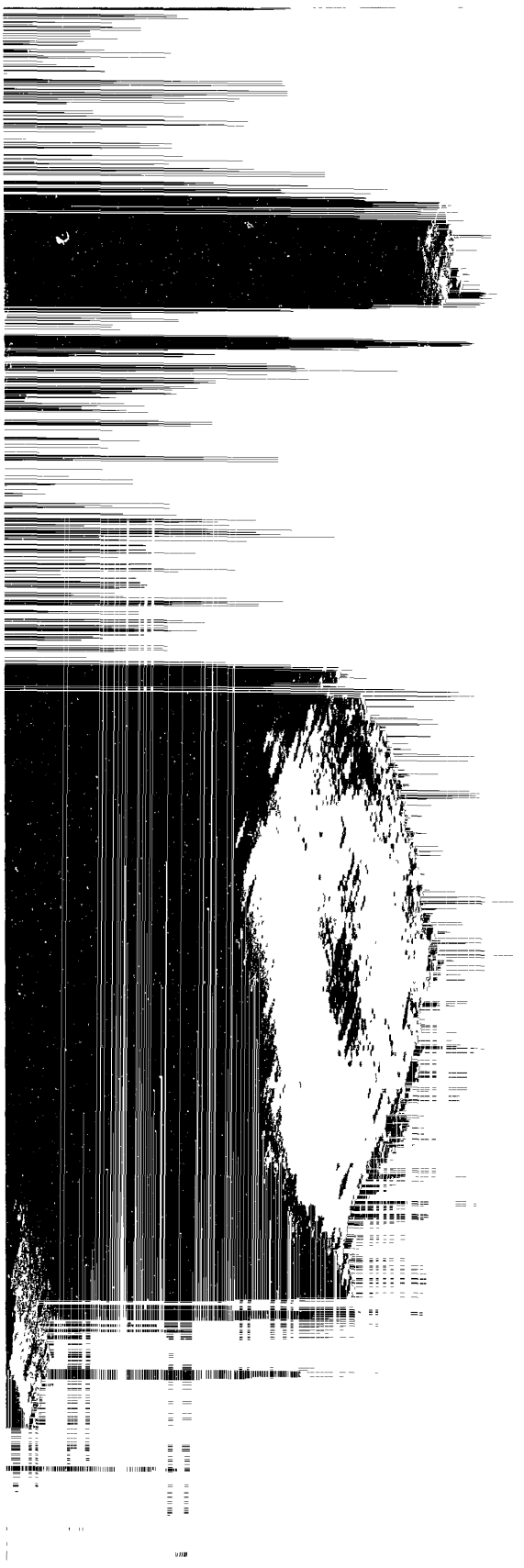
4. We have finished the investigation of the properties of a magnetically delayed low-pressure gas switch, which was designed here at LLNL.

5. We are applying statistical electromagnetic theory techniques to help assess microwave effects on electronic subsystems, by using a mode stirred chamber as our measurement tool.

6. We are investigating the generation of perfluoroisobutylene (PFIB) in proposed CFC replacement fluids when they are subjected to high electrical stresses and breakdown environments.

E. Karl Freytag
Thrust Area Leader

Section 7



7. Microwave and Pulsed Power

Overview

E. Karl Freytag, Thrust Area Leader

Pulsed Plasma Processing of Effluent Pollutants and Toxic Chemicals

George E. Vogtlin **7-1**

Ground Penetrating Imaging Radar for Bridge Inspection

John P. Warhus, Scott D. Nelson, Jose M. Hernandez, Erik M. Johansson, Hua Lee, and Brett Douglass **7-5**

High-Average-Power, Electron Beam-Controlled Switching in Diamond

W. Wayne Hofer, Don R. Kania, Karl H. Schoenbach, Ravindra Joshi, and Ralf P. Brinkmann **7-13**

Testing of CFC Replacement Fluids for Arc-Induced Toxic By-Products

W. Ray Cravey, Wayne R. Luedtka, Ruth A. Hawley-Fedder, and Linda Foiles **7-19**

Applying Statistical Electromagnetic Theory to Mode Stirred Chamber Measurements

Richard A. Zacharias and Carlos A. Avalle **7-23**

Magnetically Delayed Low-Pressure Gas Discharge Switching

Stephen E. Sampayan, Hugh C. Kirbie, Anthony N. Payne, Eugene Lauer, and Donald Prosnitz **7-27**

Pulsed Plasma Processing of Effluent Pollutants and Toxic Chemicals

George E. Vogtlin
Defense Sciences Engineering Division
Electronics Engineering

We are exploring the efficiency of pulsed plasma processing in the removal of NO_x and other pollutants. Our ultimate goal is a flow-through system where gases would be treated during a single pass. We are currently using a closed-loop system with mixtures of bottled gas. The closed-loop system permits testing of processes, without a requirement for the development of complex and expensive power supplies for the one-pass treatment. We have constructed a new processor this year that can accommodate many electrode shapes at temperatures up to 400°F.

Introduction

The efficient removal of NO_x from effluent sources is essential to meet the requirements of the Clean Air Act. NO_x is a mixture of nitric oxide, NO, and nitrogen dioxide, NO_2. We are exploring the efficiency of pulsed plasma processing in the removal of NO_x and other pollutants. Pulsed plasma appropriate for processing is generated by a short high-voltage pulse between two electrodes. The electrons from this discharge create radicals from the air molecules. These radicals can then react with the pollutants to give harmless or removable substances.

Our ultimate goal is a flow-through system where gases would be treated during a single pass. We are currently using a closed-loop system with mixtures of bottled gas. The closed-loop system permits testing of processes, without a requirement for the development of complex and expensive power supplies for the one-pass treatment. We also believe that flow through the reactor should be in turbulent flow. Turbulent flow means that all the gases in the processor flow through at the same velocity, including that at the wall. The closed-loop system permits these high flow rates without an extensive gas-mixing and heating system.

Progress

Experimental System

The experimental system permits the introduction of various gas combinations prior to testing. Analysis can be conducted during or after these tests. We have constructed a new processor this year that can accommodate many electrode shapes at temperatures up to 400°F. This processor is shown in **Fig. 1**.

Electrode geometries can have a crucial role in the efficiency of this process. It is essential to efficiently couple the energy uniformly into the gas. The geometry can affect the power supply coupling efficiency, the discharge uniformity, and the pressure losses due to turbulence.

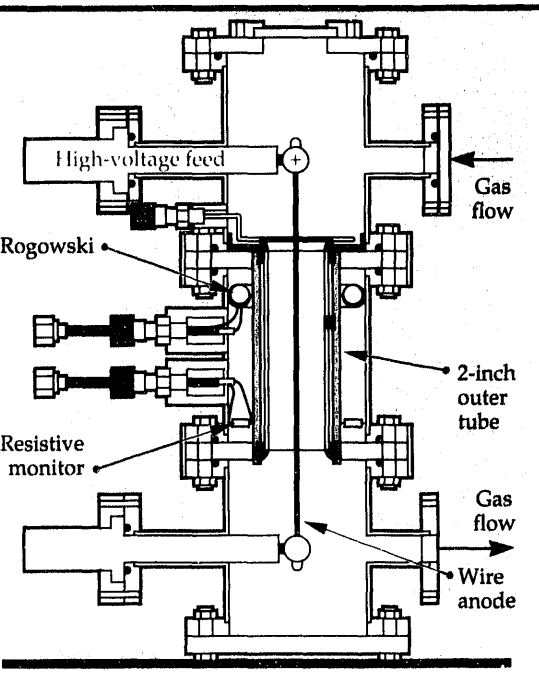


Figure 1. Processor for removing effluent pollutants.

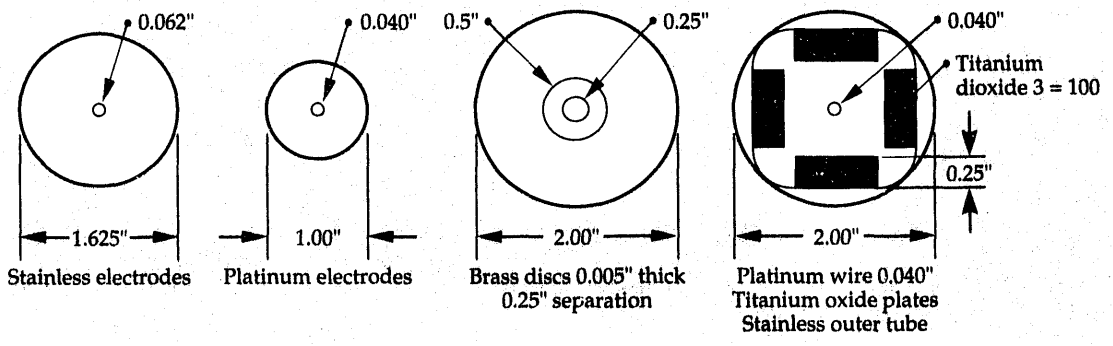


Figure 2. Electrode geometry for processor.

The processing chamber has been designed with an outer pipe two inches in diameter. This tube can be used as an electrode; other geometries of smaller dimensions can be placed inside. The reaction chamber can be increased in length as needed to match the impedance of the high voltage feed to that of the processor, for maximum energy transfer. Configurations tested for NO removal are shown in **Fig. 2**.

Nitric Oxide Removal

We measure the efficiency of removal in eV/molecule. The performance of the removal in ev/NO molecule is a function of the NO concentration. We are presently charging the system to approximately 600 ppmv (parts per million by volume). NO reacts with itself in the presence of air, and the change in concentration is proportional to the square of the concentration. This means that the natural rate of NO reduction at 500 ppmv is 25 times that at 100 ppmv. This effect must be subtracted from the reduction due to pulse plasma processing. The natural reduction of NO has a negative temperature coefficient, which means the reduction is less at higher temperatures. At room temperature, the natural decay at 500 ppmv is approximately equivalent to 20-pulses-per-second pulse plasma processing. We feel the present system gives good data to the 500 ppmv level and can go to higher ppmv at higher temperatures. The efficiency of NO removal has shown to be sensitive to concentration. **Figure 3** shows this effect. It appears that the removal of NO_2 increases once the NO has been removed.

Initial measurements will be made with the closed-loop system; however, we intend to convert this system to a flow-through system, which will permit steady-state mixing.

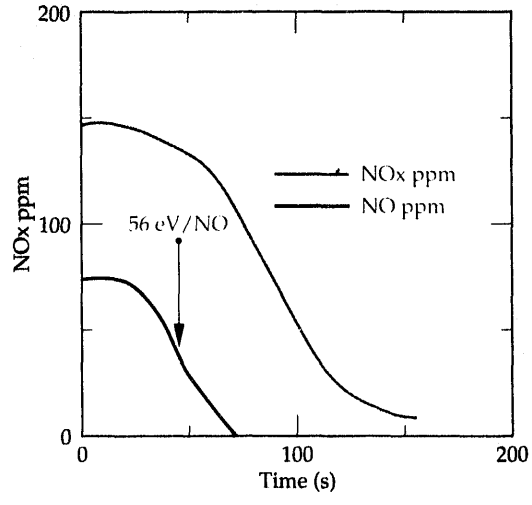


Figure 3. Nitric oxide removal. Plots show data from NO and NO_x processing.

Additives: N-octane and Water

The addition of n-octane has improved the efficiency of NO removal. **Figure 4** shows efficiency improvements with 0, 1850 ppmv of n-octane. N-octane is similar to gasoline and has a flammability limit of 8000 ppmv at room temperature. The addition of n-octane, suggested by R. Atkinson,[1] uses a process that effectively burns organics by recycling the OH radical. This process should be possible with many organic compounds. It seems likely that the work by Fujii,[2] using an oil that is vaporized in the processing chamber, is a similar process.

Tests to date have been with dry air and with approximately one percent water. The tests with dry air have an efficiency similar to wet air. We plan tests in the near future with water up to eight percent by volume at 200°F.

Diagnostics

Diagnostic systems have been used to analyze the light emissions of the discharge. These include a monochrometer and an open shutter camera. Devices to measure the results of chemical reac-

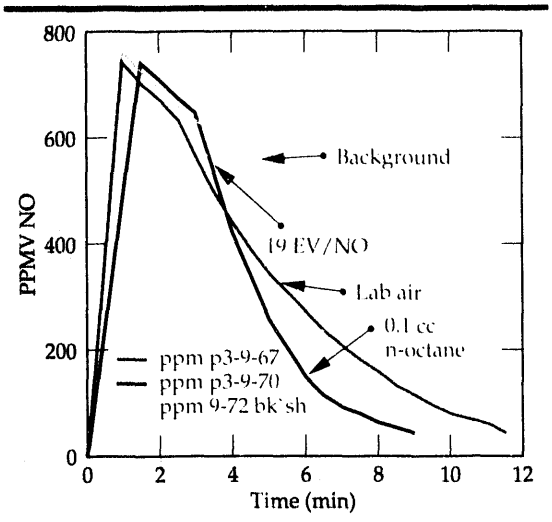


Figure 4. Results from n-octane addition in NO removal.

tions include a chemical NO_x meter, a chemiluminescence NO_x meter, and an IR and FTIR analyzer.

The measurement of energy is essential to determine the efficiency. We measure the energy by recording the voltage and current as a function of time, and then integrate the product. This gives us the joules per pulse. The pulse rate is measured by a counter. The total energy in any time period is then the product of the time pulse rate and joules per pulse. To prevent reflections, a load resistor is included at the end of a short transmission line, where the voltage is measured with a voltage divider. The current is measured with a 0.1-Ω resistor in the return path.

Applications

Diesel Application. The application of this technology to diesel exhaust cleanup poses many challenges due to weight, size, life, cost, and efficiency requirements. We are developing power supplies with similar requirements as part of the Laser Isotope Separation Program at Lawrence Livermore National Laboratory. It may be possible to meet these requirements if the efficiency of NO_x removal is sufficiently improved. Our goal would be to remove the required NO_x with less than two percent of the engine power output. This would require 8 hp or 6 kW for a 400 hp engine.

The efficiency of the overall process is the key to diesel applications. The higher the efficiency, the more likely this process will be practical. It may be that the electrical energy can be generated by thermal-electric generators that use the engine waste heat. These devices would require no loss in engine power to clean the exhaust.

Other pollutants may be removed or destroyed by this plasma process. It is likely that fuel droplets, carbon monoxide, and volatile organic hydrocarbons may be oxidized to harmless compounds. It may also be possible to remove particulates using this process in the presence of liquid droplets.

Coal-Fired Power Plants. This technology can be applied to coal-fired power plants. It may be possible to simultaneously remove NO_x, SO_x, mercury, and particulates from the effluent. The removal of both NO_2 and SO_x would be accomplished by their reaction with ammonia. This reaction gives ammonium sulfate and nitrate, which can be sold as fertilizer.

The critical issue in this application is cost. It must be competitive with other processes. Improvement of efficiency results in reduced capital costs and operating costs. Our primary goal is improvement of the efficiency of simultaneous removal of NO_x and SO_x.

Other Applications. There are many applications for pulsed plasma processing. These include destruction of volatile organic hydrocarbons, elimination of hydrogen sulfide from fuel gases, and others where plasmas can induce or accelerate a chemical reaction.

1. Roger Atkinson, *Chem. Rev.* **85**, 69 (1985).

2. K. Fujii, 19th Int. Conf. Phenomena in Ionized Gases (Il Ciocco, Italy), 1991.

Ground Penetrating Imaging Radar for Bridge Inspection

John P. Warhus, Scott D. Nelson, and Jose M. Hernandez
Defense Sciences Engineering Division
Electronics Engineering

Hua Lee and Brett Douglass
Electronic and Computer Engineering Department
University of California Santa Barbara

Erik M. Johansson
Laser Engineering Division
Electronics Engineering

We have developed conceptual designs, completed requirements analyses, and performed experiments, modeling, and image reconstructions to study the feasibility of improving ground-penetrating imaging radar technology for efficient and reliable nondestructive evaluation of bridges and other high-value concrete structures. In our feasibility study, we made experimental measurements of frequency-dependent electrical properties of cement, from which we derived an electromagnetic (EM) model for concrete, to use in system-level simulations. We performed parameter studies to evaluate key system design issues, using two- and three-dimensional, finite-difference, time-domain EM analysis codes to simulate an ultra-wideband synthetic aperture radar and produce simulated radar data for a variety of concrete structures. Images produced from simulated radar data were analyzed to evaluate important radar system performance parameters and characterize imaging algorithms we are developing.

Introduction

Ground-penetrating imaging radar (GPIR) radiates very-short-baseband (i.e., without a high frequency carrier) electromagnetic (EM) pulses into ground media such as soil and concrete to probe for features of interest, without disturbing the media. This technology is attractive for use as a bridge inspection tool because it is non-contacting and can produce high-resolution reconstructed images of imbedded structural features using a vehicle moving at highway speeds. However, the full capability of the technology has not been exploited at a commercial level. Limitations preventing current ground-penetrating radar (GPR) systems from being more widely used for bridge inspection include difficult data interpretation (no image reconstruction); inaccurate depth and position measurement; relatively poor spatial resolution; and limited area coverage, which limits operating efficiency.

In an improved bridge inspection GPR system, a mobile ultra-wideband (UWB) radar gathers data for high-resolution image reconstruction of features and defects embedded within the structure. Performance enhancements are achieved by increasing transmitted pulse bandwidth and power, using receiving antenna arrays and synthetic aperture radar processing techniques, and adding high-resolution imaging.

An advanced GPIR and imaging system has the near-term (2- to 4-year) potential of addressing critical national and international needs for reliable, cost-effective nondestructive evaluation of bridges and other reinforced concrete structures. There are more than 578,000 highway bridges in the U.S., and more than 40% of them are either structurally deficient or functionally obsolete.[1] These conditions limit usefulness and can pose a safety threat to the bridge users if the bridges are not properly monitored and maintained.

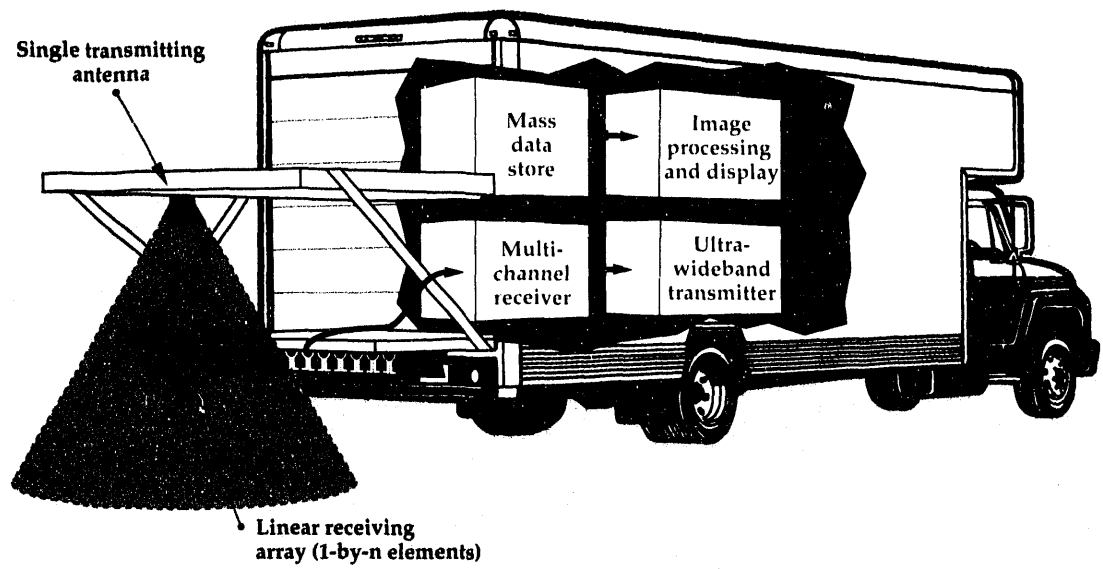


Figure 1. GPIR bridge inspection concept.

The bridge deck and its wearing surface are the most vulnerable parts of any bridge, undergoing damage from routine service. They are particularly well suited for inspection using a vehicle-mounted inspection system. The deck has a shorter average service life (35 years) than the bridge itself (68 years). The wearing surface, which provides the driving surface and protects the deck beneath it, is usually designed to be replaced many times over the life of the bridge. Concrete slabs with concrete or asphalt cover are the most widely used decks and wearing surfaces in all types of bridge construction.[2]

Our approach in this study has been to use system-level design, supported by experiments and analytical modeling, to evaluate key system performance parameters and requirements and to determine feasibility.

Progress

During FY-92, our efforts were aimed at defining requirements for improving GPIR performance and evaluating the capability of available technologies to satisfy those requirements. We developed an overall system design concept for an improved inspection system and analyzed its requirements. We investigated the electrical properties of cement (a key constituent of concrete) to develop a model for concrete and to gain insight into imaging enhancement and correction issues. We modeled radar systems and concrete target structures to simulate and evaluate interactions of UWB pulses with clutter sources (aggregate in the concrete, concrete surface reflections, antenna cross-coupling) and targets (structural features like reinforcing bars, or flaws like voids or delaminations). Using analytical capabilities that were improved during FY-92, we also conducted parametric studies to evaluate imaging resolution and performance issues, especially with respect to dispersion effects.

Our FY-92 work showed that current technologies can provide the performance required to implement an improved GPIR with some limited needs for technology development. Additional development work is required in image reconstruction and enhancement, UWB antennas and arrays, and low-power, high-repetition-rate UWB transmitters. Our assessment of these developmental needs indicates a high probability of success in achieving project goals.

System-Level Requirements Analysis

To establish a design baseline for our feasibility study, we formed a basic system-operational concept from which a system conceptual design was developed and performance requirements were defined. Basic operational concept guidelines included: (1) the inspection vehicle moves over the bridge deck at a speed of at least 30 mph; (2) data is acquired for one traffic lane-width of bridge deck with each pass of the vehicle; (3) bridge deck structures are inspected to a depth of 0.5 m; (4) images are reconstructed in three dimensions, with resolution on the order of 50 mm; and (5) image reconstruction is done off-line, at rates permitting 10 to 20 bridges to be covered per day.

As shown in **Fig. 1**, a transmitting antenna and a linear array of receivers travel over the bridge

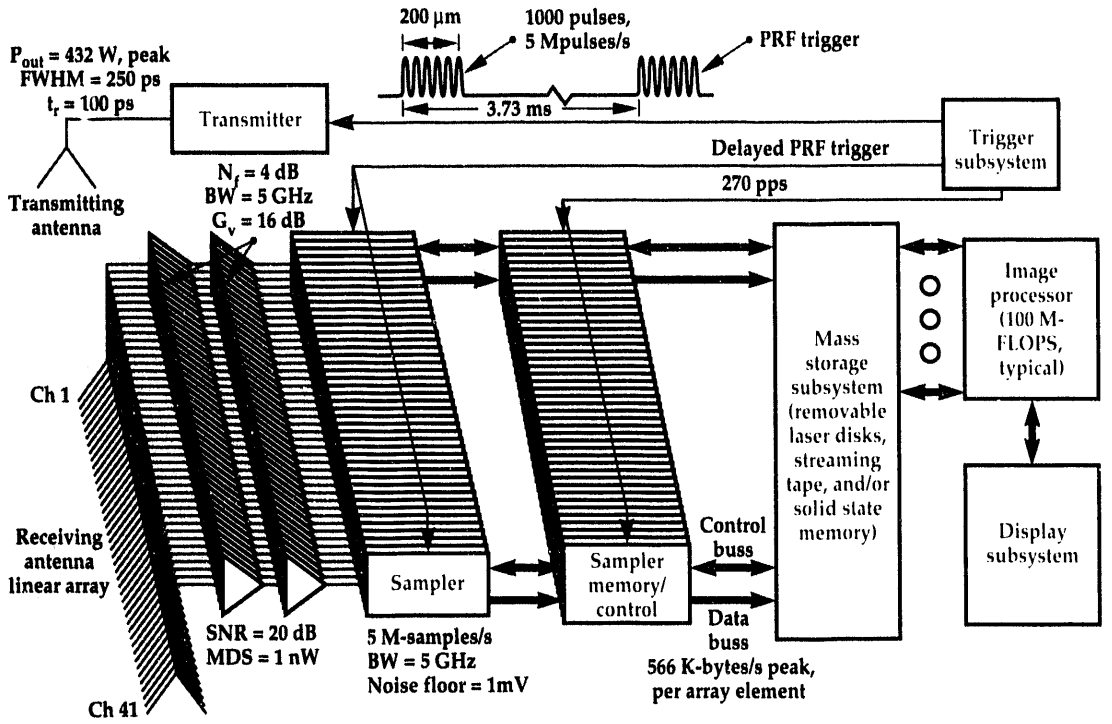


Figure 2. Improved GPIR block diagram.

deck surface, sweeping out a traffic lane-wide synthetic aperture. Data recorded from the receivers is transferred via multiple data streams to a mass data storage subsystem, from which it can be accessed for image reconstruction. Image processors in the vehicle or at centrally located processing centers, reconstruct three-dimensional (3-D) images of the bridge deck structure for evaluation by a bridge inspector. Images for a bridge 100 m long and four lanes wide are reconstructed in less than an hour.

Mobile data acquisition, at speeds approaching highway speed limits, will permit efficient and cost-effective evaluation of large areas of bridge decks in very short times. Evaluation of reconstructed images produced from the radar data will allow bridge inspectors to determine bridge conditions, and prioritize maintenance and repair activities and expenditures on the basis of high-quality inspection data.

A more detailed system-level concept is illustrated in block diagram form in **Fig. 2**. Key system requirements that we identified and defined for this design include: receiver dynamic range and minimum discernible signal; the number of receiving channels (and array elements) required to achieve the desired traffic lane-width coverage and image resolution; peak and average transmitter power; transmitted waveform characteristics; transmitting antenna characteristics and pulse repetition frequency; data acquisition and transfer rates; and computational power required to provide efficient image reconstruction turn-around.

Material Characterization

To better understand the problem of collecting radar data and producing images of features embedded in a lossy heterogeneous material like concrete, we performed broadband (0.1 to 4 GHz) S-parameter measurements of transmission (attenuation) and reflectivity of cement samples. These measurements were made with a network analyzer and a coaxial line in which the dielectric material surrounding the center conductor was formed from cement. From the S-parameter data, we calculated the complex dielectric constant. Many measurements were made over a period of about nine months to observe variations of these properties as the cement cured.

Figure 3 shows typical results of measurements and calculations for a cement sample at eight and 204 days after it was poured. The decreases in relative dielectric constant (ε_r) and attenuation over time reflect the reduction of the amount of free water within the curing cement. Important consequences of the frequency dependence of ε_r and attenuation are that concrete is dispersive and acts as a bandpass filter. Dispersion distorts the propagating and scattered EM waves in the media by reducing risetime and increasing pulse width, and attenuation reduces the effective bandwidth of the

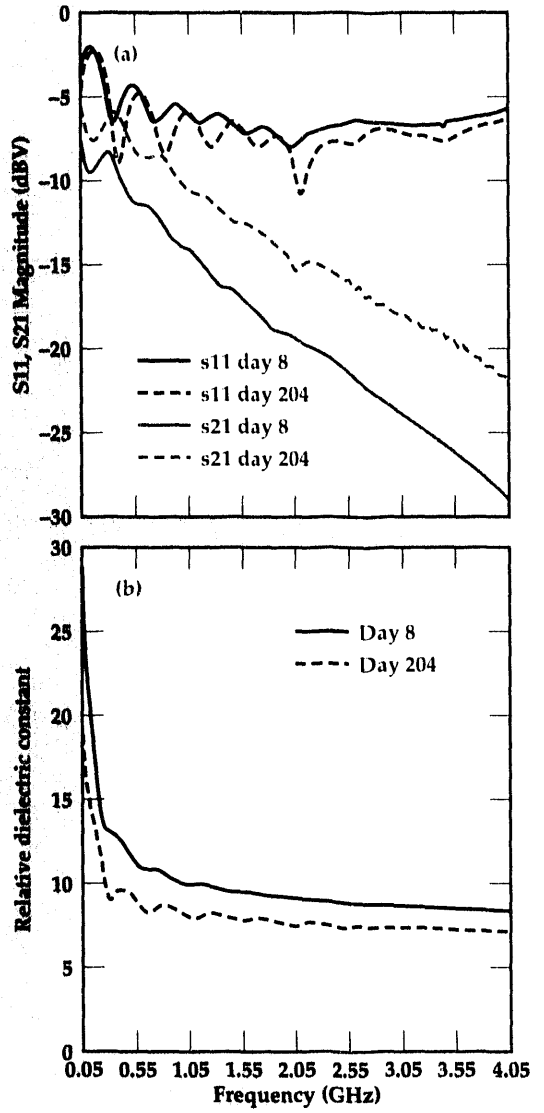


Figure 3. ***(a) Measured transmission (S21) and reflection (S11), and (b) calculated relative dielectric constant (ε_r) for cement sample.***

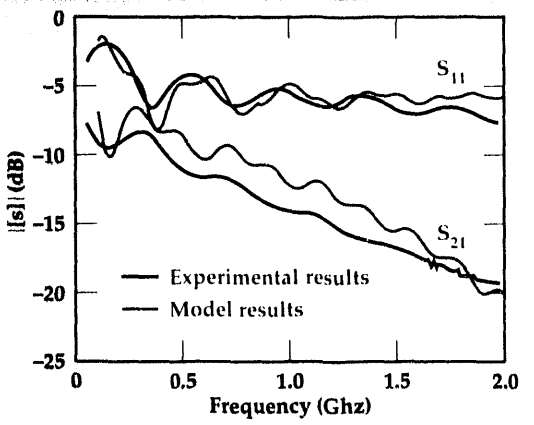


Figure 4. Results from (a) 1-D experimental data and (b) 1-D cement model.

EM energy. These effects degrade the resolution of images reconstructed from the UWB radar data and cause errors in depth measurement.

Data obtained from these measurements were useful in defining the dominant electrical parameters of concrete required for accurate modeling using analytical codes like AMOS or TSAR. In addition, knowledge gained from these measurements provided insights into ways in which correction and enhancement of radar data can be achieved.

Modeling and Image Reconstruction

A key element in our studies of system performance requirements was EM modeling of radar system components and targets. Our modeling requirements included needs for one-, two-, and three-dimensional simulations. Those requirements covered a variety of issues important to our feasibility study, including: developing and validating an EM model for concrete; modeling complex bridge-like structures to support our image reconstruction algorithm development effort by providing a means to evaluate algorithm performance; and performing trade studies that examined radar system configuration options and performance parameters.

To satisfy our requirements, we used two finite-difference time-domain EM analysis codes, which were developed and are maintained by Lawrence Livermore National Laboratory. Those codes, AMOS and TSAR, permitted us to evaluate a wide range of technical issues, without requiring the investment of limited project resources to produce physical hardware or execute numerous experiments. In support of our need to model a dispersive material, both codes were upgraded to permit modeling EM wave propagation and scattering in a medium whose dielectric properties are frequency dependent.

By combining results from early material characterization experiments with analytical modeling tools, we developed an EM model for concrete. Permittivity data derived from material characterization measurements were used in a one-dimensional (1-D) model to simulate those measurements. **Figure 4** is a plot of measured data overlaid with results from the 1-D model, showing good agreement between measurements and the model.

To extend these results to two dimensions, an additional experiment was performed. UWB pulses were launched through a concrete block, using a broadband antenna, and detected and recorded on the other side of the block with a UWB sensor and recording system. A two-dimensional (2-D) model of the experiment was constructed and run using AMOS. The model used the complex permittivity data from the 1-D case, and included

such details as the antenna beamwidth, radiated electric field waveform, dimensions of the block, dispersive effects of the cement, and clutter effects produced by the aggregate within it. **Figure 5** shows plots of both the experimental measurements and simulated results. Again, fairly good agreement between experiment and simulation provided validation of the concrete EM model.

After confirming the validity of our 2-D EM model for concrete, this analytical tool was used extensively to evaluate radar system design parameters and image reconstruction algorithm performance. Parametric studies we conducted using this model are summarized in **Table 1**. Image reconstructions were made from the simulated radar data for most of the cases listed in the table. The images aided assessments of the impacts of design parameter changes on overall system performance and complexity, and in evaluating image quality for specific imaging techniques. In addition to the studies listed in the table, 2-D simulations permitting evaluation of air/concrete boundary and antenna cross-coupling effects, and of the impact of using multiple transmitters as well as multiple receivers were also run. These studies helped to confirm conceptual design conclusions and permitted us to consider other more complex system configurations.

An example of results from one important study is shown in **Fig. 6**. A series of simulations was run in which a single target of fixed size was embedded at increasing depths within the concrete. The purpose of this series was to evaluate imaged range and cross-range resolution of the reconstructed target and to assess the accuracy of the position location of the target after corrections for dispersion effects had been made. The image sequence was reconstructed using a multi-frequency holographic method, which includes a correction for the effects of dispersion.

Table 1. Model-based parametric studies.

Parameter	Variations
Radiated pulsewidth	100 to 1000 ps
Array element spacing, spatial sample rate	5 to 45 mm
Target cross-section size	5 to 75 mm
Target depth	10 to 150 mm
Clutter source (aggregate) density	10 to 50%
Target density/type	No targets, 1 void, 2 reinforcing bars (rebars), 2 rebars plus 1 void, a rebar grate, and rebar shaded from EM radiation by other rebars

The degrading effects of dispersion and frequency-selective filtering on resolution are clearly shown in these images. As predicted from material characterization measurements and analysis, while the target's dimensions remain fixed, its image increases in size in both down-range and cross-range directions as its depth increases. However, careful analysis of the images shows that the target depth, as indicated by the peak of the image intensity, is located quite accurately when dispersion corrections are applied. The only exception is the case very close (10 mm) to the transmitter and receiving array. The most likely cause for the error, in that case, was EM wave interference that occurred because the reflected pulse from the target was incident on the receiving antennas, while the tail of the transmitted pulse was still propagating past the receivers.

The three-dimensional (3-D) EM modeling effort was started late in the year, to support testing

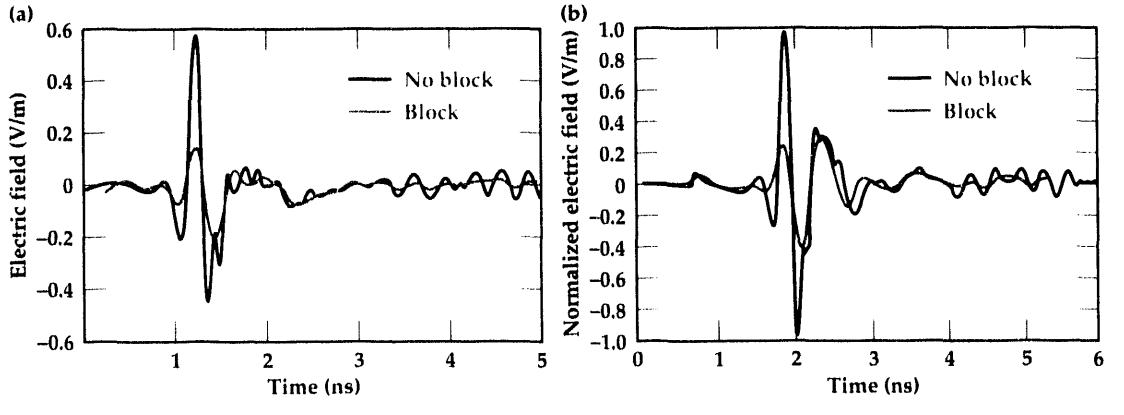


Figure 5. Results from (a) 2-D experimental data and (b) 2-D concrete model.

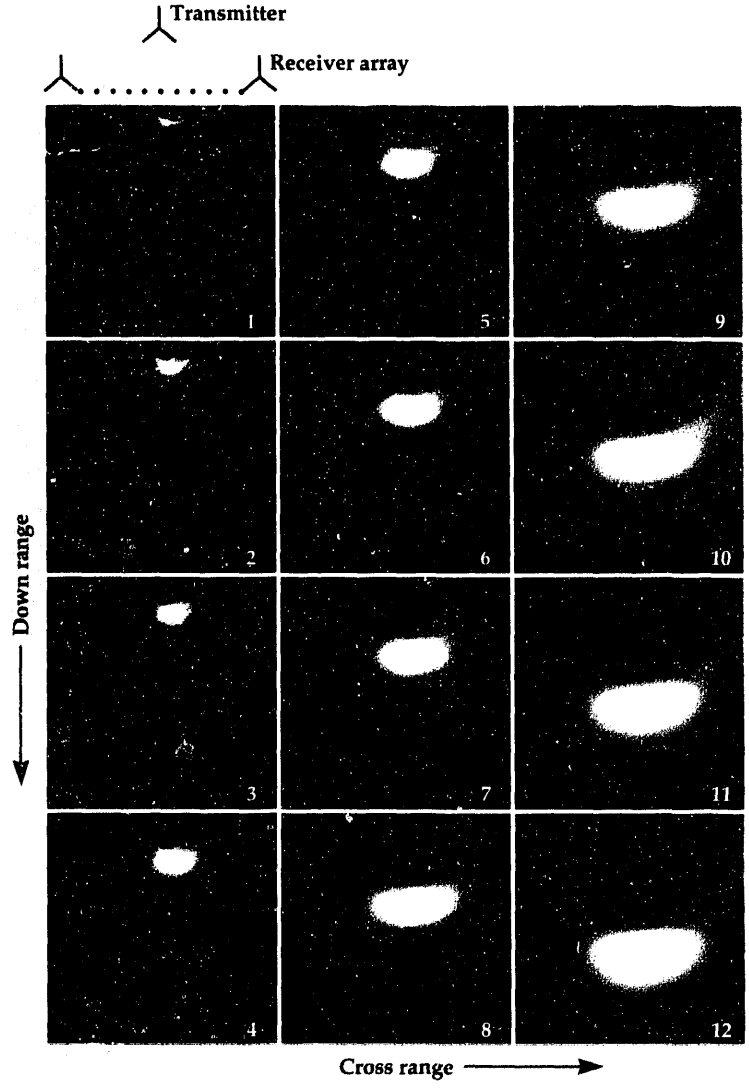


Figure 6. 2-D image sequence of rebar target at increasing depths.

Simulated depth (mm)	Reconstructed image depth (mm)
10	16
20	21
30	31
40	41
50	50
60	60
70	69
80	80
90	89
100	100
110	111
120	120

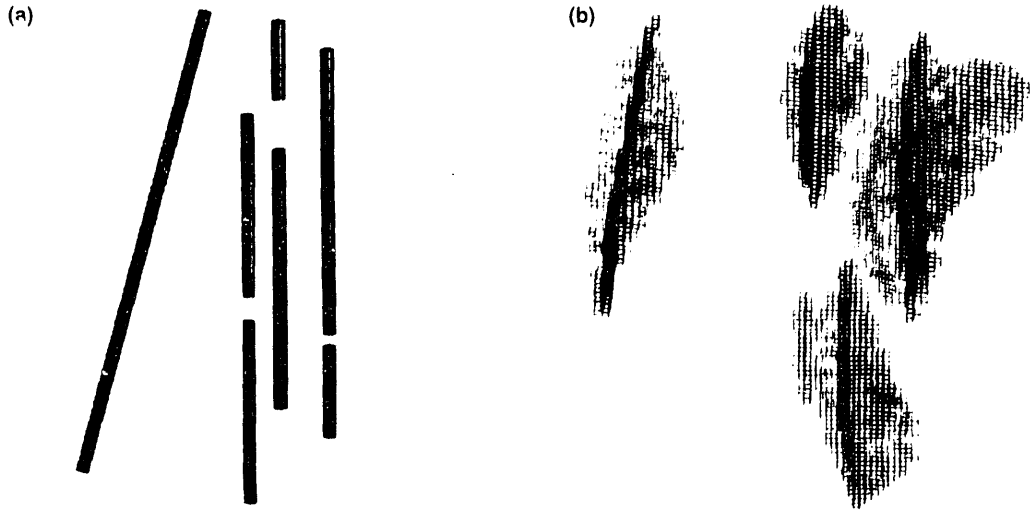


Figure 7. (a) Physical model of four rebars in free space and (b) volumetric rendering of 3-D image reconstruction.

of 3-D image reconstruction software. Very simple physical models were used in the early tests. An example of a test case is shown in **Fig. 7.** in which four rebars, three of which have gaps, are assembled in free space. This simple model permitted evaluation of the imaging software without including the complicating effects of clutter, dispersion, and filtering. After the image was reconstructed, it was processed and enhanced with some rudimentary techniques to provide a means for viewing the 3-D rendering as shown in the figure.

Images of the rebars are clearly visible in the rendering; however, the gaps in the bars are not. The gaps are not seen in the rendering because they did not produce any reflection of the EM energy launched in the simulation, and the energy scattered from the rebars tends to fill in the voids. In the case where air-filled gaps in rebars are embedded in concrete, we do expect to detect and image the gaps because the air/concrete interface at the gap will produce a significant phase reversed reflection.

Future Work

Our continuing efforts are aimed toward a field demonstration of a limited-capability prototype system late in the next fiscal year. To support that effort, we will complete a series of experiments to confirm key modeling results and verify system design parameters. Those experiments will be conducted using a concrete test slab that was designed to simulate bridge construction features, and in which we embedded several flaw simulants. UWB antenna and transmitter development will be pursued, and image reconstruction algorithm development, testing, and refinement will continue with the goal of having optimized radar hardware and imaging code available late in the year to support the demonstration. A low-cost prototype system will be designed to permit demonstration of base data acquisition and image reconstruction performance. The objective of the demonstration will be to show improved performance in resolution, and accurate reconstruction of embedded structure.

Acknowledgements

We wish to thank Jim Brase, Remote Sensing, Imaging, and Signal Engineering Thrust Area Leader, and John DeFord, Computational Electronics and Electromagnetics Thrust Area Leader, for their support in supplying the resources needed to develop and evaluate imaging techniques, to perform EM modeling, and to enhance the capabilities of EM modeling codes for this project.

1. *Our Nation's Highways: Selected Facts and Figures*, U.S. Department of Transportation, Federal Highway Administration, Publ. No. FHWA-PL-90-024.

2. N.P. Jones and B.R. Ellingwood, "NDE of Concrete Bridges: Opportunities and Research Needs," Federal Highway Administration Conf. on NDE for Bridges (Arlington, Virginia), (August 25–27, 1992).

High-Average-Power, Electron Beam-Controlled Switching in Diamond

W. Wayne Hofer
Defense Sciences Engineering Division
Electronics Engineering

Karl H. Schoenbach, Ravindra Joshi, and Ralf P. Brinkmann
Old Dominion University
Norfolk, Virginia

Don R. Kania
Inertial Confinement Fusion Program
Laser Programs

The superior electronic and thermal properties of diamond make it an ideal material for a high power solid-state switch. Our FY-92 goals were to identify and address technical issues that could potentially limit the anticipated performance of electron beam-triggered, high power switching in diamond. In particular, we concentrated on the role of contacts and non-linear effects at high electric fields, electron beam range in diamond, and carrier transport modeling.

Introduction

The superior electronic and thermal properties of diamond make it an ideal material for a high power solid-state switch. We predict that an electron beam-controlled diamond device could switch well over 100 kW average power, at megahertz repetition rates, with greater than 95% efficiency and voltages greater than 5 kV.

High power diamond switches could significantly increase the performance of high power-switched power supplies, modulators, and power converters. Commercial applications include high power radar, control for electric vehicles, high power industrial controllers, and possibly solid-state switching at utility substations.

The crystal structure of diamond is relatively well characterized. It is a semiconductor with a band-gap of 5.5 eV at 300K. By comparison, the band-gap of GaAs is 1.4 eV. The high band-gap of diamond results in a small dark current compared to Si or GaAs. As a result, the breakdown field or holding voltage is very high, i.e., 1–10 MV/cm. The electron and hole mobility are approximately 2000 cm^2/Vs. At room temperature, diamond has the highest thermal conductivity of any solid, 20 W/K cm, about five times that of copper. The electronic properties of chemical vapor desposition(CVD) diamond now exceed those of the best natural diamond (**Table 1**). CVD diamond substrates can be cooled using microchannel cooling, a highly effective thermal management technology developed at Lawrence Livermore National Laboratory (LLNL). When the electronic proper-

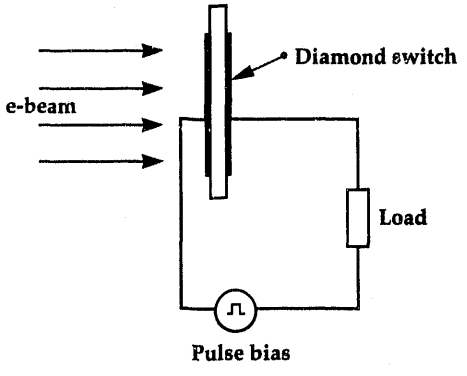


Figure 1. Electron beam-controlled switch. In our switch, 100 keV electrons are absorbed in a thin diamond film, and by ionization, generate a high concentration of electric carriers in the diamond.

Table 1. Electronic properties of diamonds.

	Microwave films	Single-crystal natural II-A	Homoepitaxial film
Thickness	25–500 μm	$1 \times 1 \times 3$ mm^3	120 μm
Grain size	10–100's μm	single-crystal	single-crystal
Electrical resistivity (Ω-cm)	10^5–10^{11}	$> 10^{11}$	$> 10^{11}$
Lifetime (ps)	100–800	100–1000	150
Mobility (cm^2/Vs)	500–4000	2800	3000
Raman	1330–1334 FWHM 4–9/cm	1332.4 peak (FWHM 2.4/cm)	1332/cm peak (FWHM 2.9/cm)

ties of diamond and its superior thermal properties are combined with microchannel cooling and the rapid advance in CVD technology, diamond becomes an excellent solid-state material for advanced, high performance power electronics.

Progress

In our switch (**Fig. 1**), a high concentration of carriers (corresponding to kA's/cm^2) are created by ionization when a high voltage, low current electron beam is absorbed in the diamond film. The controlling electron beam can be generated and modulated at megahertz frequencies by compact, long-life, commercially available grid-controlled thermionic cathodes. As an alternate to thermionic cathodes, recent research and development of new high-current-density electron sources, such as ferroelectric cathodes, micro-field emission cathodes, and even diamond field emission cathodes, could potentially be combined with a thin film diamond to create a very compact, robust, high power solid-state diamond switch.

Our FY-92 goals were to identify and address technical issues that could potentially limit the anticipated performance of electron beam-triggered, high power switching in diamond. In particular, we concentrated on the role of contacts and nonlinear effects at high electric fields, electron beam range in diamond, and carrier transport modeling. Our work is a combined effort with researchers at Old Dominion University at Norfolk, Virginia, and at LLNL.

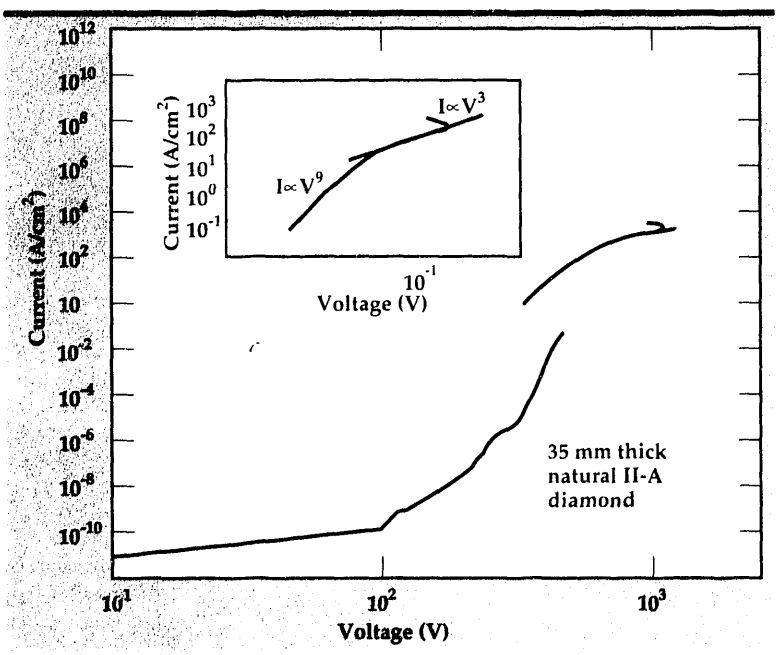


Figure 2. Dark current vs applied voltage for 35-μm natural II-A diamond film. The current increases rapidly once the trap-filled limit is reached (~200 kV/cm). Similar results were obtained for 50-μm-thick diamond, but the dark current increased rapidly at about 400 kV/cm.

Contacts and Nonlinear Effects

Diamond is normally thought to be an excellent high voltage insulator. However, we have shown that for both natural and CVD diamond on silicon, the dark current (no electron beam-generated carriers) increases by 10–11 orders of magnitude at high electric fields.

For natural II-A diamond, the strong nonlinear increase starts at 200-kV/cm for 35 μm-thick diamond (**Fig. 2**) and at 400-kV/cm for 50 μm-thick diamond. We believe that this threshold corre-

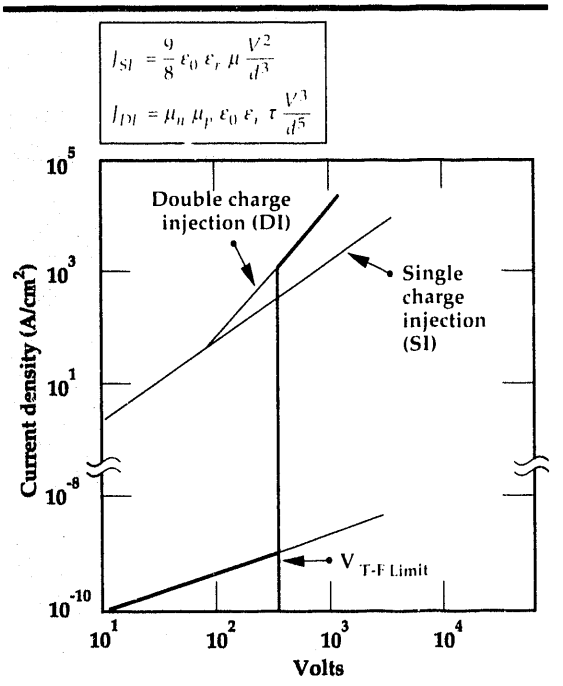


Figure 3. Drift-diffusion modeling results. A model based on charge injection qualitatively predicts the rapid increase in current at high electric fields.

sponds to the trap-filled limit. The thickness dependence correlates well with photoelectronic theory where the trap-filled-limit electric field is proportional to the square of the material thickness. As the field is further increased, dark current is dominated by charge injection at the contacts. Both our qualitative and more comprehensive drift-diffusion modeling results (**Figs. 3** and **4**) predict these results. At even higher fields, a negative differential resistivity phase quickly leads to de-

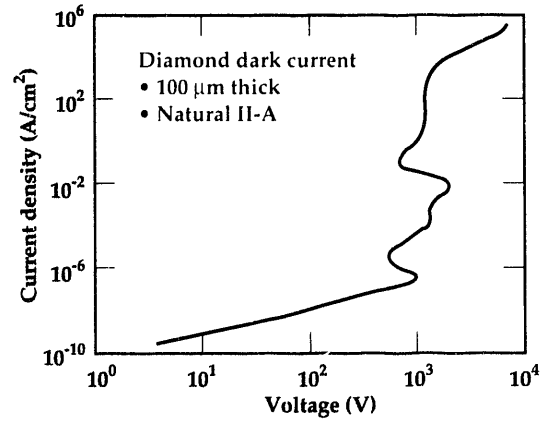


Figure 4. Drift-diffusion modeling results. Data from our drift-diffusion model with injecting contact and deep trap centers agree with the behavior observed qualitatively and experimentally.

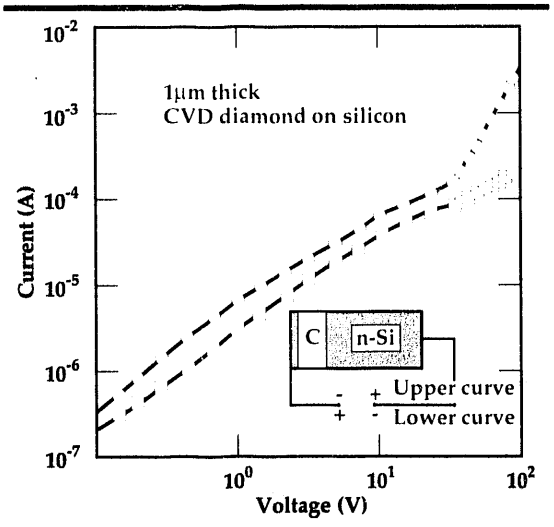


Figure 5. Dark current vs voltage for 1-μm-thick CVD diamond on n-silicon. The onset of rapidly increasing dark current depends on the polarity of the bias. When the silicon is biased negative, the hold-off voltage is highest.

struction of the switch, most likely due to current filaments.

Similar to natural diamond, the dark current in CVD diamond on a heavily doped (~10^{18} cm^{-3}) n-silicon increases rapidly at high fields (**Fig. 5**). However, the electric field threshold was in excess of 0.9 MV/cm, much higher than that for natural diamond, and it depends strongly on the bias polarity. Apparently silicon is a very

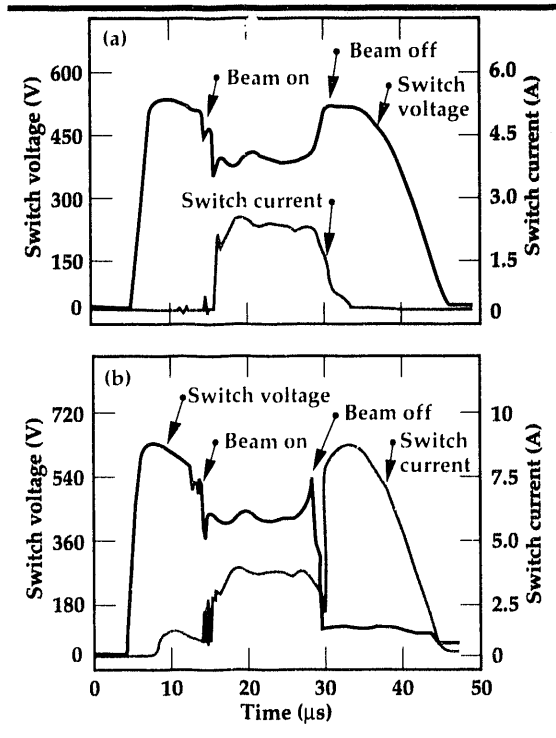


Figure 6. Voltage vs time for electron beam-switched, 35-μm-thick, natural II-A diamond at two different bias levels. At the lower field (a), the conductivity follows the electron beam, but at the higher field (b), the current continues after the beam is turned off, due to charge injection.

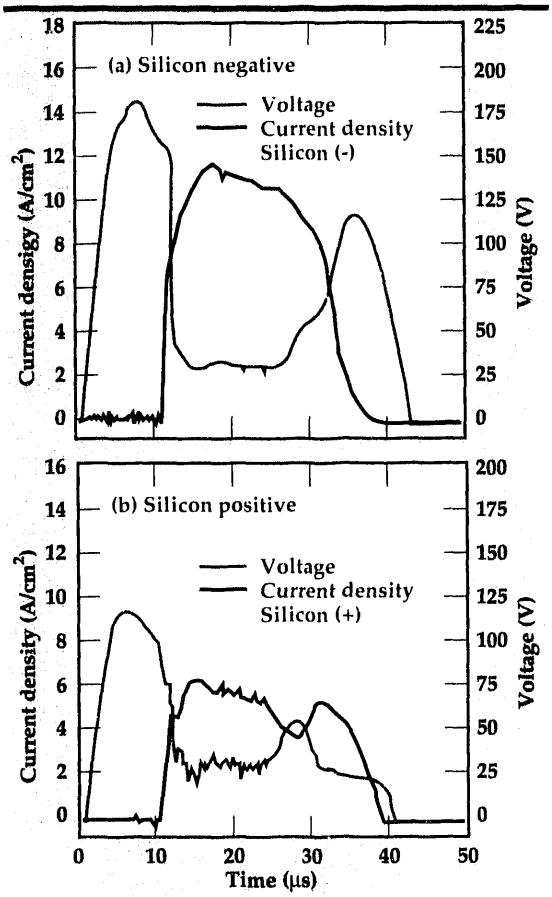


Figure 7. Electron beam-induced conductivity in 1-μm-thick CVD diamond on n-silicon for opposite polarities. When the n-silicon is negative (a), the switch conductivity follows the beam current. However, when the n-silicon is positive (b), the switch conductivity persists after the beam is turned off.

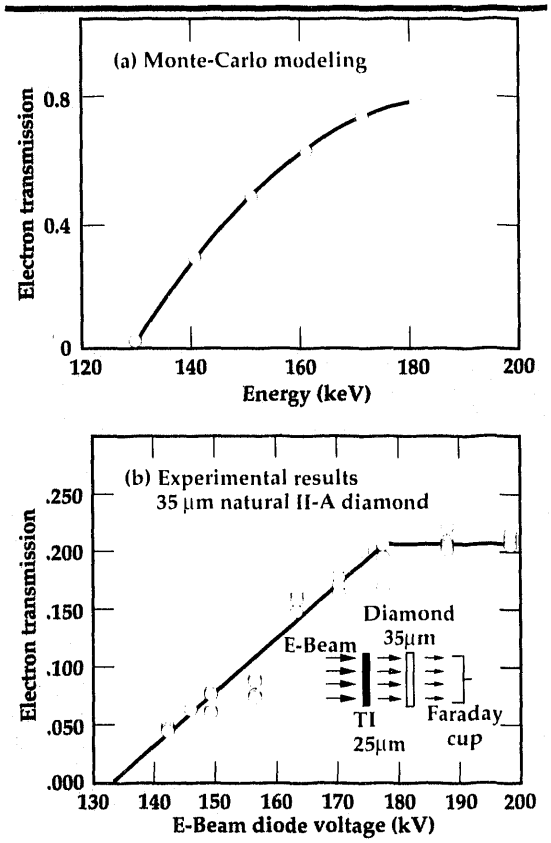


Figure 8. (a) Monte-Carlo calculations of 150-keV electron beam penetration in 35-μm diamond. Scattering effects in a 25-μm Ti anode foil are included. (b) Transmission of the electron beam in a 35-μm-thick natural diamond film. The electrons are completely absorbed at about 130 keV.

good blocking contact.

Dark current in diamond is highly dependent on carrier recombination and trapping at deep energy levels in the diamond and whether contacts are blocking or injecting.

Electron Beam-Induced Conductivity

When the natural diamond switch is irradiated with an electron beam, the switch conductivity follows the beam profile at lower electric fields; but at higher fields, the switch remains conducting when the beam is turned off (**Fig. 6**). At even higher fields, the switch failure is likely due to filamentary current. We believe this corresponds to operation in the negative differential resistivity mode.

When irradiated with an electron beam, the conductivity of the CVD diamond follows the beam profile (**Fig. 7**) up to 1.8 MV/cm when the silicon is biased negative. However, when the silicon is biased positive, the CVD diamond remains conducting even when the beam is turned off.

Based on these data, it appears that CVD diamond on silicon may limit carrier injection, thus enabling us to switch higher voltages.

Based on our current data, we should be able to switch voltages ranging from 400 to 4000 V. If the lock-on field can be extended 5–10 times by tailoring the diamond growth and contacts, we may be able to produce an on-off switch operating at 20 to 40 kV.

To hold-off and switch higher voltages, we must extend the onset voltage threshold of the rapidly increasing dark current by using better blocking contacts.

Electron Penetration Depth

Electron penetration depth determines the maximum diamond thickness for a given electron beam energy. The diamond thickness in turn determines the maximum hold-off voltage.

Modeling results show that the penetration depth for 150 kV electrons in diamond is about

35 μm (**Fig. 8a**). These results are confirmed by our experimental data (**Fig. 8b**).

Future Work

To switch the highest possible voltage and power, we must extend the electric field threshold of the nonlinear increase in dark current. This will be accomplished with non-injecting contacts (blocking contacts) and by understanding and controlling deep-trap center impurities in the diamond switch. We will concentrate our efforts on CVD diamond grown on highly doped silicon and determine how voltage hold-off scales with diamond thickness.

We will obtain data on deep-center impurities and defects in diamond by using Electron Beam-Induced Current Transient Spectroscopy (EBICTS). EBICTS spatially resolves the activation energy and density of deep-level traps. These data will be used in our drift-diffusion models that calculate carrier transport in diamond at high electric fields.

Finally, we will construct a prototype switch to demonstrate kilovolts switching in diamond and a range of hundreds of amperes.

1. R.H. Bube, *Photoelectric Properties of Semiconductors*, Cambridge University Press (Cambridge, England), 1992.

Testing of CFC Replacement Fluids for Arc-Induced Toxic By-Products

W. Ray Cravey
Defense Sciences Engineering Division
Electronics Engineering

Ruth A. Hawley-Fedder and Linda Foiles
Condensed Matter and Analytical Sciences Division
Chemistry and Materials Science Department

Wayne R. Luedtka
Computer and Communication Engineering Division
Electronics Engineering

We have developed a unique test-stand for quantifying the generation of perfluoroisobutylene (PFIB) in chlorofluorocarbon (CFC) replacement fluids when they are subjected to high electrical stress/breakdown environments. PFIB is an extremely toxic gas with a threshold limit value of 10 ppbv as set by the American Conference of Governmental Industrial Hygienists. We have tested several new fluids from various manufacturers for their potential to generate PFIB. Our goal is to determine breakdown characteristics and quantify toxic by-products of these replacement fluids to determine a safe, usable alternative for present CFC's. We are currently working with 3M, DuPont, and Ausimont, key manufacturers of these replacement fluids, to test them for potential PFIB generation.

Introduction

Restrictions on the use of chlorofluorocarbons (CFC's) worldwide, nationally, and at Lawrence Livermore National Laboratory (LLNL) will have an enormous impact on industry and government laboratories. On September 16, 1987, several CFC-producing nations including the United States signed the Montreal Protocol, which called for the phase-out of production of CFC's no later than the year 1997. President Bush, reacting to scientific data showing the imminent danger of an ozone hole over the Northern Hemisphere, ordered the phase-out of CFC's by the year 1995, two years early than the Montreal protocol.[1] Worldwide usage of CFC is estimated at 750,000 metric tons[2] (see **Fig. 1**). Almost half of this amount is used either as cleaning agents in electronic PC board manufacturing (24%), or in the foam blowing industry (another 24%). The majority of the remaining CFC use is in refrigeration and aerosol sprays. CFC's are used extensively in the automobile industry, as well, as a refrigerant for vehicle air conditioners. Although several replacements for existing CFC's do exist, many of these replace-

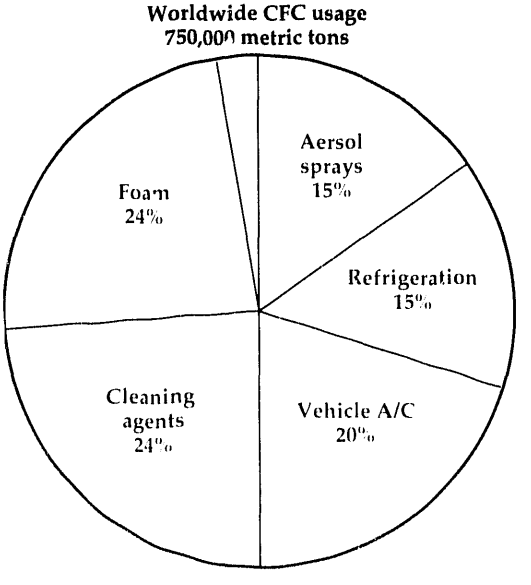


Figure 1. Worldwide use of CFC. Total use is 750,000 metric tons.

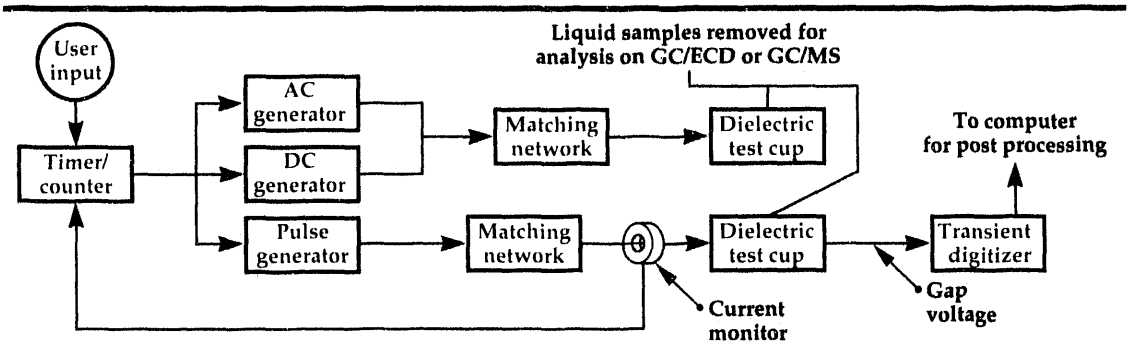


Figure 2. Block diagram of CFC replacement fluid test-stand. The user sets the time or number of pulses for the desired test: AC, DC, or pulse. The output of the corresponding generator is fed through a matching network into the dielectric test cup. The test cup houses the gap electrodes and contains the fluid under test. Samples are removed from the test cup and analyzed on a gas chromatograph (GC) or GC mass spectrometer, to quantify the amount of PFIB that was produced, if any. The voltage across the gap and the coinciding current are digitized and recorded on a computer. There they are analyzed, and the power and energy are determined.

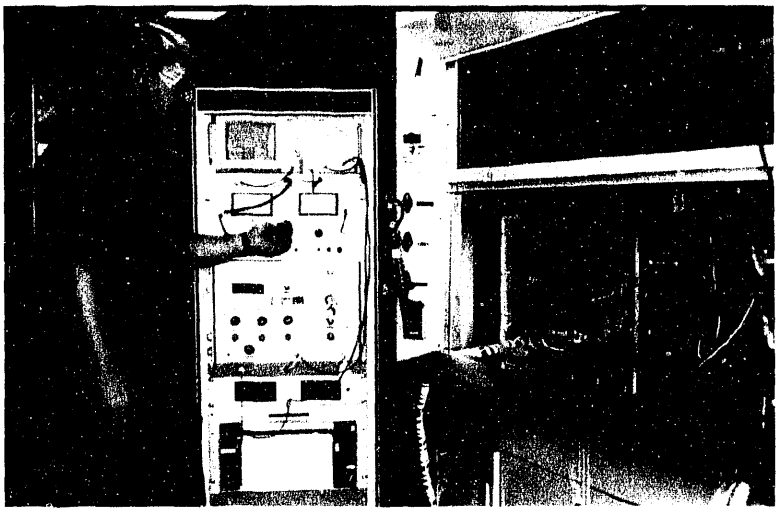


Figure 3. Photo of test-stand.

ments are toxic,[3] flammable, or highly expensive. In addition, two replacements, although not toxic, have been identified as ozone depletion agents, and their use will be phased out in the near future.[4]

We are studying several new replacement fluids for CFC's that would have similar electrical and thermal characteristics. CFC's are currently used as high voltage electrical insulators and dielectric coolants for the high-average-power modulators used to drive the copper vapor lasers at the Atomic Vapor Laser Isotope Separation (AVLIS) facility at LLNL. Several substitutes have been suggested to replace these fluids. But a stumbling block associated with these replacements is the potential to generate toxic gases, such as PFIB, when the fluids are subjected to high electrical stresses and/or breakdowns.

Progress

Over the last year, we designed, fabricated, and tested the new CFC replacement fluid test-stand, which was our foremost goal. Work is proceeding on the compressive data collection and analysis that is necessary to quantify the toxic by-product production for given breakdown conditions. Initial results from tests conducted with the test-stand show a linear trend with respect to arc energy. We have received funding from the Superconducting Super Collider (SSC) for analysis and testing of the cooling fluid used in their low-energy booster (LEB) cavities. A more descriptive summary of our progress is given in the subsequent sections.

Replacement Fluid Test-Stand

We have identified three electrical breakdown/stress environments that may contribute to PFIB production in CFC replacement candidates: AC breakdown; high DC field stress; and pulsed breakdown. Our test-stand was designed to simulate all three of these environments. A block diagram and description of the operation of the test-stand are given in **Fig. 2**; **Fig. 3** is a photograph of the test-stand.

An important test that has application in the refrigeration industry is breakdown due to high AC voltages. We have the capability of producing 60-Hz AC, high voltage breakdown conditions with voltages ranging from 0 to 40 kV, and any desired gap spacing across the test cell electrodes. Our capabilities allow us to simulate the inside environment of the ordinary refrigeration compressor/motor assembly.

High DC field stress occurs in many situations where there are high voltages present. Electrical arcing is not associated with these conditions; however, there can be extremely high fields and corona onset. With our system, we are able to generate these high pre-breakdown fields in the fluid-under-test for any preset time limit. After the given time limit, the fluid is analyzed for PFIB formation.

Pulsed breakdown environments are produced with many high voltage modulators. Typically, these modulators are used for driving radars, lasers, and accelerator cells. We have a wide range of pulsed conditions that can be produced with our present test-stand. In addition to being able to produce the pulsed conditions mentioned, we have implemented a complete electrical diagnostic system for measuring the breakdown voltage, discharge current, arc power, and energy that are associated with each pulse. Through the use of the data that is recorded by the diagnostics, we are able to correlate the quantity of PFIB produced with the pertinent control variables, such as voltage, current, pulse-width, pulse repetition frequency (prf), and energy.

Experimental Results

Measurements have been made on a candidate replacement fluid for the copper vapor laser modulators. The fluid was subjected to a 30-kV pulse at a pulse repetition rate of 75 Hz, with a full-width half-maximum of 100 μs. A 10-mil gap spacing with brass electrodes conforming to ASTM standards[5] was used for testing. The dielectric test cup was filled with 85 ml of the fluid under test. Samples were taken from the dielectric test cup before exciting the fluid, at 20 kV, 40 kV, and 100 kV pulses. Between each sample interval, the voltage and current waveforms were recorded, and the instantaneous power and energy were calculated for the arc. A typical data set is illustrated in **Fig. 4**.

There are two loss mechanisms associated with the arc discharge: resistive/inductive phase[6] losses and conduction loss. The resistive/inductive phase losses are produced during the time the voltage collapses across the gap and the current begins to rise. During this phase, there is a power and energy loss associated with the instantaneous current rise and voltage collapse. The faster the voltage collapses, the less appreciable these losses are. The second loss mechanism, which appears to be the most dominant in this experiment, is the loss associated with the voltage drop across the arc itself. The forward drop of the spark gap was observed to jump between 40 and 60 volts. The maximum output current is 1 amp for the present configuration, limited by the output transformer.

The samples were analyzed on a dual column gas chromatograph equipped with an electron capture detector. Results from the analysis are shown in **Fig. 5**. The reduced data shows a clear trend with respect to energy. We would like to predict the amount of PFIB that is formed for higher energies, by using the data we collect on the test-stand. Initial data looks promising. The calculated field for the breakdown voltage is much lower than what was expected based on the published fluid characteristics. Further investigation showed that the breakdown level is much higher at lower prf due to fluid recovery. Although this is no surprise, it is a key variable in the pulsed breakdown of these fluids.

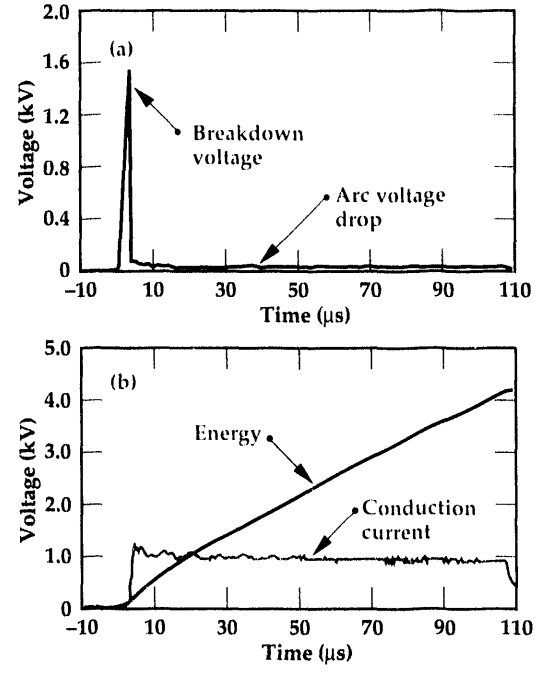


Figure 4. Typical data for (a) voltage and (b) current waveforms.

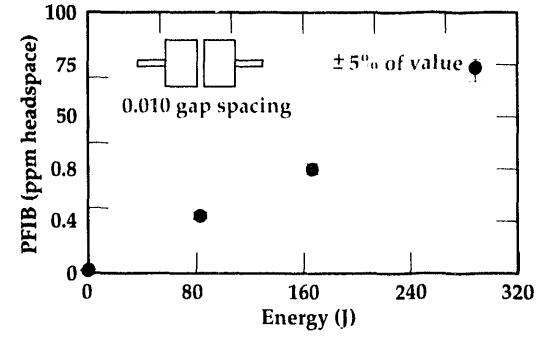


Figure 5. Typical data set for pulse breakdown-induced PFIB generation. The fluid sample was pulsed with a 30-kV, 1-μs pulse at a prf of 75 Hz. This data was generated with brass electrodes and a 10-mil gap spacing in the dielectric test cup. The energy per pulse is 4 mJ.

Superconducting Super Collider Tests

We have been funded by the SSC to test for the amount of PFIB that is potentially produced in their LEB cavities. We have set up a working agreement with SSC to establish PFIB generation data for LEB cavities. Also, we are providing measurement capabilities for quantifying

PFIB in samples they provide. The first sample generated by SSC was subjected to 60 kV for 20 hours of operation. An electrostatic model of the LEB cavity was generated using Ansoft's Maxwell 2D[7], and the maximum electric field in the fluid was calculated to be 83 kV/cm. They reported that during the operation, there were no detectable breakdowns. The samples were analyzed, and no PFIB was measured above the 60 ppbv detection limit.

Future Work

Our plans for the next fiscal year are focused on three key areas.

(1) We are currently working to generate a comprehensive and complete data set for various arc-induced breakdown conditions for various fluids. The last year was dedicated to developing the needed test-stand and diagnostics as well as chemical and analytical techniques. Now that the testing facility is in place, we will fill in our data set and analyze the results for PFIB generation trends, with respect to parameters such as energy, power, and prf.

(2) In addition to the continued testing of the replacement fluids, we propose to test, in cooperation with 3M, a fluid contamination detection system. The system uses a UV source (not specified) and detector to measure the transparency of the fluid. Initial tests have been conducted by 3M, which show that the fluid's transmission changes by as much as 20% when the fluid is subjected to high thermal stress. We will test various detector configuration and breakdown parameters to decide if the detector can be used as a reliable means to signal the possibility of PFIB contamination in hostile working environments.

(3) The third area that has been identified is the refrigeration industry. In the majority of industrial refrigeration systems, the refrigerant is circulated through the compressor motor, for cooling and lubrication.[8] In large systems, the voltage level can exceed 1000 volts, leading to the occasional and often catastrophic electrical breakdown of the fluid. Our test-stand is capable of reproducing these breakdown conditions in the laboratory where a comprehensive analysis can be performed.

1. Statement by Presidential Press Secretary Fitzwater on the Phaseout of Ozone Depleting Substances, February 11, 1992.
2. P. Elmer-Dewitt, "How Do You Patch a Hole in the Sky That Could Be as Big as Alaska?," *Time* **139** (7), (February 17,1992).
3. A.K. Naj, "CFC Substitutes Might Be Toxic, Rat Study Finds," *Wall Street Journal*, July 2, 1991.
4. M. Weisskopf, "Study Finds CFC Alternatives More Damaging Than Believed," *The Washington Post*, February 23, 1992.
5. ASTM Standard D 877-87, *Dielectric Breakdown Voltage of Insulating Liquids Using Disk Electrodes* (1989).
6. J.C. Martin, *Nanosecond Pulse Techniques*, Atomic Weapons Research Establishment, United Kingdom Note 4, 1970.
7. *Maxwell 2D Field Simulator*, Version 4.33, Ansoft Corporation.
8. Althouse, Turnquist, and Bracciano, *Modern Refrigeration and Air Conditioning*, The Goodheart-Willcox Company, Inc., (South Holland, Illinois) 1988.

Applying Statistical Electromagnetic Theory to Mode Stirred Chamber Measurements

Richard A. Zacharias and Carlos A. Avalle
Defense Sciences Engineering Division
Electronics Engineering

We are developing measurement and analysis tools to assess microwave effects on electronic subsystems that operate in large metal cavities, such as avionics boxes in aircraft. The measurement tool is the mode stirred chamber (MSC), which is a metal-walled chamber, large relative to a wavelength, into which electromagnetic energy is injected. It usually contains a stirrer paddle to randomly change resonant mode patterns as a function of time. The analysis tools are based on statistical electromagnetics. This theory predicts that the microwave power measured at an arbitrary point (not near the walls) within an overmoded, randomly complex cavity is a Chi-squared distribution with two degrees of freedom. This is a single-parameter distribution. Therefore, the mean power density measured at an arbitrary point in the cavity is sufficient to develop a complete statistical model of the power at any arbitrary interior point. By showing that a randomized aircraft equipment bay has sufficient Q and ensemble variations to behave as such a random complex cavity, we have determined that the mean coupling measured at a point in the cavity would be sufficient to predict the microwave stress (statistical distribution of fields) to which an avionics box would be exposed over an ensemble of like aircraft (the fleet). A MSC could be used to generate the same distribution for testing avionics boxes uninstalled. This method could provide tremendous cost savings in testing. In FY-92, we developed a small, low-power MSC and verified that its interior power distribution is indeed predicted by the theory. We also made measurements in two equipment bays of a Boeing 707 aircraft and verified that the power measured at various points in these cavities has the same distribution. The aircraft tests were funded by the NASA Langley Research Center and were conducted in collaboration with the U.S. Naval Surface Warfare Center, Dahlgren.

Introduction

Modern transportation systems and military systems are increasingly dependent on sophisticated electronic controls. At the same time, the potential for electromagnetic (EM) susceptibility of electronic systems is increasing for several reasons: modern integrated circuits with higher densities and speed are often more sensitive to EM transients; modern composite material structures may provide poorer EM shielding; the EM power in the environment is increasing as more users share the airwaves. Because of these factors, the need to ensure the survivability of electronic systems that may be exposed to high-power EM signals has become of great interest. Advisory regulation has recently been drafted for aircraft that would require testing and/or analysis to assure the EM hardness of installed flight critical and flight essential equipment. Similar safety assurance certification procedures may be imposed on other electronically controlled transportation systems of the future. The Department of Defense (DOD) is particularly interested in developing methods for assessing potential EM effects on military systems, since the normal environment these

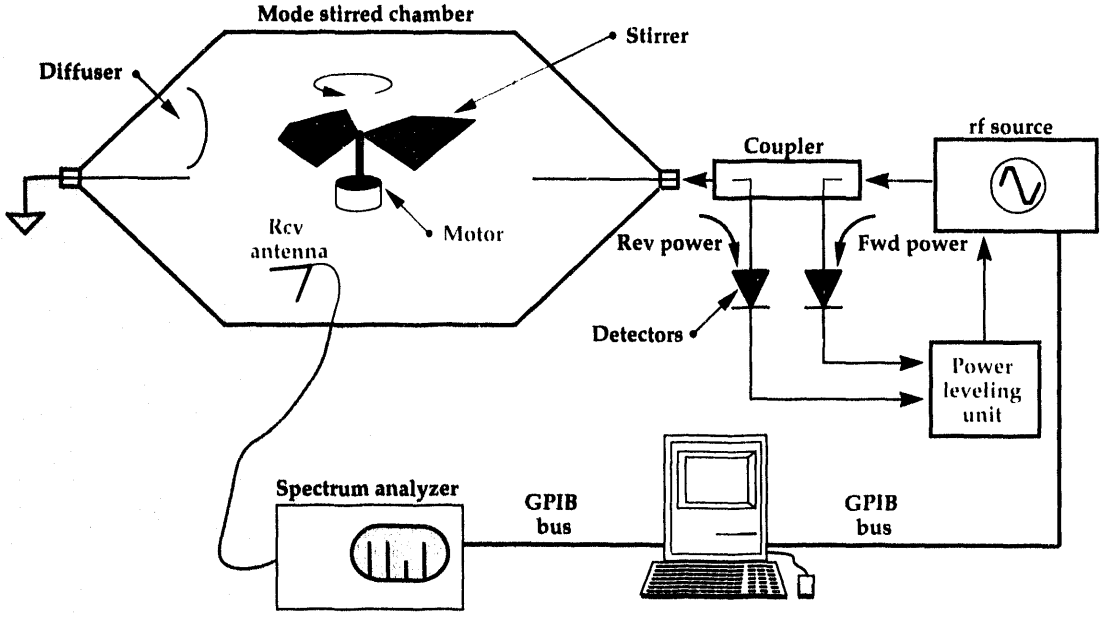


Figure 1. Block diagram of Mode Stirred Chamber instrumentation.

systems operate in is often quite severe (e.g., an aircraft carrier deck), and since these systems may be exposed to high-power signals intentionally transmitted by an adversary. The DOD also needs similar methods to assess the lethality of proposed EM weapons against candidate targets.[1]

While full-system, full-threat testing may be the most thorough manner to determine susceptibility, it is often too expensive to be practical for large systems. Computer models alone have been unsuccessful in accurately predicting EM susceptibility, especially at high frequencies. At Lawrence Livermore National Laboratory (LLNL), we have developed economical assessment techniques based on measuring and comparing EM stress and strength.[2] The power to a device or circuit (stress) is extrapolated from a low-power EM coupling measurement and is compared against the upset threshold (strength) of the device or circuit. A system model is used to relate device or circuit effects to system effects. This technique works well for small systems such as missiles and land mines, but for large systems, the number of measurements and analyses becomes large. In addition, as the system size becomes many wavelengths, the coupling as a function of frequency and angle also becomes extremely complex to the point that deterministic descriptions cannot be made. In this regime, new methods of testing and modeling must be developed. We believe that the mode stirred chamber (MSC) and statistical electromagnetics will provide a new method to make viable stress-vs-strength comparisons for subsystems.[3,4,5]

Progress

Our first major accomplishment was to develop a small, low-power MSC. The chamber was equipped with a power leveling circuit and with automated control and data acquisition instrumentation. Chamber performance was characterized. Our second major accomplishment was to measure and analyze the statistical distribution of microwave power in equipment bays of a full-sized transport aircraft. The measured distribution matched that measured in the MSC and that predicted by theory.

MSC Development

Our MSC was built from an existing transverse electromagnetic (TEM) cell. A TEM cell is designed to operate single mode and produces a well-known field pattern usually used to calibrate sensors. Above its cutoff frequency, the TEM cell becomes increasingly overmoded. To promote more effective overmoding, we removed a section of the septum (center conductor), and added large reflective diffusers, to scatter the energy in random directions within the cell. A motor-spun stirrer paddle, large compared to the wavelength, randomizes the field pattern as a function of time. At frequencies greater than several times the cutoff ($f > 400$ MHz), the mode density is large, and the stirring produces random field patterns.

Figure 1 shows a block diagram of the MSC instrumentation. A low-power microwave source

delivers the EM energy into the chamber. The rf signal can be fed into the chamber through the usual TEM cell input port or by a separate horn antenna placed inside the chamber. A power leveling unit was developed to dynamically adjust the power to compensate for power reflected at the transmit antenna due to changes in the voltage standing wave ratio (VSWR) as the stirrer turns. We used a variety of sensors and antennas, and a spectrum analyzer to measure power at various points in the chamber. Fully automatic control and data acquisition were implemented with a personal computer.

Boeing 707 Tests

Tests were conducted at Davis Monthan Air Force Base in Tucson, Arizona. The experimental setup is shown in **Fig. 2**. The interior of the aircraft equipment bay under study was instrumented with two transmit and two receive antennas. Each pair consisted of a long wire and a horn antenna. The transmit antennas were driven by low-power microwave sources installed in an adjacent instrumentation van, and were oriented to avoid direct illumination of the receiving antennas. Spectrum analyzers were used to measure power at various points within the bays, while a motorized stirrer paddle randomly changed the mode pattern as a function of time. The stirrer motion simulated the random variations in the position of equipment in the bays throughout the fleet of aircraft. Portable computers were used for data acquisition and analysis. Power was measured at discrete frequencies for each of the four possible transmit and receive antenna combinations. The measurements were repeated for several antenna locations within the aircraft bays.

Analysis and Results

The time (or stirrer position) waveforms for both the MSC and aircraft data looked similar in that the received amplitude varied rapidly with excursions of 20 dB or more. The time data was statistically analyzed to generate probability density histograms of the power (in dBm). These are shown in **Fig. 3** (**a** and **b**), for the aircraft and MSC, respectively, along with the theoretical probability density (dashed curves). The predicted density was derived from a two-degree-of-freedom, Chi-squared density using a variable transformation to convert to dB. The aircraft and MSC data compare well to each other and to theory.

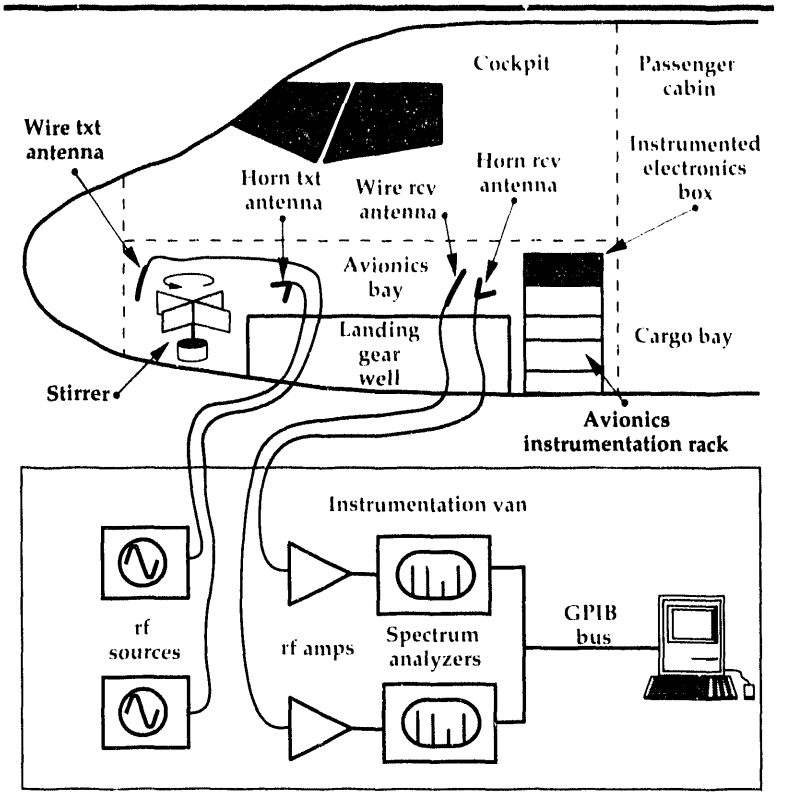


Figure 2. Block diagram of instrumentation for microwave power statistical distribution measurements made at Davis Monthan Airbase on a Boeing 707.

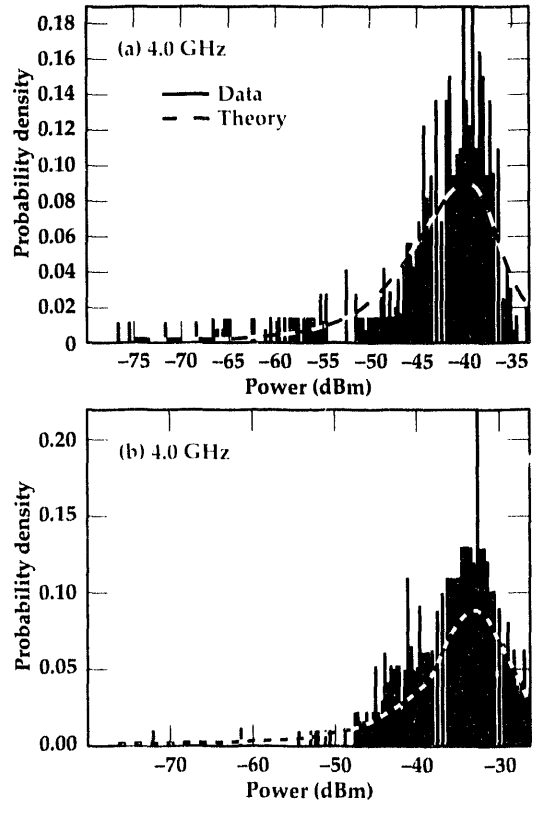


Figure 3. Probability density of measured power in (a) the mode stirred chamber and (b) the equipment bay of a Boeing 707.

Future Work

Four issues must be resolved before certified, quantitative subsystem assessments can be made in MSC's:

1. MSC tests must be shown to be repeatable, predictable, and rigorous. This is necessary for the technique to be accepted by government regulators and industry.
2. The power distribution in the MSC must match that found in the system cavity. This ensures that the subsystem sees the same EM power environment as if it were installed.
3. Coherence length must be understood and the effects of structures near the subsystem taken into account. A structure installed in the cavity will interact with a subsystem if their separation is within a coherence length. Therefore, structures within this length will need to be simulated in the MSC for accurate results.
4. The transfer function describing coupling to devices in the subsystem must be shown to be separable into a product of coupling from outside the system cavity to its interior, and coupling from the cavity interior to the devices within the subsystem. An example of a non-separable case is when the major coupling into the subsystem is through a waveguide that exits the cavity. In this case, the random field environment in the cavity is immaterial, and MSC tests would not provide accurate results.

In FY-92, we developed a small MSC facility and instrumentation. We made measurements in the MSC and in a commercial aircraft that showed that the power distributions matched each other and the theory. In FY-93, we are planning experiments to resolve the remaining issues. We will use statistical EM theory to make predictions of the coupling as a function of frequency, onto simple wire antennas in the MSC, and compare those predictions against measurements. This will help to demonstrate that MSC tests can be rigorous and predictable. Since the prediction will be based on the correlation of the field components picked up by wire segments of the antenna, this will also serve to validate our understanding of coherence length. Separability will be demonstrated experimentally in our anechoic chamber and MSC using simple metal boxes representing an airframe and an avionics box. Once shown for a simple case, the measurements will be repeated with various cable and transmission line connections to the avionics box to establish the cases where separability does and does not hold.

The end product of this study will be a set of theoretically and experimentally validated EM susceptibility assessment tools that would allow accurate EM effects testing of subsystems at high frequency, while avoiding expensive high-power full-system tests.

1. A. Pesta and N. Chesser, *Department of Defense Methodology Guidelines for High Power Microwave (HPM) Susceptibility Assessments*, Office of the Secretary of Defense, HPM Methodology Panel Report, Draft (1989).
2. R. Mensing, R.J. King, and H.S. Cabayan, *A Method for Estimating the Susceptibility of Electronic Systems to HPM*, Lawrence Livermore National Laboratory, Livermore, California, UCID-21430 (1988).
3. M. Crawford and G. Koepke, *Design, Evaluation, and Use of a Reverberation Chamber for Performing Electromagnetic Susceptibility/Vulnerability Measurements*, NBS Technical Note 1092 (1986).
4. R. Price, H. T. Davis, R. H. Bonn, E. P. Wenaas, R. Achenbach, V. Gieri, R. Thomas, J. Alcala, J. Hanson, W. Haynes, C. McCrea, C. Montano, R.Peterson, B. Trautlein, and R. Umber, *Determination of the Statistical Distribution of Electromagnetic Field Amplitudes in Complex Cavities*, Jaycor Report No. 88JAL129 (1988).
5. T. Lehman, *The Statistics of Electromagnetic Fields in Cavities with Complex Shapes*, Phillips Laboratory Interaction Note (1993).

Magnetically Delayed Low-Pressure Gas Discharge Switching

Stephen E. Sampayan, Hugh C. Kirbie, and Anthony N. Payne
Laser Engineering Division
Electronics Engineering

Eugene Lauer and Donald Prosnitz
Advanced Applications Program
Laser Programs

We have investigated the properties of a magnetically delayed, low-pressure gas discharge switch. We performed measurements of the closure and recovery properties of the switch; performed quantitative erosion measurements; and observed the onset of x ray production in order to compare switch properties with and without delay. Further, we performed qualitative optical measurements of transition line spectra to correlate our electrical recovery measurements with plasma deionization.

Introduction

Fast-closure-rate, high-voltage (> 100 kV), high-current (> 10 kA), high-repetition-rate (> 1 kHz) switching has remained a major area of research in the pulsed power field.[1,2,3] Solid-state switching has generally been limited to several tens of kilovolts; high-pressure gas discharge switching is limited to repetition rates below 1 kHz; vacuum switching is generally a slow closure process; and magnetic switching requires extremely precise voltage and reset state control to minimize jitter.

Low-pressure gas discharge switches have shown promise as a fast-closing, high-repetition-rate device such that if sufficiently fast closure times can be achieved, single-stage power conditioning chains would become feasible.[4,5]

The primary difficulty with this switch, however, is anode electrode damage during closure initiation, resulting in short lifetimes. Once triggered, electrons emitted from the cathode plasma can form a pinched beam and deposit significant enough localized energy to vaporize anode material. Inserting a series delay element, which inhibits the application of full voltage and current until such time that the discharge plasma has filled the gap, minimizes this effect. It is this version of the low-pressure switch that we are presently studying.

Progress

Our magnetically delayed low-pressure switch (MDLPS) test-stand was built primarily to support the long-pulse, relativistic klystron (RK) and free electron laser (FEL) work at Lawrence Livermore National Laboratory (LLNL).[6] In this application, a closing switch initiates a pulse, which is delivered to an induction accelerator cell.[7] The induction cell accelerates an injected electron beam to a sufficient energy suitable for the RK or FEL.

Apparatus

The MDLPS test-stand (**Fig. 1**) consists of a single water-filled, 12-Ω, 70-ns Blumlein from the advanced test accelerator at LLNL. The Blumlein is attached to a liquid load and charged from a single dual-resonant transformer. The transformer is powered by two charged capacitor banks discharged through separate thyratrons, diode isolated and fired sequentially to produce two charging pulses. A trigger pulser initiates a single closure event at the peak of the first charging pulse; the second, variable timing, charging pulse is allowed to ring to zero and is used as a test pulse to verify gap recovery.

The low-pressure gas gap consists of an anode-cathode electrode pair separated a sufficient dis-

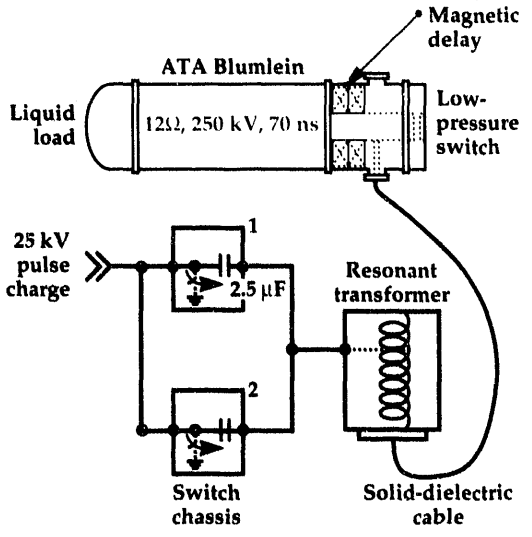


Figure 1. Switch test-stand used to study magnetically delayed switch properties.

tance to prevent self breakdown (at approximately 100 to 150 kV/cm). A surface flashover triggering device, embedded within the cathode, initiates the ionization processes that render the gas highly conductive. A saturable inductor placed in series with the switch, delays the onset of full current, allowing the ionization processes to spread throughout the gap volume prior to full closure. The saturable inductor is designed to limit current flow below the threshold for constricted discharges, and hold off the full anode-cathode voltage until the discharge has filled the gap volume.

Diagnostics for the test-stand consisted of current and voltage sensors for the switch and Blumlein. Other diagnostics consisted of a fast x ray detector, a fast gated camera, and a 0.25-m monochromator. At preset, we used an image-intensified gated camera to observe the monochromator output. We are installing a gated photomultiplier system for future work.

Experiments

We performed standard parametric studies of the magnetically delayed low-pressure switch and compared the performance with and without the saturable inductor. A comparison of typical closure properties is shown in **Fig. 2**, and of typical recovery properties in **Fig. 3**. For these data, the gap spacing was 1 cm, the gas was nitrogen, and the anode-cathode voltage was 90 kV.

Closure time, defined as the 10 to 90% transition time of the voltage across the low-pressure gas gap, showed a factor-of-two improvement at lower pressures with the magnetic delay, i.e., with the saturable inductor. At higher pressures, above approximately 7 mT, the saturable inductor had little effect.

Recovery time with the series saturable inductor improved significantly and was extremely reliable: 99% recovery probability was estimated. At lower pressures, extremely good recovery times (approximately 50 μs or 20 kHz-equivalent pulse repetition frequency) were observed, while at higher pressures, recovery time was observed to be 5 μs. By contrast, recovery probability without the series saturable inductor was not reliable and was measured to be between 80 and 90%. Although at lower pressures and this recovery probability, faster recovery times were observed, recovery did not occur above 7 mT.

We made qualitative spectroscopic measurements of late-time line emission from the gap in order to verify our recovery measurements performed electrically. Spectroscopic observations of the discharge showed that line radiation from the nitrogen decayed within 10 μs after current cessation. Line radiation characteristic of the anode material, however, required greater than 50 μs to decay. This result was consistent with our electrical measurements.

Erosion rates with the series saturable inductor were a factor of 60 less than those of a similar lifetime test without the series saturable inductor. Photographic comparisons are shown in **Fig. 4**. These tests were conducted at 90 kV, with nitrogen at 8 mT pressure and using aluminum anodes. In the first test (**Fig. 4a**), severe anode damage was observed over the entire surface of the electrode and particularly across from the trigger electrode. The total number of shots was

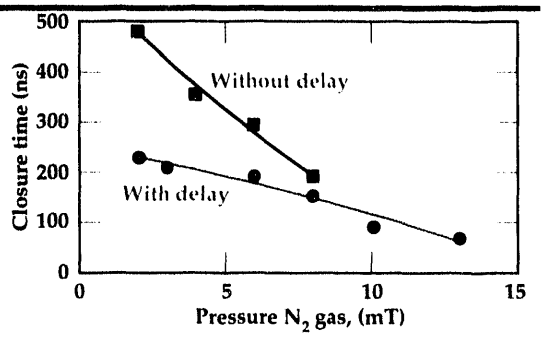


Figure 2. Measured closure results with and without delay.

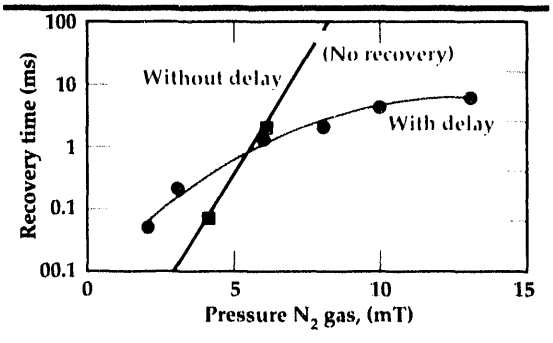


Figure 3. Electrically measured recovery results.

approximately 16,000. In the second test (**Fig. 4b**), a less pronounced indentation resulted from the test, with minimal damage having occurred throughout the anode surface. The total number of shots during this latter test was approximately 400,000.

We performed x ray measurements (**Fig. 5**) to understand the time evolution of electron emission from the cathode of the low-pressure switch. Our relative measurements of the integrated x ray output during switch closure showed an order-of-magnitude decrease with the series saturable inductor. Further, we observed the most intense x ray output from the low-pressure switch during the initiation or 'trigger delay' period without the series saturable inductor, and after the closure process had begun with the series saturable inductor.

We also measured the variation of the x ray output from the gap at various gas pressures. From this measurement, we observed the x ray output decrease by about 30% when the gas pressure was increased from 1 to 9 mT.

Modeling

We developed a one-dimensional model for the closure regime of the low-pressure switch. In this model, the motions of ions and electrons are modeled by cold fluid equations that include collision ionization and space charge effects. The model equations are parameterized in terms of gas type (ionization coefficient) and pressure; switch gap length and cross-sectional area; and the initial electron density produced by the trigger pulse. This model permits us to follow the space–time evolution of the electric field and the ion and electron current densities in the gap, as well as the total switch current and terminal voltage during switch closure.

We have implemented the model in a general-purpose network and system simulation code. This implementation permits the construction of a detailed system simulation model of the test-stand that includes the charging circuit, Blumlein, magnetic switch, and load. We use a magnetic switch model that includes rate-dependent loop-widening of the hysteresis loop, hysteretic losses, minor loops, and hysteresis effects.[8] Presently, we are validating the low-pressure switch and magnetic switch models against experimental data. Once validated, the complete system model should permit us to study the sensitivity of switch closure performance to magnetic core parameters and to low-pressure switch parameters such as electrode spacing and initial electron density, and thereby provide us with a tool for making switch design tradeoffs.

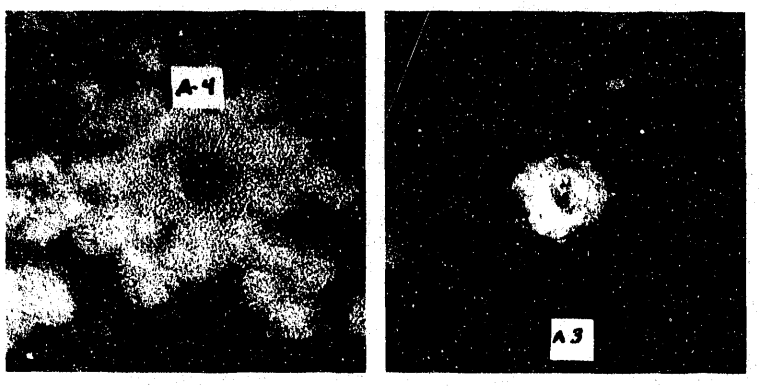


Figure 4. Comparison of electrode erosion (a) without and (b) with magnetic delay.

Future Work

We have demonstrated that the use of a series saturable inductor placed in series with a low-pressure gas spark gap greatly enhances performance. From our measurements, we understand this improved performance to be primarily due to minimizing anode material vaporization during the initial closure of the gap. Without the series saturable inductor, x ray emission occurs

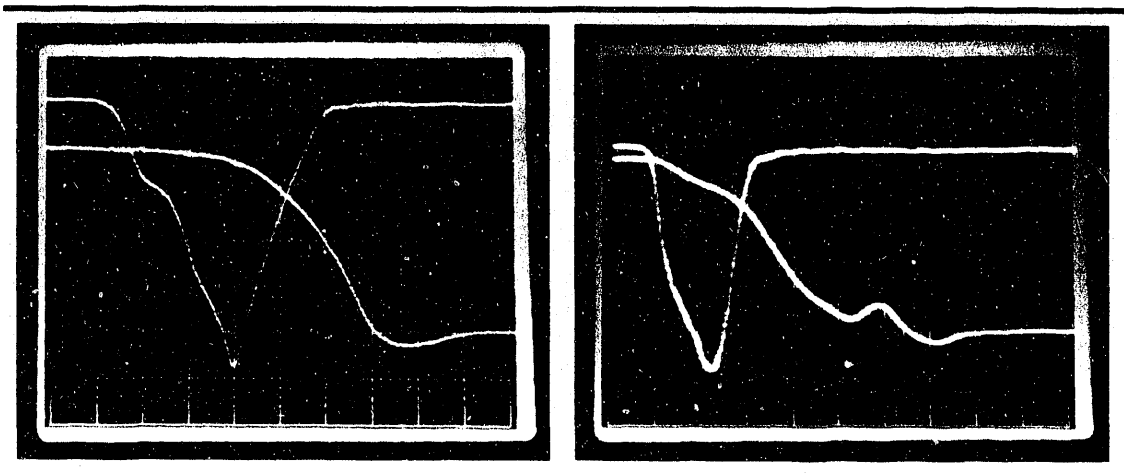


Figure 5. Measured x ray pulse (top curve) and closure (bottom curve) of the low-pressure switch (a) without magnetic delay (0.5 V/div.) and (b) with magnetic delay (0.1 V/div.). Horizontal scale is 100 ns/div.

from the point of triggering until the initial collapse of the gap impedance. The energy deposition into the anode is large as determined by the integrated x ray intensity. With the series saturable inductor, energy deposition into the anode is initiated at the instant the collapse of the gap impedance occurs. The net effect is lower energy deposition into the anode.

From our data, we conclude that with our present triggering method, this switch is capable of operating as either a low-repetition-rate final output switch or, because of the slower closure times at low pressure, as a high-repetition-rate initial commutation switch, i.e., in the initial stages of the power conditioning chain. Although the present triggering device appears adequate, it is difficult to couple a significant portion of the trigger electrical energy into the low-pressure gas. In future work, we will install newly developed, simple, ferroelectric electron emitters as a triggering device.[9] Current densities from 0.1 to 1 kA/cm^2 have been extracted from such an emitter for several hundred nanoseconds. Such a device should allow better coupling of the trigger electrical energy to the low-pressure gas. We would therefore expect much faster closure times even at lower pressures.

Our spectral observations indicate that recovery is primarily inhibited by anode vapor remaining ionized in the gap. It is well established that recombination times for metal vapor exceed those of gasses by at least an order of magnitude. Thus, to enhance recovery, we will investigate the use of anode materials with low heat of vaporization, in order to minimize the accumulation of anode material vapor in the gap.

1. G. Schaefer, M. Kristiansen, and A. Guenther, *Gas Discharge Closing Switches*, Plenum Press (New York, New York), 1990.
2. H.C. Kirbie, G.J. Caporaso, M.A. Newton, and S. Yu, "Evolution of High-Repetition-Rate Induction Accelerators Through Advancements in Switching," *1992 Linear Accelerator Conf. Proc.*, 595 (1992).
3. R.A. Dougal, G.D. Volakakis, and M.D. Abdalla, "Magnetically Delayed Vacuum Switching," *Proc. 6th IEEE Pulsed Power Conf.*, 21 (1987).
4. E.J. Lauer and D.L. Birx, "Low Pressure Spark Gap," *Proc. 3rd Int. Pulsed Power Conf.*, 380 (1981).
5. E.J. Lauer and D.L. Birx, "Tests of a Low Pressure Switch Protected by a Saturable Inductor," *IEEE Conf. Record 1982 15th Power Modulator Symposium*, 47 (1982).
6. T.L. Houck and G.A. Westenskow, "Status of the Choppertron Experiments," *1992 Linear Accelerator Conf. Proc.*, 498 (1992).
7. S. Humphries, *Principles of Charged Particle Acceleration*, John Wiley and Sons, Inc. (New York, New York), 283ff (1986).
8. A.N. Payne, "Modeling Magnetic Pulse Compressors," *Conf. Record 1991 Particle Accelerator Conf.*, 3091 (1991).
9. H. Riege, *New Ways of Electron Emission for Power Switching and Electron Beam Generation*, European Organization for Nuclear Research, Report CERN-PS 89/42(AR) (1989).

Nondestructive Evaluation

The Nondestructive Evaluation (NDE) thrust area supports initiatives that advance inspection science and technology. The goal of the NDE thrust area is to provide cutting-edge technologies that have promise of inspection tools three to five years in the future. In selecting projects, the thrust area anticipates the needs of existing and future Lawrence Livermore National Laboratory (LLNL) programs.

NDE provides materials characterization inspections, finished parts, and complex objects to find flaws and fabrication defects and to determine their physical and chemical characteristics. NDE also encompasses process monitoring and control sensors and the monitoring of in-service damage. For concurrent engineering, NDE becomes a frontline technology and strongly impacts issues of certification and of life prediction and extension.

In FY-92, in addition to supporting LLNL programs and the activities of nuclear weapons contractors, NDE has initiated several projects with government agencies and private industries to study aging infrastructures and to advance manufacturing processes. Examples of these projects are (1) the Aging Airplanes Inspection Program for the Federal Aviation Administration; (2) Signal Processing of Acoustic Signatures of Heart Valves for Shiley, Inc.; and (3) Turbine Blade Inspection for the Air Force, jointly with Southwest Research Institute and Garrett.

In FY-92, the primary contributions of the NDE thrust area, described in the reports that follow, were in fieldable chemical sensor systems, computed tomography, and laser generation and detection of ultrasonic energy.

Fieldable Chemical Sensor Systems

Our objective for this project is to develop diagnostic instruments for quantitative measurements of the levels of chemical contaminants and concentrations in the field or on-line in operating processes. We believe that the integration of Raman spectroscopy and fiber-optic sensors will allow a revolution in chemical analysis by developing the capability for field analysis rather than the more conventional methods requiring extraction of samples for later evaluation in a laboratory environment. We are in the second phase of a two-phase project. In the initial phase, we determined the widespread needs for chemical sensors to measure contaminant levels in liquids and gases and on solid surfaces. We selected Raman spectroscopy as the first system to develop because of its wide applicability as a chemical monitoring technique. During the second phase, we have developed a state-of-the-art micro-Raman spectrometer, designed two unique fiber-optic-based sensors for remote coupling of the spectrometer to either liquid/gas phase samples or solid surfaces, and purchased a new imaging spectrometer and low-light-level detector.

Computed Tomography

We continue to develop computed tomography (CT) scanners covering a broad range of object sizes. An integral part of this work is to derive the

reconstruction and display algorithms. The overall goal of this work is to improve the three performance parameters (spatial and contrast resolutions and system speed) that characterize CT imaging systems. In addition, we have addressed related topics such as elemental or effective-Z imaging, model-based imaging using *a priori* information, parallel processor architectures for image reconstruction, and scientific visualization of reconstructed data.

In FY-92, we completed the California Competitive Technology Cone-Beam CT Project with Advanced Research and Analysis Corporation as our industrial partner. We continued to work on two other projects: the active/passive CT of radioactive drums and characterization of high explosives for the Pantex plant; and high-purity single-crystal germanium detectors in collaboration with the University of California, San Francisco.

Laser Generation and Detection of Ultrasonic Energy

Finally, we have developed a facility to generate and detect ultrasonic energy with lasers. Laser-generated ultrasonics is an attractive alternative to traditional ultrasonic NDE, because it allows remote, noncontacting, ultrasonic NDE. We are developing NDE applications for use on contamination-sensitive components and in hostile environments. Laser ultrasonics has several other advantages, such as broadband excitation, multimode acoustic energy generation, and adaptability to scanning complex shapes.

Satish V. Kulkarni
Thrust Area Leader

Section 8

8. Nondestructive Evaluation

Overview

Satish V. Kulkarni, Thrust Area Leader

Fieldable Chemical Sensor Systems

Billy J. McKinley and Fred P. Milanovich **8-1**

Computed Tomography

Harry E. Martz, Stephen G. Azevedo, Daniel J. Schneberk, and George P. Roberson **8-5**

Laser Generation and Detection of Ultrasonic Energy

Graham H. Thomas **8-23**

Fieldable Chemical Sensor Systems

Billy J. McKinley
Engineering Sciences
Mechanical Engineering

Fred P. Milanovich
Environmental Sciences Division
Biomedical and Environmental
Research Program

In the initial phase of this project, we determined the widespread needs for chemical sensors to measure contaminant levels in liquids and gases and on solid surfaces. We selected Raman spectroscopy as the first system to develop because of its wide applicability as a chemical measuring technique.

During FY-91, we developed a state-of-the-art micro-Raman spectrometer capability, designed two unique fiber-optic-based sensors for remote coupling of the spectrometer to either liquid/gas phase samples or solid surfaces, and purchased a new imaging spectrometer and low-light-level detector. During FY-92, we developed two complete systems around these two new sensors and demonstrated the application of the solid surface sensor in the analysis of diamond coatings.

Introduction

Our objective is to develop diagnostic instruments for quantitative measurements of the levels of chemical contaminants and concentrations in the field or on-line. We believe fiber-optic coupled Raman spectroscopy will make a significant impact in chemical analysis by developing the capability for field analysis, as opposed to the more conventional methods requiring extraction of samples for later evaluation in a laboratory environment. We are in the second phase of a two-phase project.

In the initial phase of this work, we surveyed the needs for sensors, particularly with respect to environmental restoration and waste management concerns.[1] The most obvious needs were for chemical sensors that can be used in the field, thus eliminating the costly, time-consuming, and often impossible process of bringing samples to the laboratory for analysis. From our involvement with Nuclear Weapons Complex reconfiguration planning, we also observed a need for on-line or near-line chemical analysis in chemical processing. The combined set of needs ranges from *in-situ* analysis of contaminants in groundwater to on-line monitoring and feedback control at multiple locations along the process line for chemical-processing operations. We are addressing these needs by developing diagnostic instruments and sensor systems that are simple, robust, portable, and sensitive enough for field operation and decision making.

We have chosen Raman spectroscopy for development for a number of reasons,[1] primarily because of its wide applicability as a chemical sensor. The major problem in the application of Raman spectroscopy is in the signal-to-noise ratio, which is related to the extremely low scattering cross section (10^{-29} cm^2/mol-sr). In a typical measurement, coherent scatter from the laser is 10^8 times greater than the Raman shifted incoherent scatter. The major thrust of our project, therefore, is to create sensors that maximize the Raman scatter and the acceptance angle of the optical system, which collects the scattered light for the spectrometer.

Progress

Micro-Raman Spectroscopy Capability

During FY-91, we established a micro-Raman spectroscopy capability.[2] Although the micro-Raman spectrometer developed represents major progress in our facilities and capability, it is limited in various ways: (1) it can only accept 0.5-cm-dia samples; (2) it requires considerable alignment of the optical components on a regular basis; and (3) the old spectrometer has considerable scattering noise, and the detector system's sensitivity and noise figure are not as good as that of more modern detectors.

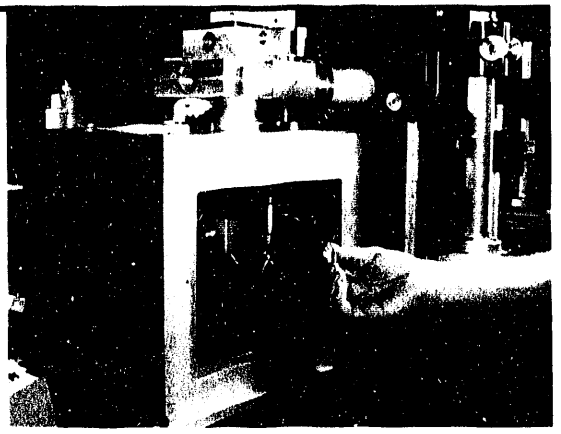


Figure 1. Miniaturized fiber-optic coupled sensor. The sensor replaces all the other optical apparatus shown in the photograph.

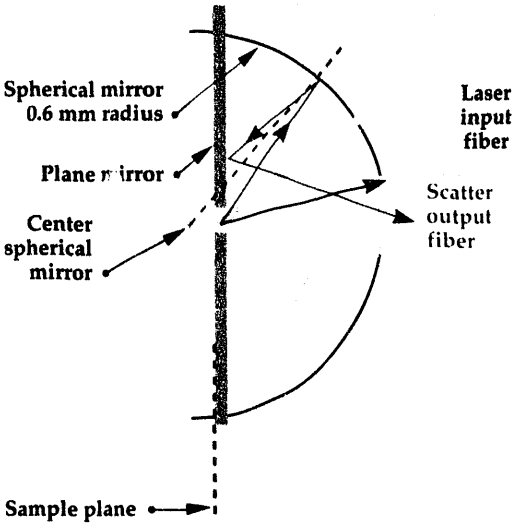


Figure 2. Illustration of the principle of operation of the one-sided Raman sensor for solid surfaces.

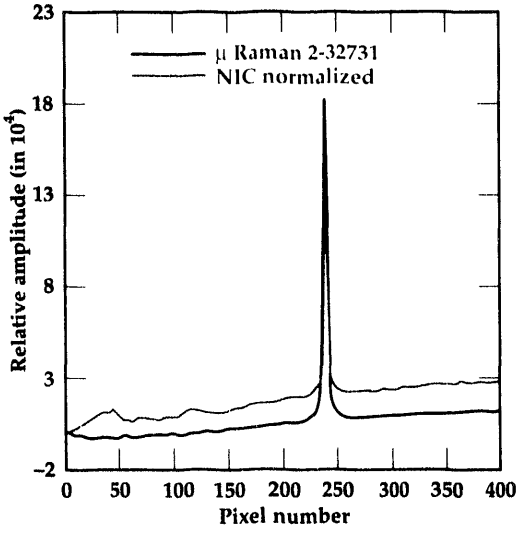


Figure 3. Comparison of data for one-sided (NIG) sensor and micro-Raman system.

To correct the deficiencies of this system and to modify it to serve as the testbed for more advanced systems, we purchased an imaging spectrometer and a low-light-level, liquid-nitrogen-cooled, two-dimensional charge-couple-device detector. The combination of these new instruments is a significant advancement in our capability to address multipoint, high-spectral-resolution, broad-bandwidth spectral analysis and the very demanding low-light-level recording issues associated with Raman spectroscopy.

Fiber-Optic Remote Coupling Devices

As previously stated, the core objective of this research is to develop the instrumentation necessary to perform field analysis. This includes two fiber-optic-based devices for remote coupling of the spectrometer to the sample. The first, to be used in analysis of solid surfaces, is referred to as the one-sided sensor; the second, for analysis of trace concentrations in solutions, is a liquid-core optical waveguide. The liquid core serves as the sample cell and as a low-loss light transmission line. The waveguide not only retains the laser excitation light, but is an efficient concentrator of the Raman scatter. Both of these designs are discussed further below.

One-Sided Sensor. Our goals with the one-sided sensor are to achieve a major advance in miniaturization of instrumentation; to eliminate numerous interactive, optical-alignment adjustments; to accommodate large solid-surface evaluation; and to allow more practical, robust, quantitative field applications of spectroscopy. The fiber-optic-coupled sensor shown in **Fig. 1** replaces all the other optical apparatus shown in the photograph, reducing the system size by at least a factor of 100. The fiber-optic sensor needs no adjustments; only its photons move.

Figure 2 shows the principles of the sensor operation: light from the laser is focused onto the surface of the object of interest through an aperture (20-μm-dia) in a plane mirror. The mirror is a thin (2-μm) aluminum-coated polymer membrane. The plane mirror is off-set from the center of the spherical mirror at the appropriate distance for maximum light-gathering efficiency. Light scattered back through the aperture is collected directly by the fiber up to an incident angle roughly equal to the numerical aperture (N.A.) of the fiber (the acceptance angle of the fiber is proportional to the N.A.). Scattered light incident at greater angles is reflected by the two mirrors back to the fiber at a 'shallower' angle within the N.A. of the fiber.

The prototype for this sensor has been evaluated, and a comparison is made with the micro-Raman system (**Fig. 3**). Although preliminary results are very good, further improvements are possible by appropriate optical filtering to elimi-

nate the Raman scatter being produced in the coupling fibers.

Liquid-Core Optical Waveguide Design. Our goal with the liquid-core optical waveguide is to make a major breakthrough in lowering the threshold achievable in on-line analysis for trace concentrations in aqueous solutions.

As mentioned earlier, the major limitation of using Raman spectroscopy for chemical analysis is the very low molecular Raman scattering cross section. For analysis of low (ppm) concentrations in solutions, the problem is orders of magnitude more severe than for concentrated solutions.

To overcome this limitation, we have designed the optical system to maximize the sampling interaction path length in the solution and the acceptance angle of the scattered light returned to the detector. The test bed for evaluating these new waveguides is shown in **Fig. 4**.

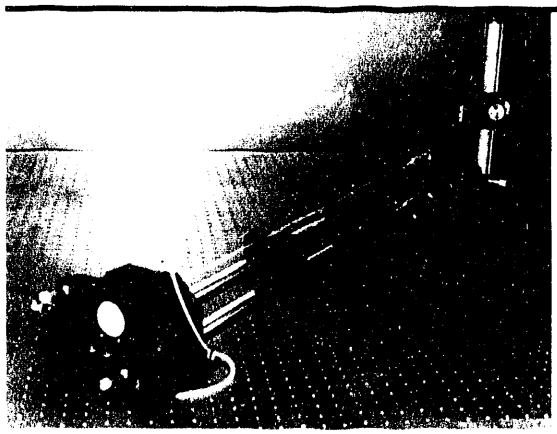


Figure 4. Testbed for evaluating new waveguides.

Future Work

Our plans for future work include three areas:

(1) Optimization of optical system throughput. Greater efficiency can be achieved by optimizing the transfer of light from the sensors to the spectrometer. The spectrometer has a numerical aperture of .24 and a minimum slit width of 10 μm, which set the boundary conditions for the entire optical system. We have begun the optimization design for the one-sided sensor. Many of the limitations and the optimization methods apply directly to the liquid-core waveguide as well.

(2) Testing the liquid-core optical waveguide (cell). We intend to test the system on two categories of problems: (1) analysis of trace contaminant concentrations in groundwater, and (2) chemical process monitoring and feedback control. In the latter case, we will choose a chemical from the uranium-processing line identified in the Nuclear Weapons Complex Reconfiguration Study, as the model for the analytical system in the design of the cell. We will also develop the spectral analysis algorithms that will be necessary in a feedback control loop.

(3) Spectroscopy. To complete this phase of the project, creating a capability for Raman analysis, we must develop more understanding and have more experience in spectroscopic analysis. The first steps have been taken in the purchase of new spectral-analysis software and a Raman spectra-comparative database. We are working with specific examples, such as the diamond coating evaluation.

Acknowledgements

The authors would like to acknowledge John Lutz and Sang Sheem for their contributions to this project.

1. B.J. McKinley, F.P. Milanovich, M.S. Angel, and H.K. McCue, "Fieldable Sensor Systems for Environmental Contaminants," *Engineering Research and Development*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-90, 7-9 (1991).

2. B.J. McKinley and F.P. Milanovich, "Fieldable Chemical Sensor Systems," *Engineering Research, Development, and Technology*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-91, 7-1 (1992).

Computed Tomography

Harry E. Martz
Engineering Sciences
Mechanical Engineering

Daniel J. Schneberk
Applications Systems Division
Computation Directorate

Stephen G. Azevedo
Engineering Research Division
Electronics Engineering

George P. Roberson
Defense Sciences Engineering Division
Electronics Engineering

We are developing several data-acquisition scanners for computed tomography (CT), along with associated computational techniques for image reconstruction, analysis, and display. This report describes recent progress in active and passive CT, cone-beam CT, high-energy CT, and specialized applications research. We have sought to advance the state of the art in CT technology, while at the same time actively supporting programs at Lawrence Livermore National Laboratory and new business initiatives. Our goal is to provide reliable and efficient nondestructive evaluation techniques for use in probing the internal structure of fabricated objects and materials associated with a broad spectrum of applications.

Introduction

Nondestructive evaluation (NDE) is being used in an ever broadening array of industrial and military applications. One area in recent years where growth is evident is computed tomography (CT). First used in the 1970's as a medical diagnostic tool, CT was adapted to industrial and other nonmedical purposes in the mid-1980's. Standard radiographic techniques, such as single projection radiography, hide crucial information: the overlapping of features obscures parts of these features, and the depth of the features is unknown. CT was developed to retrieve three-dimensional (3-D) information.

For CT, several radiographic images of the object are acquired at different angles, and the intensity information collected by one or many detectors is processed in a computer. The final 3-D image, generated by mathematically combining these images, gives the exact locations and dimensions of internal features within the object, as well as external details. Over the past six years, we have worked on research and development (R&D) of many CT topics, concentrating on three main areas: (1) scanners, (2) software tools, and (3) applications.[1-9] Two years ago, we began R&D on a combined active and passive computed tomography (A&PCT) system.[1,2]

In this report, we describe our major progress in A&PCT, cone-beam CT and high-energy CT. We also present advances in the application of these and other capabilities for both Lawrence Livermore National Laboratory (LLNL) programs and business. Lastly, we outline our future plans.

Progress

A&PCT Research

Characterization of mixed (radioactive and hazardous) wastes requires that the identity and strengths of intrinsic radioactive sources be determined accurately. In collaboration with LLNL's Nuclear Chemistry Division, we have developed a three-phased plan to address the nondestructive assay (NDA) of 208-L (i.e., 55-gallon) drums. These phases are (1) experimental A&PCT research and development, (2) simulated A&PCT research and development, and (3) determination of minimum-detectable limits vs waste-matrix attenuation. We report here on the experimental and simulated

A&PCT efforts. The determination of minimum-detectable limits vs waste-matrix attenuation effort was funded by the Office of Safeguards and Security and is described elsewhere.[10]

A&PCT Scanner. Experimental data were acquired on a small-scale canister containing mock wastes and two passive sources, 95-μCi ^{133}Ba and 74-μCi ^{228}Th, using a medium-energy CT scanner (MECAT) built at LLNL.[11,12] These data were used to investigate (1) ACT, to obtain images that represent cross-sectional attenuation maps of a waste-canister's contents; (2) PCT, to locate and determine the identity of any radioisotopic sources present; and (3) ACT data, to correct the PCT data so that accurate source strengths can be determined.

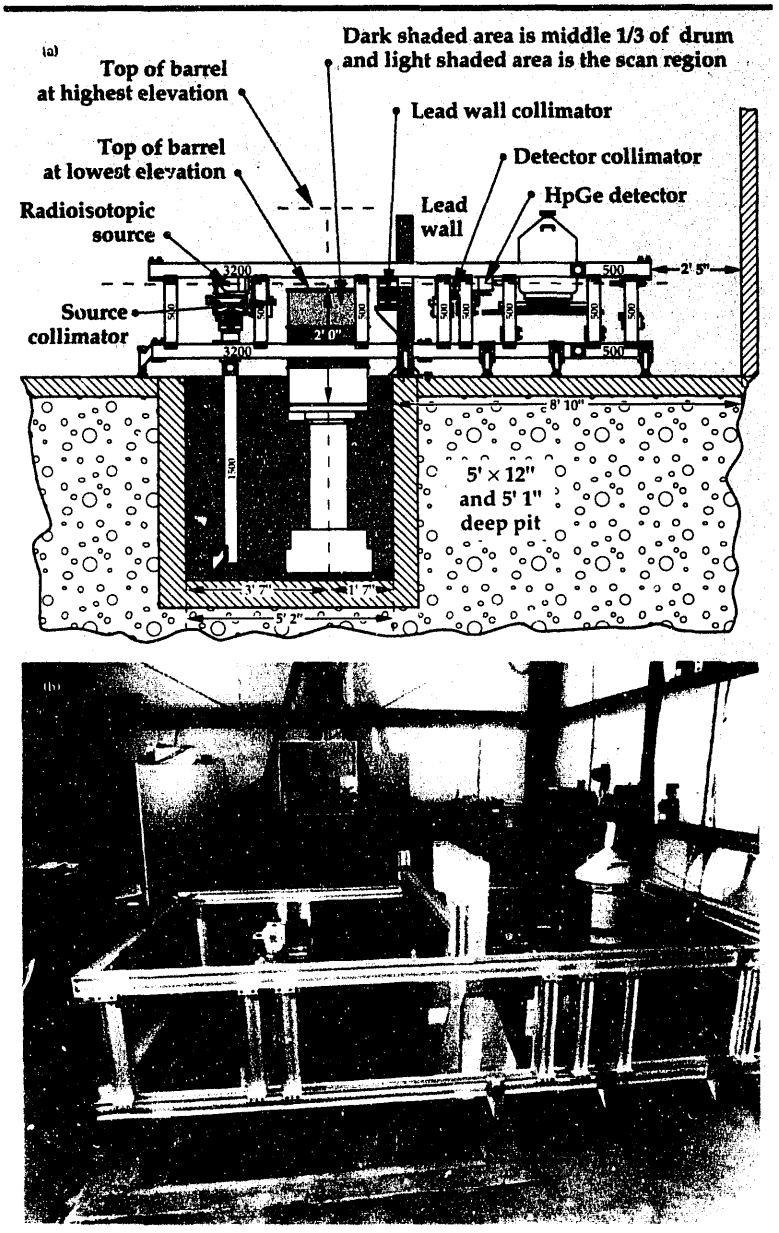


Figure 1. (a) ACT and PCT prototype scanner design: (b) scanner photo, showing construction to date.

Our experimental results reveal that ACT scans properly map the canister's attenuating matrix and, when coupled with PCT scans, yield quantitative source strengths.[13] Preliminary results suggest that heavy-metal content, which is larger than the volume-element size imaged, may be identified. These encouraging results have led us to design and construct a full-scale, 208-L prototype A&PCT drum scanner.

The full-scale, 208-L prototype A&PCT drum scanner uses a single, high-purity germanium (HPGe) detector of the type used in nuclear spectroscopy measurements. This scanner's construction is scheduled for completion early in 1993. The scanner design and progress to date are shown in **Fig. 1**. This scanner will be used to better explore and understand the relationships among the four most important CT performance or resolution parameters—spatial, contrast, energy, and temporal (or speed)—from the point of view of assaying nuclear waste drums for radioactive content. The definition of contrast resolution differs for the A&PCT measurements. In the former, it is a measure of attenuation differences that can be observed; in the latter, it is a measure of radioactive strength differences. Speed includes data-acquisition and analysis time.

Limits to improving the PCT activity results include geometrical uncertainties caused by the collimator's angular cone of acceptance, photon scattering, lack of sufficient counts, the randomness inherent in photon counting, poor detection efficiency, the energy resolution required, system noise, data-acquisition, and times required for data analysis. Quantitative assays using PCT are further complicated by the need for attenuation corrections, which are obtained from the ACT data. Unfortunately, these data are limited by many of the same performance parameters.

Reconstruction Technologies. We have developed A&PCT image-reconstruction and simulation algorithms to better characterize mixed-waste drums, in collaboration with Laboratoire d'Electronique de Technologie et d'Instrumentation (LETI) in Grenoble, France, and the University of California at San Francisco (UCSF). The A&PCT image reconstruction and analysis process consists of mapping the activity of intrinsic radioactive sources, using PCT data, and correcting this

data for attenuation, by using an attenuation matrix obtained from an ACT scan. Simulated data are necessary to better understand A&PCT reconstruction algorithms and measure their performance, and to better interpret experimental data. The simulation program is based on a forward-projection algorithm[14] that discretely computes the projections; i.e., integrated counts per unit time per unit volume, of an emitting object attenuated by a user-specified matrix. We use two algebraic, iterative, A&PCT reconstruction codes: a weighted-least-squares, steepest-descent (WLS-SD) algorithm and a maximum-likelihood expectation-maximization (MLEM) algorithm.[14]

We studied three simulated phantoms: (1) a large homogeneous source included in a large homogeneous attenuator; (2) a mock-waste drum involving small sources; and (3) a spatial-resolution phantom.[15] All three examples involve attenuation but are not strong enough to cause missing data, and the last example includes noise. The noise is generated to match experimental uncertainties. Results showed that both reconstruction algorithms recover the activity values to within experimental uncertainty (counting statistics). The WLS-SD algorithm produces more spreading of the activity over multiple pixels, but also performs slightly better than MLEM (i.e., it gives more accurate activity values) in the case of noise. We are working on a number of improvements in these algorithms; e.g., incorporation of collimator geometry, addition of the effects of very strong attenuation, and optimization of the code for speed and activity accuracy. We are also investigating other algorithms.

A&PCT Applications. In addition to continuing experiments on the small-scale canister of mock waste, we studied the attenuation of both passive sources by a uniformly attenuating Cu cylinder. A representative ACT image of the Cu cylinder with both passive sources is shown in **Fig. 2**. Note that the locations of the cylinder and both passive sources are easily visible. It is also interesting to point out the 30% difference in the Cu cylinder wall attenuation value. We found that this difference may be due to a wall thickness variation of ~ 1 mm from one side of the cylinder to the other, and porosity. A portion of the resultant change in the wall attenuation value can be attributed to partial volume effects (due to the crude 3-mm spatial sampling) instead of a noticeable change in wall thickness.

For the attenuated PCT scan, a combination of three factors yielded noisy net passive projection data: (1) 3-x-3-mm aperture, (2) the short data-acquisition time (150 s), and (3) the attenuation of the passive sources by the copper cylinder (~ 10.1 cm o.d. and ~ 8.9 cm i.d.). The passive sources' energy peaks were within the Compton and background distributions. It is important to point out that the method of extracting the net-passive projection data (gross counts minus spectral background) from the gross projection data is not adequate for low-count-rate PCT data, and resulted in meaningless net-peak intensities for the passive projection data sets. **Figure 3** shows a representative comparison between the gross and net projection data for the ^{228}Th source at 238/240 keV, and for the ^{133}Ba source at 384 keV. Neither the ^{228}Th nor the ^{133}Ba net projection data reveal any

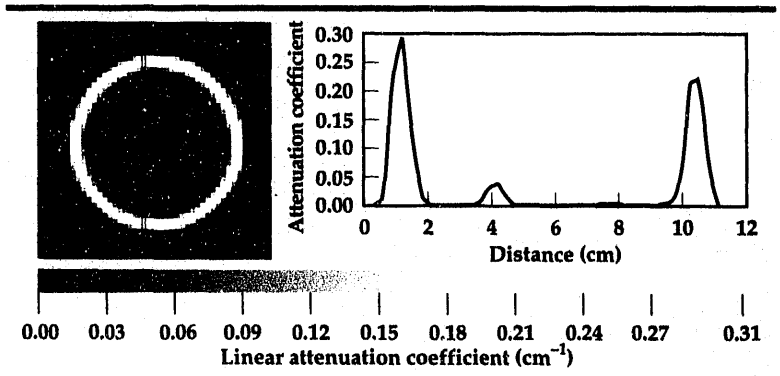


Figure 2. Representative ACT image of a Cu cylinder with passive sources. A 1-D profile of this data is on the right.

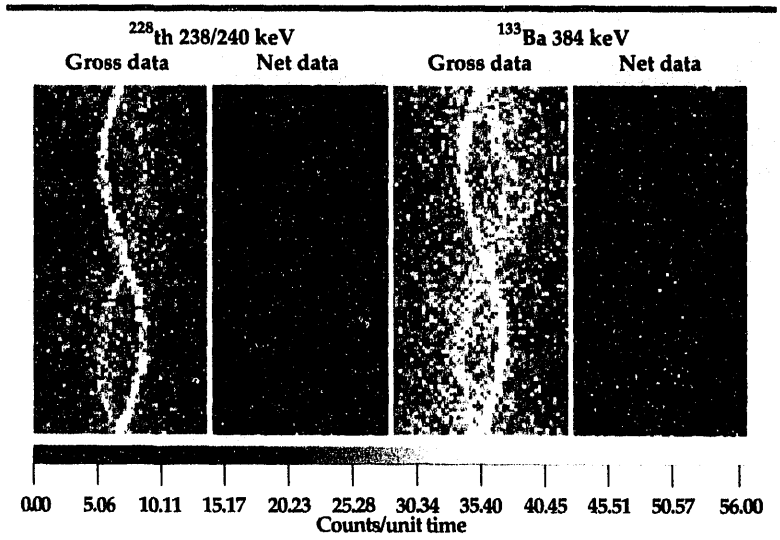


Figure 3. Representative gross and net passive sinogram data at 238/240 keV for the ^{228}Th source, and at 384 keV for the ^{133}Ba passive source.

internal source distribution nor did the final attenuation corrected PCT images (see **Fig. 4**). The resultant, noisy, passive sinogram data are due to a low count rate.

Since the net projection data did not result in passive source identification or localization of activity, we analyzed the gross projection data. As expected, these data are distorted. For example, the gross ^{133}Ba projection A&PCT WLS-SD-reconstructed PCT-image data (**Fig. 4**) appear to have three internal source distributions: (1) a ^{133}Ba passive source at the location, as expected from the ACT image in **Fig. 2**; (2) an apparent ^{133}Ba source at the location of the ^{228}Th source, and (3) an apparent ^{133}Ba ring source. Only the first distribution is real; the latter two are artifacts and are very misleading when these data are analyzed for the ^{133}Ba source activity. For comparison, the results of PCT scan data without attenuation (i.e., no Cu cylinder or ^{228}Th passive source) are shown in **Fig. 5**. We are investigating the use of other gamma-ray spectrum processing methods that will improve the extraction of the net-peak data from noisy gamma-ray spectra. Using a better method is important, since we expect low-count-rate data to be the norm in mixed-waste drum assay scans containing LLW amounts of activity.

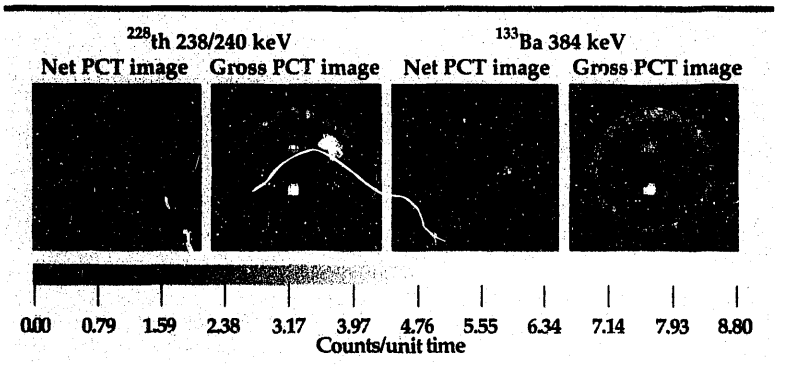


Figure 4. Corrected PCT images obtained by reconstructing the net and gross sinogram data, using the WLS/SD algorithm.

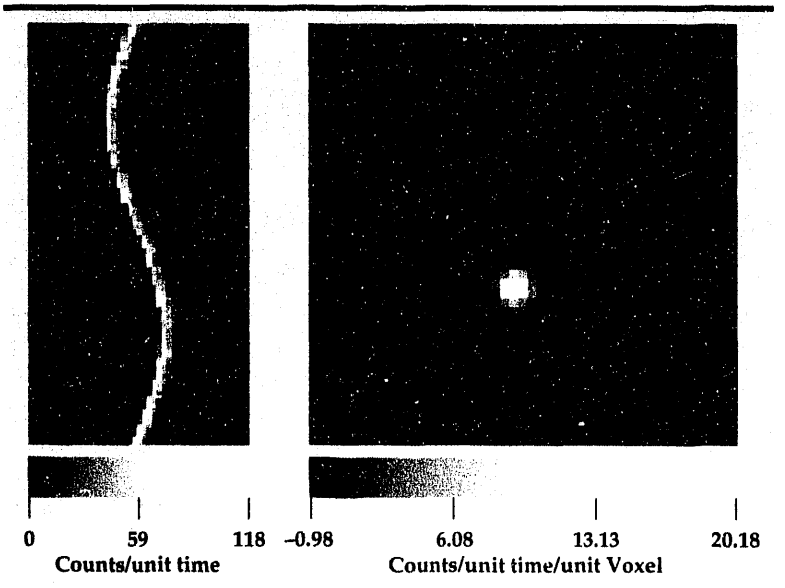


Figure 5. Representative unattenuated passive sinogram and PCT image data for the ^{133}Ba source at 384 keV.

Cone-Beam CT Research

We are expanding our research in cone-beam CT imaging methods. We begin this section with an introduction to cone-beam CT imaging, and follow with our progress during FY-92.

Computed tomography of the 1970's and early 1980's has been an inherently two-dimensional (2-D) process. Typically, a single detector or linear [one-dimensional (1-D)] array of detectors is used to gather x ray-attenuation transmission gauge measurements through the object under inspection. One gauge measurement is called a ray sum; multiple ray sum measurements along a single line are called a projection. Many projections are acquired at various angles about the object, but always through the same cross-sectional plane (see **Fig. 6a**), creating a 2-D data set called a sinogram. Image reconstruction of such 2-D sinograms usually involves filtering and backprojection operations that are well characterized and understood, and results in 2-D, cross-sectional images.

With the ready access of microfocus (spot sizes of about 1 to 50 μm) x ray machines, good 2-D (planar) x ray detectors and improved video technology in the last decade, CT research has concentrated on direct 2-D projection (or radiographic) measurements and one-step, 3-D, volumetric image-reconstruction methods. This speeds up the data-acquisition process (since multiple slices are acquired simultaneously) and results in a more efficient use of the x ray source. Until recently, 2-D projection measurements were acquired with the x ray source far from the detector, so the radiation penetrates the object with parallel-beam rays and standard 2-D image-reconstruction methods could be used (**Fig. 6b**). Currently, the more interesting case is to acquire projection data with the source close to the object and detector. This mode results in cone-beam x ray imaging (**Fig. 6c**). Cone-beam CT allows the use of geometric magnification to improve spatial resolution, and it makes the most efficient use of the source radiation. Problems to be solved with cone-beam CT include scanner alignment and the need for more complex image-re-

construction algorithms. We have addressed these problems and will show some results throughout the following sections.

Cone-Beam CT Scanners. Several of our CT scanners are inherently cone-beam systems.[2,5] Our recent collaborative work on cone-beam image-reconstruction methods (discussed below) has enabled us to use these scanners in a variety of new applications, and to take full advantage of improvements in source and detector technology. This has resulted in a better understanding of the components involved in cone-beam scanners.

A crucial component of many cone-beam scanners is the scintillator, the mechanism for converting the x ray photons into visible light. Recent developments in glass scintillator fabrication and manufacture provide the potential for increased performance of our lens-coupled, camera-based scanners such as the MicroCAT and the high-energy CT scanner (HECAT). To examine the possibility of increased scanner performance, we have applied our lens-coupled scanners to evaluate this new type of scintillator glass, called Lockheed high density (LHD) glass. In general, we seek to establish the spatial resolution, speed, energy response, and contrast limits of this glass for the wide variety of different digital radiography and CT applications we usually encounter within the Nondestructive Evaluation section.

The promise of this new glass is high brightness and high spatial resolution, a combination which is not readily available without substantial costs. High brightness can be obtained with off-the-shelf image intensifiers, or image-intensified charged couple devices (CCD's). Often these imagers include spatial distortions due to their physical shape, and image intensification components. Distortions must be accounted for before meaningful CT images can be obtained.[2] Furthermore, intensifier-based CT scanners can have other limitations due to the actual spatial resolutions that can beobtained.[1]

Initial measurements by other researchers have shown that the LHD glass composes a flat-field, distortion-free projection image with at least ~ 25- to 35-μm (or 20- to 14-lp/mm) inherent spatial resolution, and is bright enough for standard, non-intensified, visible-light CCD cameras.[16] We are currently evaluating a Cohu 4910 camera, operated in two modes, variable integration mode and RS-170 mode. We have also used the ER210 Image-Intensified Silicon-Intensified Target (SIT) camera and a Photometrics CH200 CCD-based camera to further explore the properties of the LHD glass at high energies (4 and 9 MeV). If the reported properties can be realized routinely on general purpose, lens-coupled, x-ray imaging systems, with a variety of cameras, then a relatively inexpensive and high-performance, x-ray imaging alternative has been established.

To evaluate the new scintillator glass over a broad range of energies, we ran three separate series of experiments at the following energies: (1) at 90 to 130 kVp, using the MicroCAT scanner; (2) at 200 to 320 kVp, using a PHILLIPS 320-kVp medium energy source; and (3) at 4 and 9 MeV using two VARIAN Linatrons and the HECAT scanner detector. In all three experiments, we used the scanners in a similar fashion to that with their original scintillating glass materials. The only difference was that the original scintillators were replaced with the new, clear, LHD glass scintillators.

Our low-energy studies have produced data with spatial resolutions (~ 14 to 20 lp/mm) consistent with earlier results,[17] but with a slight variation in technique. Scintillating glass can be fabricated as a clear sheet, or drawn into fiber-optic bundles. We are interested in detectors that can support cone-beam imaging modes, and consequently restricted our evaluations to the clear, LHD scintillator material. It is known that cone-beams of x rays impinging on a fiber-optic scintillator will induce cross-talk in the fibers, and result

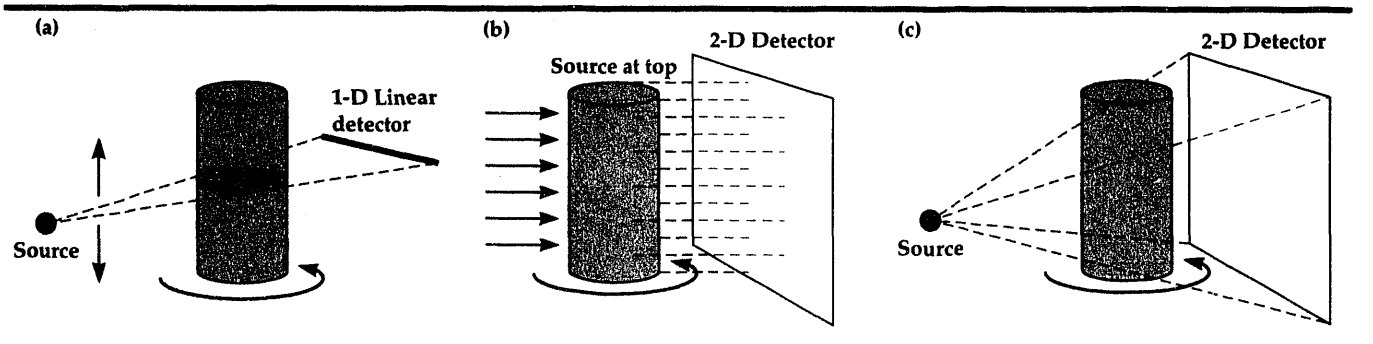


Figure 6. Methods of acquiring 3-D CT data: (a) multi-slice CT by acquiring each 2-D cross-sectional slice independently; (b) 3-D parallel CT with a 2-D detector and source at infinity (one rotation and standard reconstruction codes are used); (c) cone-beam CT by moving the source close to the object. New reconstruction algorithms are needed to process this data.

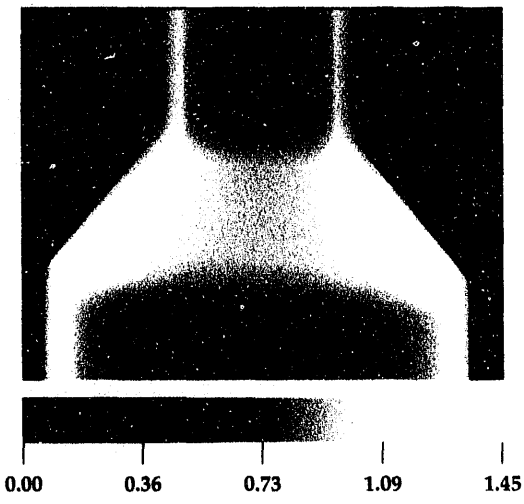
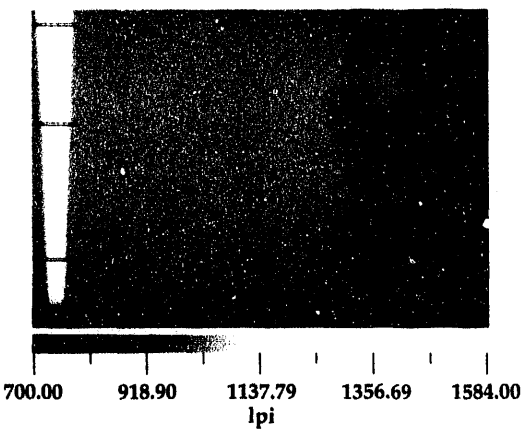


Figure 7. Digital radiograph of hollow thermocouple plug, taken with MicroCAT system.



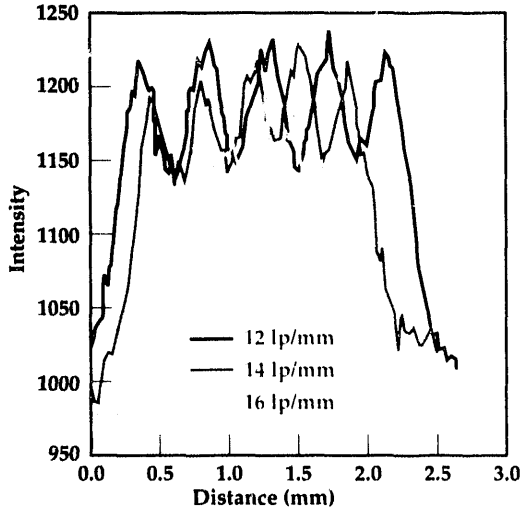


Figure 8. The spatial resolution performance of MicroCAT studied by analysis of a line-pair gauge.

in a loss of spatial resolution as cone angle increases.[18] In addition, the unprocessed image can be viewed directly, without any subtraction of the cross-hatch in the fiber-bundle usually present in unprocessed images from fiber-optic scintillators. **Figure 7** is a digital radiograph of a hollow thermocouple plug cover, taken on our MicroCAT system. One surprising result of this investigation is the increased spatial performance obtained when the focal plane of the camera is positioned inside the scintillator glass, as opposed to the back face plane of the glass. The advantages of this technique are visible in the image in **Fig. 7**. However, this added clarity is at the cost of some brightness, since some of the scintillator material is not in focus for the camera.

The MicroCAT scanner's spatial resolution performance using the LHD glass scintillator was further studied by analysis of a line-pair gauge (**Fig. 8**). There was almost no magnification involved in this exposure (M = 1.1), and with our Micro-Focal system (10-μm spot sizes), blur due to finite spot size is effectively eliminated from the system.[19] As illustrated in this figure, the loss in spatial modulation is less than 50% from 12 to 14 lp/mm, and there is significant modulation at 16 lp/mm. With the micro-focal source, magnifications of 2 to 3 can be achieved with a minimum of source blur, which can easily extend the spatial resolution of this scanner into the 18- to 20-lp/mm range. We have scanned different objects, and our results have shown increased performance compared to scans with the fiber-optic scintillator used previously. A comparison of the enhancement in the new LHD glass scintillator and other differences with the older fiber-optic scintillating glass for a pinch-weld CT study are shown in **Fig. 9**.

The spatial resolution performance reported by others[20] has been substantiated by some of our preliminary studies using the LHD glass scintillator at medium and high energies as well. For example, a representative digital radiograph of a ~3-cm turbine blade acquired with a lens-coupled, Cohu 4910, CCD camera detector system and a 270-kVp x ray spectrum from a Philips medium-energy machine source (spot size ~0.4 mm) is shown in **Fig. 10**. We estimate the spatial resolution of this digital radiograph to be on the order of 7 to 10 lp/mm

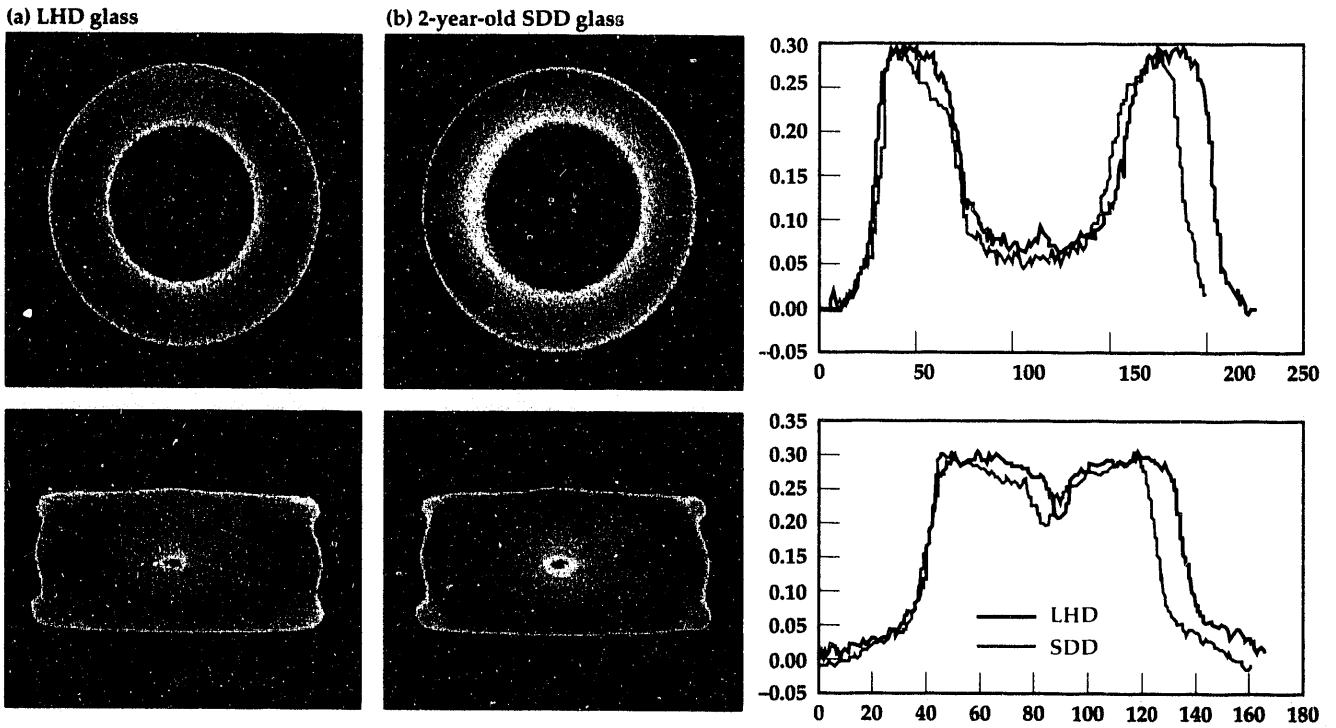


Figure 9. Comparison of enhancement in (a) LHD glass scintillator and (b) SDD fiber-optic scintillator for a pinch weld. 1-D profiles through the center of each image to the left are shown to the right.

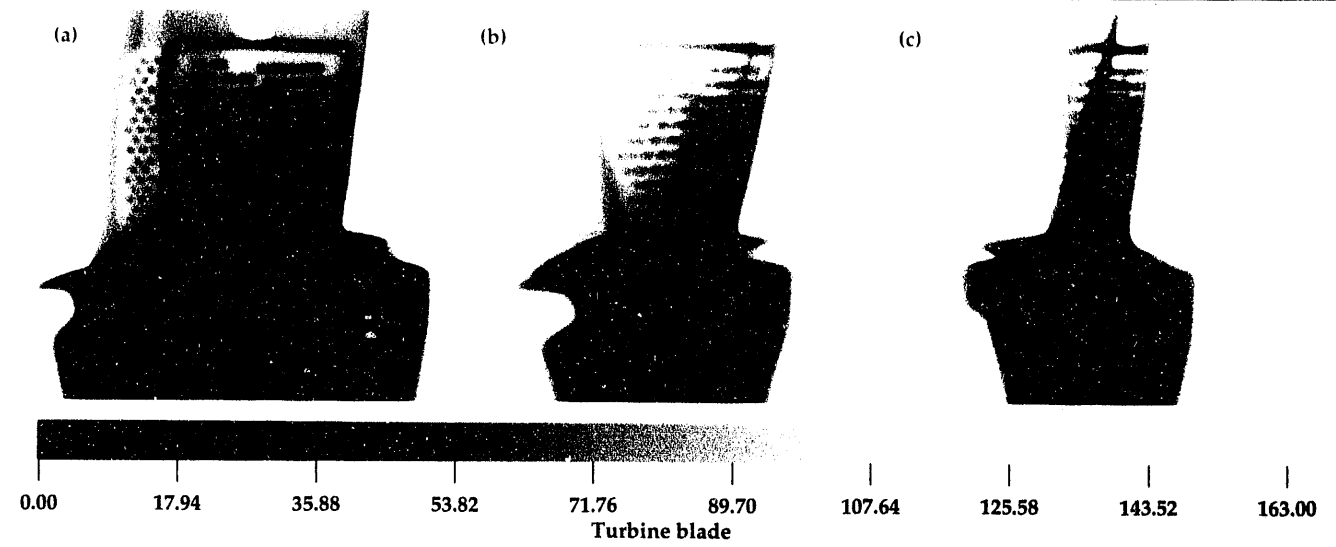


Figure 10. Representative medium-energy radiographs of turbine blade at (a) 0 , (b) 45 , and (c) 90 .

The results of the performance of the LHD glass at higher energies is shown in a pair of digital radiographs of a 5.0-cm doublet, single-crystal turbine blade acquired at both 4 and 9 MeV, using two different Linatron sources and a lens-coupled, Cohu 4910, CCD camera-based detector system (**Fig. 11**). The pedestals and features of the blade are on the order of 1 mm in spatial extent. More detailed explanation of the turbine blade study is given below. We are attempting to further quantify the spatial resolution limits of this new glass for both medium- and high-energy CT scanning applications. Our preliminary CT results for high-spatial-resolution, high-energy applications have been encouraging. **Figure 12** is a sample of images from a 3-D cube of data acquired with a 9-MeV Linatron on HECAT; they show the doublet turbine blade, measuring wall thicknesses of 500 µm.

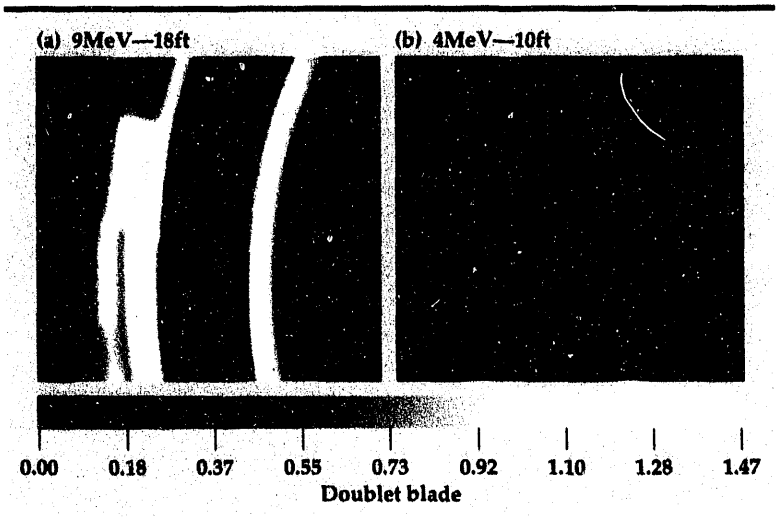


Figure 11. Representative high-energy radiographs of a doublet single-crystal turbine blade at (a) 4 MeV and (b) 9 MeV.

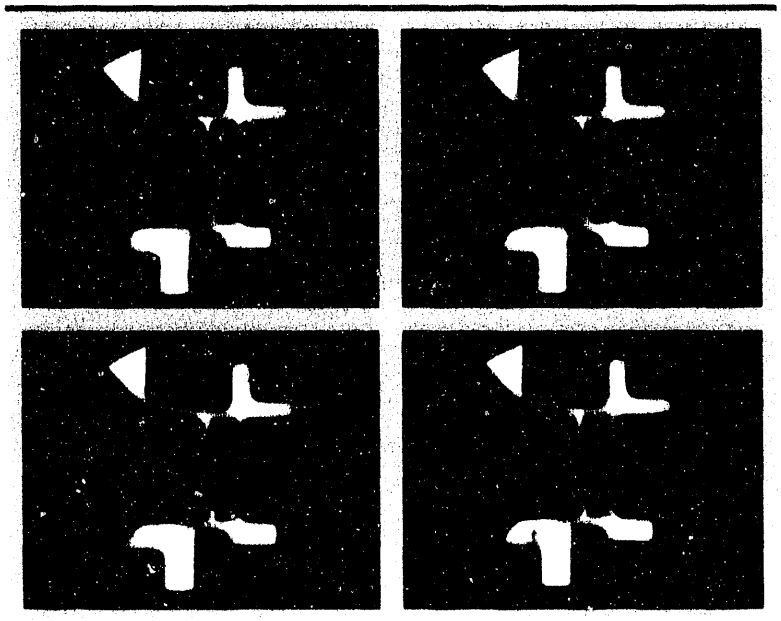


Figure 12. Representative 2-D images from a volume (3-D) CT data set of the doublet single-crystal turbine blade, acquired using the HECAT detector and a 9-MeV LINAC source.

In the next year, we will better quantify the spatial and contrast resolution performance of this new glass and the improvements this yields for lens-coupled, cone-beam CT scanners. With this proof-of-principle work as a base, and in cooperation with LLNL physicists working in astronomy, we are assembling a high-performance CCD camera with 14 to 16 bits and 2048 × 2048 detector elements, which can further explore the properties of this glass, and provide a higher performance, 2-D, cone-beam CT scanner. We will also examine this glass in slit-collimated configurations and with linear-array detectors as a means of obtaining high-contrast, high-spatial-resolution images that include less scatter.

Cone-Beam CT Reconstruction Technologies. We have implemented cone-beam reconstruction methods of others[21–25] and developed our own, all with good results.[1,2] Last year, one member of our research team (S. Azevedo) worked in France at LETI on new cone-beam methods that benefited both CT projects. New reconstruction algorithms and scanning methodologies were developed during the course of this collaboration, and a patent is pending in conjunction with the French government.[26] This and other cone-beam work are continuing in FY-93. Below, we describe some of the progress during FY-92 in image reconstruction technologies, including region-of-interest cone-beam CT, axi-symmetric CT, reverse cone-beam geometry, and a fast image-reconstruction processor.

Region-of-interest cone-beam CT. It is often necessary to view a part of an object at higher magnification than is needed over the rest of the object. Also, sometimes the object is too large for our cone-beam scanner. In these cases, the data we acquire will be 'limited'; i.e., there will be missing ray paths from our projection measurements. This type of reconstruction problem is called 'region-of-interest' (ROI) CT and is a common problem in medical and industrial imaging. There have been solutions proposed for 2-D ROI imaging, but not for the 3-D cone-beam case.

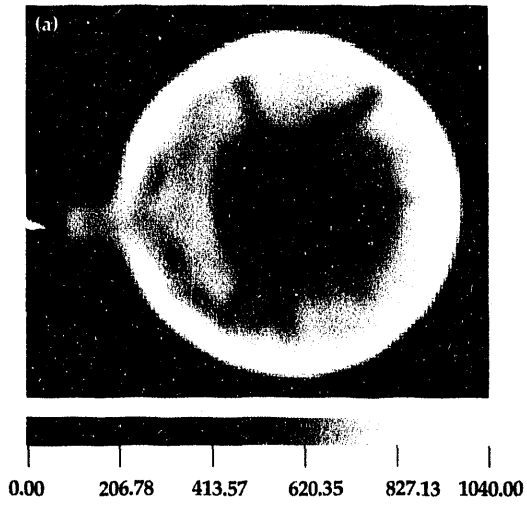
Our algorithm for reconstructing cone-beam ROI data, called Radon-ROI, uses mathematical methods similar to the Grangeat method of cone-beam image reconstruction. In Grangeat's method, the 2-D radiographic projections are mathematically converted (through weighting, filters, and re-binning steps) to an intermediate mathematical space known as the 3-D Radon domain. From this space, reconstruction of the volumetric

image is straightforward, requiring only two sets of backprojections. The Radon space is an ideal place for combining data of different resolution, so it is ideal for ROI imaging. Two scans of the same object are acquired at different resolutions, a low-resolution scan covering the entire object and a high-resolution scan covering only the ROI. These two scans are combined to form a single Radon space, which is reconstructed by the latter part of Grangeat's algorithm to form a final volume that displays the ROI at higher resolution than the surrounding part, without significant artifacts.

An example of the use of this method for ROI scanning of an automotive precombustion chamber is shown in **Fig. 13**. The image in **Fig. 13a** shows a single reconstructed slice through the object taken at low resolution. The data were acquired and reconstructed using cone-beam CT methods. The object has some shrinkage cracks just barely visible in the interior. A second scan of the precombustion chamber was acquired at a higher spatial resolution, but of the ROI only. The two scans were combined and reconstructed into a second image, as shown in **Fig. 13b**. The ROI area of this second image reveals the much higher spatial resolution obtained by this novel method.[27]

Axi-symmetric CT. We have applied the cone-beam CT reconstruction methods to the problem of obtaining 3-D exterior and interior information from axi-symmetric objects with only one 2-D radiograph or projection. This has application in several areas, such as manufacturing and high explosives testing. The problem is to perform cone-beam CT reconstruction of an object that has quasi-axial symmetry from a single radiographic view. For example, a high explosive can be radiographed during firing, with a flash x ray unit, but only one view is available. If we assume axial symmetry, greater information can be gathered from the single radiographic view. We have performed cone-beam image reconstructions of such tests and of simulated data to better understand the combustion mechanisms. Simulations provide useful information as to what kinds of artifacts to expect from any asymmetries in the object.

Another example is on-line monitoring of high-volume, axi-symmetric industrial parts. As an example, we applied this technique to a diesel engine piston. A single radiographic view of the piston was obtained (**Fig. 14a**) and reconstructed into a volume image. A representative 2-D cross section of this resultant volume is shown in **Fig. 14b**. With many such pistons being fabricated (as many as 2000 per hour), complete CT methods may be impossibly time-consuming. However, using the



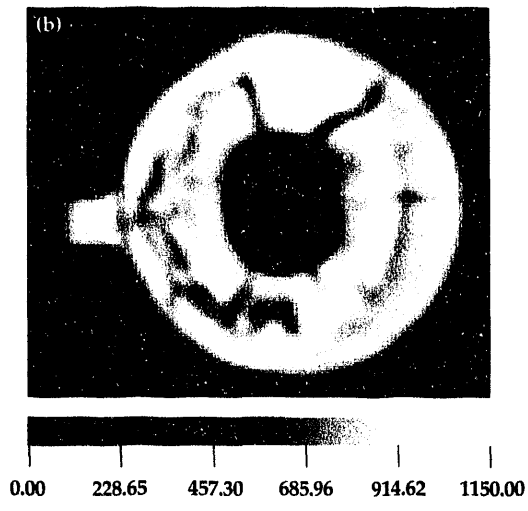


Figure 13. ROI image showing (a) single reconstructed slice at low resolution and (b) combined low- and high-resolution data reconstructed image. This is one slice out of a 3-D volume acquired by cone-beam methods. Notice the higher spatial resolution in (b). (Data courtesy of LETI, Grenoble, France.)

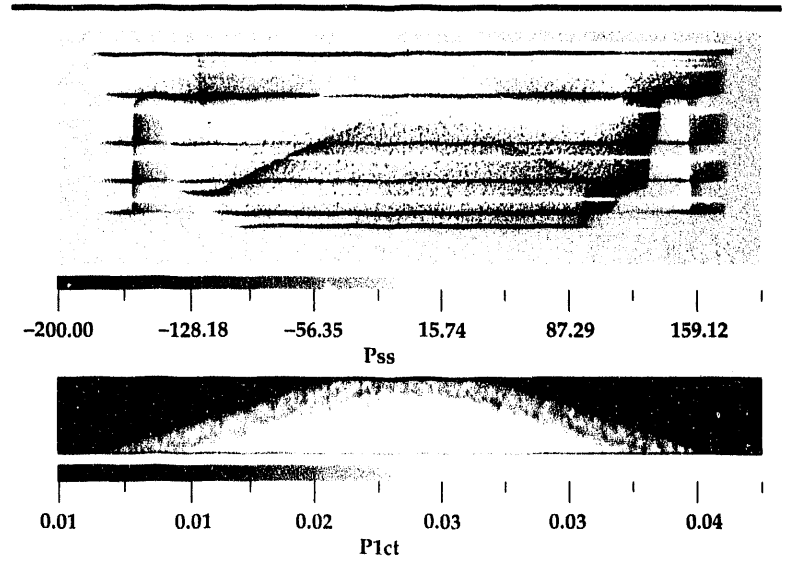


Figure 14. Representative (a) 2-D projection radiograph and (b) 2-D CT images of a diesel engine piston. The data are extracted from a volumetric image obtained from a single 2-D view.

above cone-beam reconstruction method, a single radiograph can display much more useful information about the part. For example, in **Fig. 14b**, small cracks within the interface between two different materials are much more visible than from the radiograph shown in **Fig. 14a**.

Reverse cone-beam geometry. A small company in San Ramon, California, called DigiRay, has developed a new radiographic method called "reverse geometry" cone-beam radiography. In this method, the source is a 2-D, raster-scanned flat panel, while the detector is a single element. The source raster defines the acquisition geometry, which is essentially a cone (**Fig. 15**). This system is unique and provides some inherent advantages over conventional, cone-beam, x ray-imaging systems. The use of a converging beam has the potential to produce 2-D x ray transmission images with little to no scattered photons. The DigiRay system also includes a NaI(Tl) detector that could be configured to acquire energy-specific data. Until recently, this system has been used exclusively for industrial radiographic applications, not for CT.

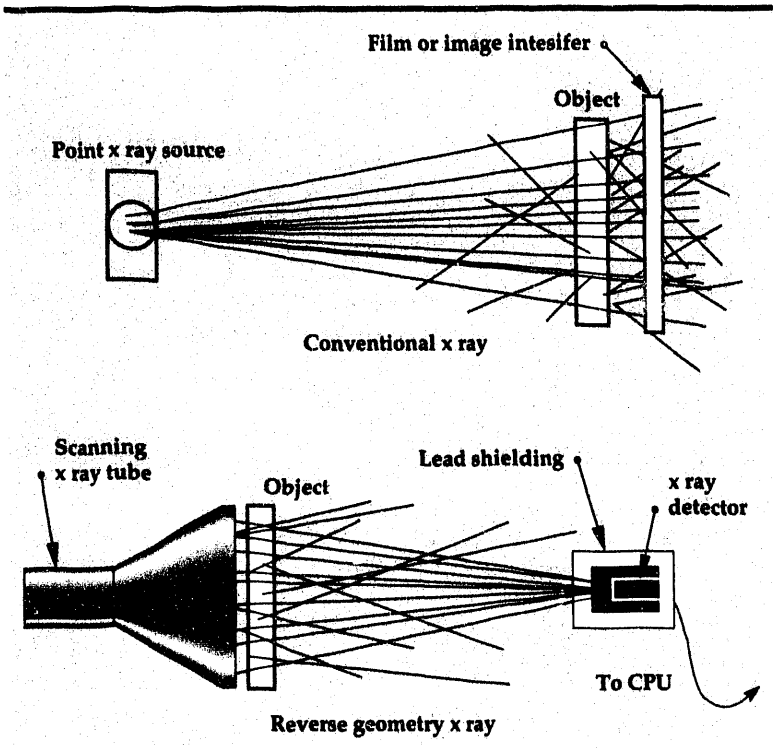


Figure 15. Comparison between (a) conventional cone-beam-projections data-acquisition and (b) reverse geometry, cone-beam-projection data-acquisition systems.

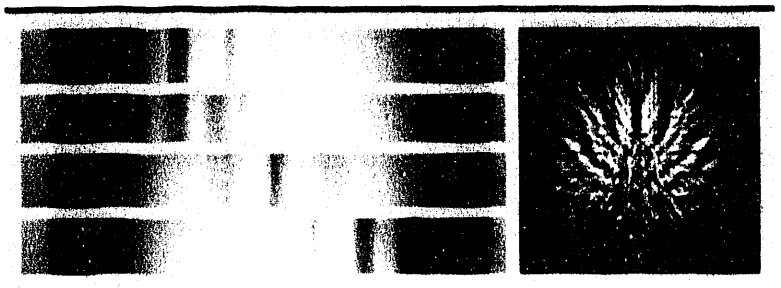


Figure 16. Representative (a) projection data and (b) CT image of a MTF phantom using reverse geometry data acquisition. (Projection data courtesy of Digiray, San Ramon, California.)

We have been evaluating the efficacy of using their unique method[28] for industrial, cone-beam, CT imaging applications. DigiRay acquired 24 2-D, inverse-geometry, cone-beam projection images as a function of angle (every 15°) for a lexan modulation-transfer-function (MTF) phantom. Representative 2-D projections are shown in **Fig. 16**. These projection data were reconstructed using a parallel reconstruction algorithm. A resultant CT image is shown in **Fig. 16**. It is difficult to determine just how useful the inverse geometry scans are from the CT images, since these data have moiré artifacts, due to the limited number of angular projections, that mask the scan results. In spite of these results, we still expect that this type of system could produce high quality 3-D imaging with little to no scattering artifacts. We are working with DigiRay to set up an experiment in which we can obtain more angular projection data that should result in CT-reconstructed images without moiré artifacts. With the use of an energy-discriminating detector system, further enhancements are expected into the regime of materials characterization.

Reconstruction hardware. Another important problem in cone-beam CT imaging is the speed of the image-reconstruction codes and architecture used. We have addressed this problem in a joint research project with a private company, Advanced Research and Applications Corporation (ARACOR) of Sunnyvale, California. In this project,[29] an advanced computational engine, called the Konoscope reconstructor, was designed to reconstruct large cone-beam data sets in reasonable computation times (within 2 h of data-acquisition times) while being low in cost. This design was successfully completed in FY-91 and was realized as a hybrid, parallel, multi-processing system.

In the last year, we have purchased, assembled, and tested the basic building blocks of this system. This prototype was constructed to demonstrate and measure the performance of various algorithms operating on the reconstructor. Several cone-beam reconstruction algorithms will be coded on the system for evaluation during FY-93. Also, the system will be used to evaluate other image-processing algorithms that may benefit from this unique design.

High-Energy CT Research

High-Energy CT Scanners. We completed the first version of the HECAT scanner in FY-92. This scanner incorporated the features and flexibility of the video-camera-based CT (VIDCT), area detector software, and built upon our past experience with film radiography using 4- and 9-MeV VARIAN Linatron sources. HECAT is currently an area detector-based scanner that can acquire data in either a variable integration (typically from 2 to 10 s), or RS-170 video-frame-rate, data-acquisition mode. Two different area detector systems have been used to acquire the CT projection data: (1) a VARIAN ER210, image-intensified SIT camera and (2) a COHU 4910 CCD camera. Both are lens-coupled to either a fiber-optic scintillator bundle or to a piece of the LHD clear glass via a visible-light 90° bending mirror. We have used the HECAT scanner to perform 3-D CT inspections of bridge members, engine parts, ceramic-metal castings, and single-crystal turbine blades.

CT systems contain four components: (1) source, (2) detector, (3) object manipulator, and (4) data acquisition and image-reconstruction and analysis computer. One of the challenges for a high-energy CT system is to reduce the effect of source blur (a blur from a finite source-spot size). High-energy sources typically involve relatively large source-spot sizes (2 mm in our case). Source blur, to first order, increases linearly with x ray magnification.

A good rule of thumb is that the source blur, δ, is equal to the spot size, A, times magnification, M, minus one, i.e., $\delta = A(M-1)$. Consequently, the source blur due to finite spot size[19] can be quite large at moderate magnification (nominally 1 mm at a magnification of 1.5). Large objects are difficult to position very close to the detector, by their very size. Also, object manipulators that can support 300 to 1000 lbs are not small, and it is difficult to minimize the source detector distance if the manipulator cannot be fixed below the source detector. It is also useful to point out that to penetrate highly attenuating sections of an object, the source must be positioned close to the detector (to increase the effective flux per volume throughout the object), which increases the magnification for a fixed object to detector distance.

Recently developed objects (e.g., single-crystal turbine blades) are small, but contain highly attenuating materials. Consequently, there is a need for high-energy, high-spatial-resolution CT scanners. This need will continue to increase as metal manufacturing achieves new levels of complexity and precision.

HECAT (**Fig. 17**) accommodates these demands in a number of ways. First we developed a highly flexible object manipulator (or stage) interface, fixturing for two different object manipulators (1) a small, 15-cm o.d., rotation-translation stage that can hold up to 25-kg objects, and (2) a rotation, translation, elevation and tilt stage that can support up to ~ 350 kg and added fixturing that enables the position of the stage and scintillator face front to be in a variety of positions. The small stage is approximately 7.6 cm from the scintillator face front, while the large stage, with a 46-cm o.d. rotational table, is a minimum of 25.4 cm from the scintillator. The fixturing for the scintillator is adjustable for a travel of 30.5 cm. The camera is seated on a NEWPORT optical rail, and can be adjusted up to 20 cm inside the leaded enclosure to enable a variety of fields of view. Using these adjustments and different lenses, we can obtain fields of view from 5 × 5 cm to 28 × 28 cm. This scanner has been built with the flexibility to allow any object to be positioned as close to the detector as physically allowable. We have performed scans with this system, for different sized objects with cone angles up to 4.8°.

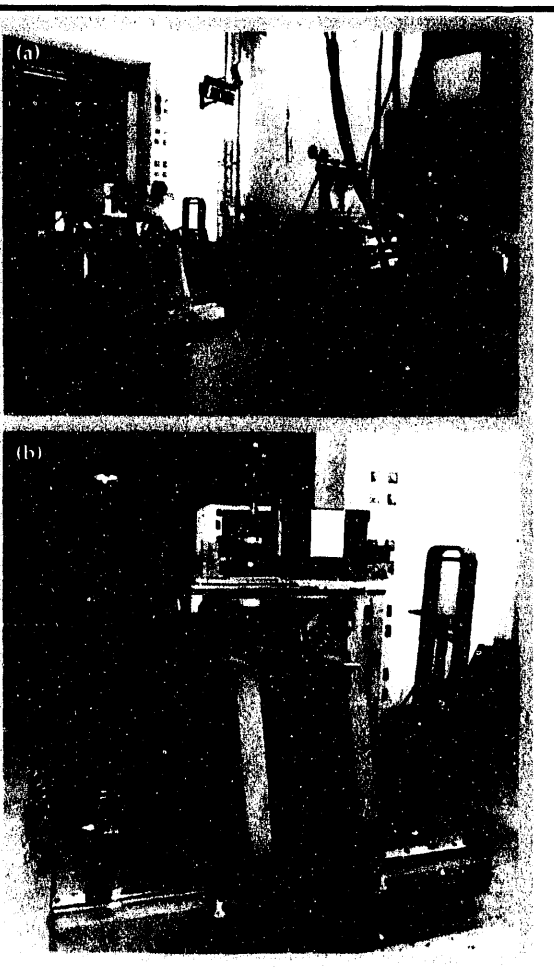


Figure 17.
HECAT photos.

This flexibility notwithstanding, one continuing limitation of this system is the relatively small maximum field of view (28 × 28 cm) of the system as a whole. One of the advantages of the 9 MeV Linatron is the ability to penetrate objects with dimensions much greater than 28 cm. The field-of-view limitation is only a result of third-generation (rotation only) scanning geometry and can be overcome by implementing and applying second-generation (translation-rotation) scanning techniques. The latter technique extends the field of view to the total distance traveled of the object manipulator. We are pursuing this enhancement to extend the capability to HECAT in the next fiscal year.

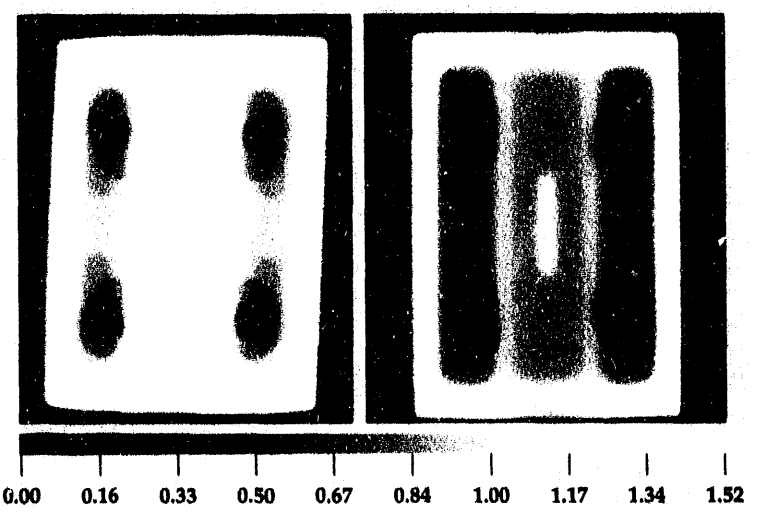


Figure 18. Representative digital radiographs of a prototype portliner from Caterpillar, Inc. The radiographs were obtained using a 9-MeV source with (a) the Varian ER210 camera-based detector; and (b) the Cohu camera-based detector. Each was lens-coupled to the LHD glass.

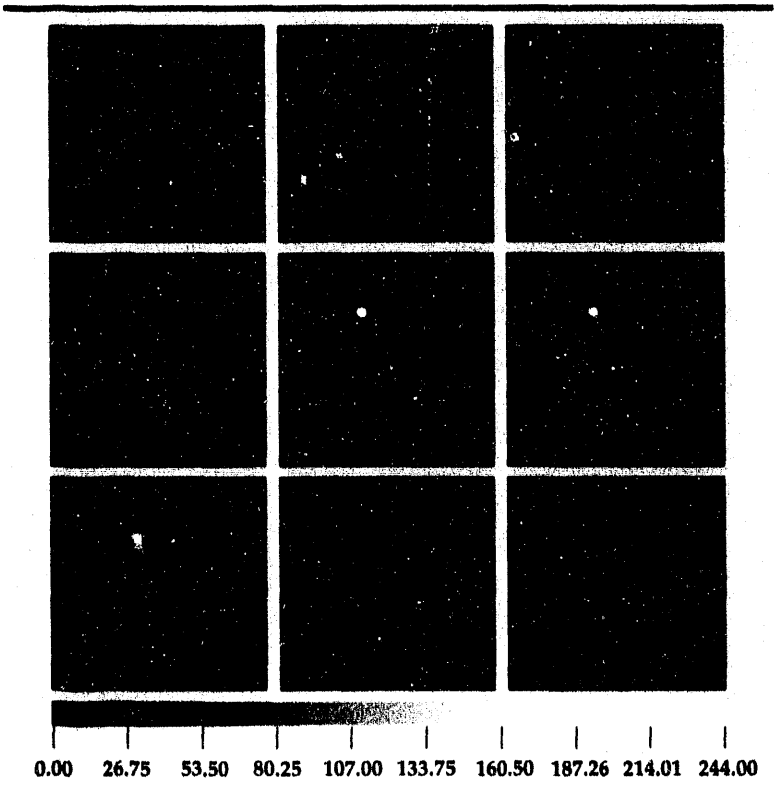


Figure 19. Representative 2-D images along the z axis from a volume image of the track of a tungsten bullet. This data was acquired with the ER210 camera and the 4 MeV source.

We have used the HECAT scanner to directly evaluate different high-energy scintillators and visible-light cameras. A comparison between the two HECAT cameras is summarized by two digital radiographs of a Caterpillar portliner. Both radiographs were acquired at 9 MeV and are shown in **Fig. 18**. Note that the CCD camera results have an increased performance over the SIT camera results.

High-Energy CT Applications. Most of our CT research has been within the low- (6 to 250 keV) to medium- (250 to 1,300 keV) energy range. We are now studying the effects of high-energy x ray beams on the available detectors, and applications that require high energies for penetration.

One application for high-energy CT is the precision tracking of tungsten projectiles in target materials. Three-dimensional CT methods can generate dimensionally accurate images of the path of a projectile through a target in all three dimensions, showing the changes in trajectory and in the character of the projectile as it passes through the object. **Figure 19** contains a set of 2-D images from a volume image in one orientation, while **Fig. 20** is another set of 2-D images for a different orientation. From these images, it is particularly interesting to study just how the projectile changed its direction by 180° to point in the direction of the initial momentum. This new capability provides an unambiguous image of the path of the projectile through the target medium.

High-energy CT has been applied to various diesel engine components as a part of a Cooperative Research and Development Agreement between LLNL and Caterpillar, Inc. The goal of this project is to combine the NDE and computational resources and expertise available at LLNL, with the diesel-engine-design and manufacturing expertise of the Caterpillar Corporation to develop in-process monitoring and inspection techniques for diesel-engine combustion chamber components and materials. Early development of these techniques will assure the optimization of the manufacturing process by design/inspection interfaces. Project goals include (1) to improve the efficiency of diesel engines; (2) to meet or exceed new environmental regulations; and (3) to develop inspection and process control technology for the production of advanced materials for improved diesel engines.

Components under study range in size from 2-mm o.d. fuel injector tips to 30-cm-x-36 cm-x-1.2 m cast iron exhaust manifolds. Most of the effort to date has involved the interrogation of cast iron exhaust assemblies and prototypes (portliners). The largest of these devices is complex and non-symmetrical, with nominal outside dimensions of 30 cm × 36 cm × 1.2 m. Our preliminary work has focused on a sample subsection of the exhaust manifold assembly, a 12-cm-x-15-cm-x-16-cm box with an inner configuration of outlet holes and ceramic sleeving.

Our research has shown a substantial increase in spatial performance of the COHU CCD camera as compared to the Varian ER210 with both coupled to the LHD glass (**Fig. 18**). The port running down the height of the object is lined with a ceramic material, which has numerous divots and cracks. A representative cross-sectional CT image of this object is shown in **Fig. 21**. The cross-sectional slice data reveal the ceramic-metal interface and features in the ceramic material.

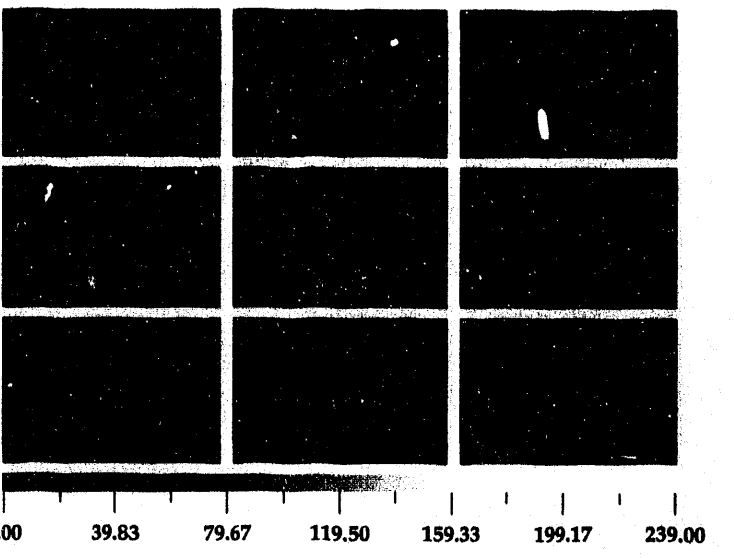


Figure 20. Representative 2-D images along the y axis from the same volume image described in Fig. 19.

Additional CT Applications

In this section, we describe some additional NDE research problems investigated during this fiscal year that are not published elsewhere.

Shape Charge. We have performed proof-of-principle CT scans on a conventional munitions shape charge to show how revealing CT is in identifying internal flaws nondestructively. To meet this end, we fabricated a few plastic inserts to mock air voids (four sets of hollow cylinders 1-, 2-, 3-, and 4-mm diameter), and 1- and 2-mm mock separations of the explosive from the copper shape charge. These plastic inserts were fixed to the explosive side of the copper wall, and the charge was filled with a mock plastic explosive.

MECAT was used to acquire CT projection data of this shape charge at 70 mm from the bottom of the charge. A summary of these results is shown in **Fig. 22**. The image on the left is the resultant image reconstructed from the projection data set. The CT image or tomogram represents a cross-sectional view of the shape charge along its longitudinal axis, with 1-mm spatial resolution and a slice-plane thickness of 1 mm. The colorbar shown here in shades of gray relates colors in the image to the linear attenuation coefficient in cm^{-1}.

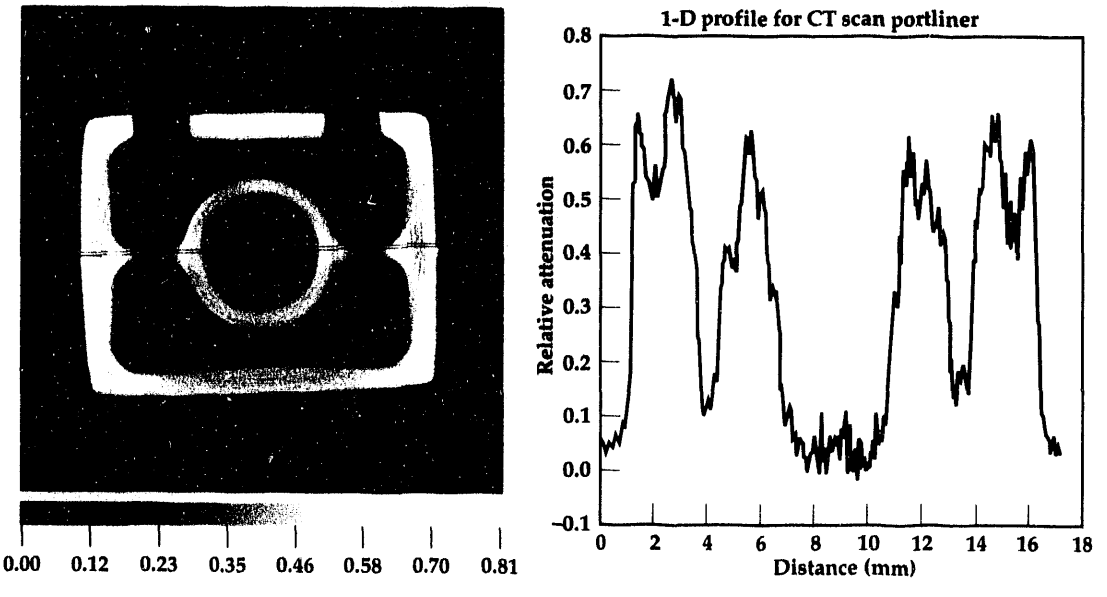


Figure 21. Representative 2-D image from a volume-image data set of the Caterpillar port liner. A 1-D profile shown by the black box is plotted to the right.

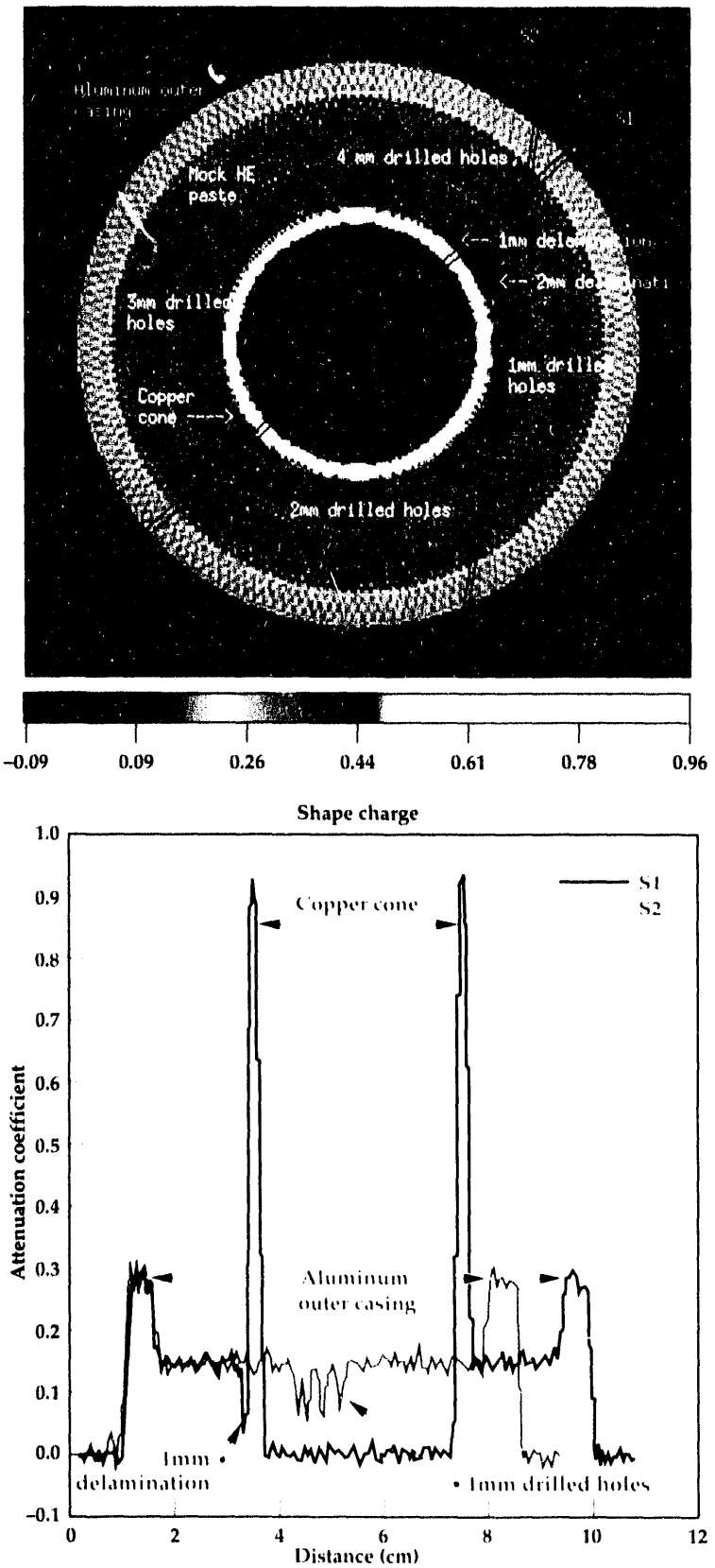


Figure 22. CT slice of a mock plastic-high-explosive shape charge with various defects. 1-D profiles are shown to the right.

The linear attenuation coefficient is related to atomic number and density of the shape charge. These results clearly reveal all four sets of voids and both the 1- and 2-mm separations between the copper wall and mock explosive. The tomographic image also provides the location and dimensions of both external and internal details of the shape charge, which is sometimes very difficult, if not impossible, to obtain from radiographic data. With respect to **Fig. 22**, two 1-D profiles labeled in the image as S1 and S2 were extracted and are displayed in the plot to the right. The next step in this study would include the correlation of the CT images to actual performance of the shape charge, determined in live test firings.

Turbine Blades. Both high- and medium-energy CT are being applied to a variety of single-crystal, hollow-core turbine blades (**Fig. 23**). These blades are required in modern high efficiency, low pollution jet engines where gas temperatures can exceed the melting point of the blade metal. Southwest Research Institute, Garrett Engine, and the NDE Section at LLNL are combining resources and expertise to evaluate NDE techniques. The goals are to accurately dimension the internal geometries and detect anomalies intrinsic to hollow cast blades. Results of this project will be used to plan future funding activities for development of advanced NDE techniques. CT is a critical NDE technique for turbine blade inspection due to their complex geometry and high aspect ratio, which limit the nature of internal information obtainable from other techniques, such as radiography, infrared imaging, and ultrasonics. We are presently scanning each of the blades shown in **Fig. 23**, using the multitude of CT scanners discussed here

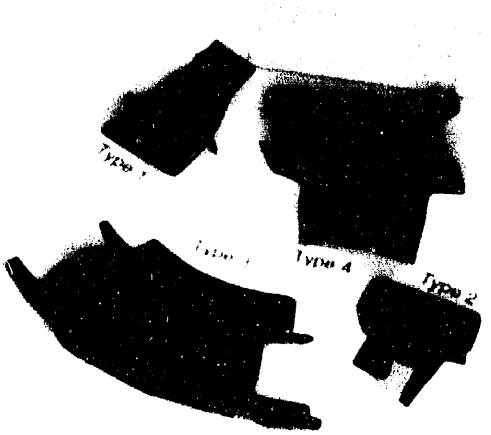


Figure 23. Photograph of four turbine blades provided by Wright Patterson Air Force Base for CT inspection studies.

and elsewhere.[5] Representative results are shown in **Fig. 24** for the type 2 blade using a medium-energy source, and in **Fig. 25** for the type 3 blade using a 9-MeV Linac.

Bridge Cable CT Imaging. In cooperation with the California Department of Transportation and High Energy Services Corporation (HESCO), Woodside,California, proof-of-principle experiments were performed in an attempt to image internal features of bridge cables and bridge cable terminuses. High-fidelity CT requires a rather large number of angular views (the ASTM recommends 1.5 times the size of the detector array in the horizontal direction). For a variety of reasons, most of which are related to functionality, components of civil bridges are highly attenuating with respect to x rays. Consequently, the internal inspection of these assemblies requires high energy (MV range) sources. Performing the radiography is complicated by the logistical issues regarding shielding the on-coming traffic from the radiation, or by developing reconstruction schemes that can obtain useful information from a limited number of views.

Using their in-depth knowledge of what can be done on bridges, HESCO acquired a limited-view CT projection (or radiographic) data set for application of our CT image-reconstruction techniques. The data set consisted of 24 film radiographs every 6° over 138°. We developed interpolation schemes, in concurrence with other algorithm work, to extend this data to 30 views over 180°.[30] Other image-processing and image manipulation tools, developed for film CT, were then applied to obtain a reconstructed image of the cable. **Figure 26** shows an example of the reconstructed images obtained using these methods. These results are encouraging enough to further pursue this method to inspect bridge cables.

Ancient Artifacts from Iraq. We have used CT to investigate two, corroded, ancient artifacts that were excavated in Iraq by University of California Berkeley archaeologists under the direction of Professor David Stronach. These artifacts, along with others, were exported to the U.S. with permission from the Iraqi archaeological authorities for scientific analysis. The artifacts are believed to be objects used for personal adornment. They were found on an ancient roadway of the Halzi Gate at the southeast corner of the city of Nineveh, the last capital of ancient Assyria. The artifacts were found among skeletal remnants in the sack (destruction) level of Nineveh dating back to 612 BC, around the fall of the Assyrian Empire. It is believed that the skeletal remnants are from an immense battle fought there.

The encrusted artifacts were nondestructively evaluated using a quantitative CT scanner to learn about their original composition and geometry.

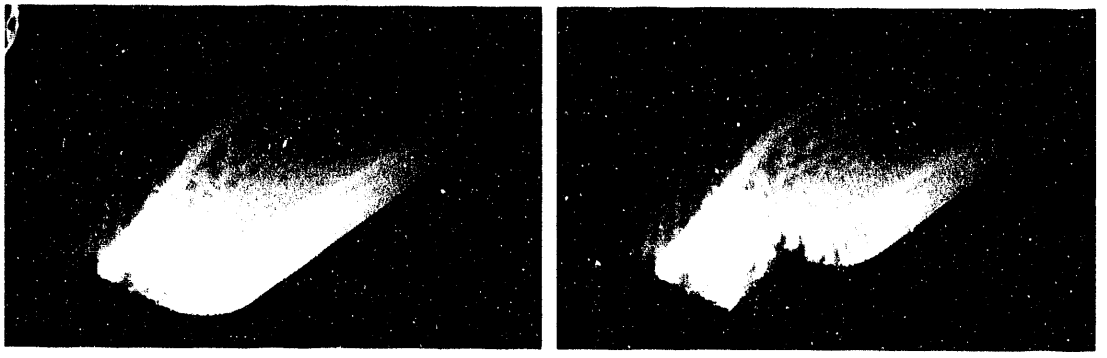


Figure 24. Representative CT images for the type 2 turbine blade, using a medium-energy source at 270 kVp and the Cohu camera coupled to the LHD glass.

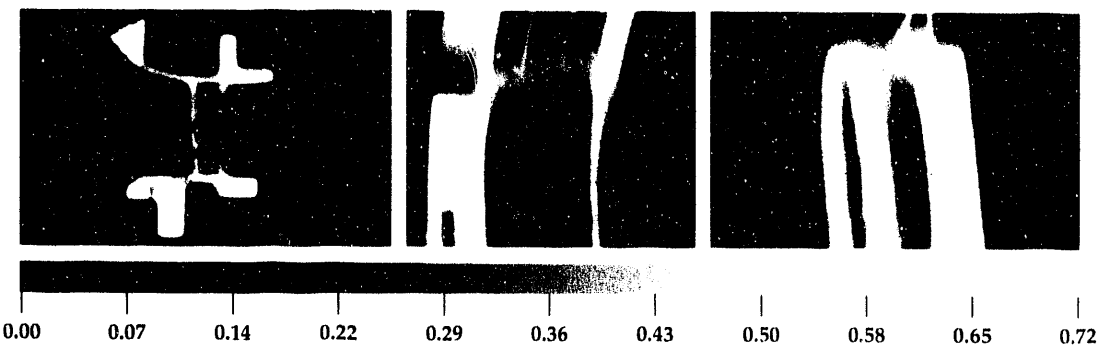


Figure 25. Representative CT results for the type 3 turbine blade, using a 9-MeV Linac source and the Cohu camera coupled to the LHD glass. Shown are 2-D images along the z, x, and y axes, respectively, of a volume data set.

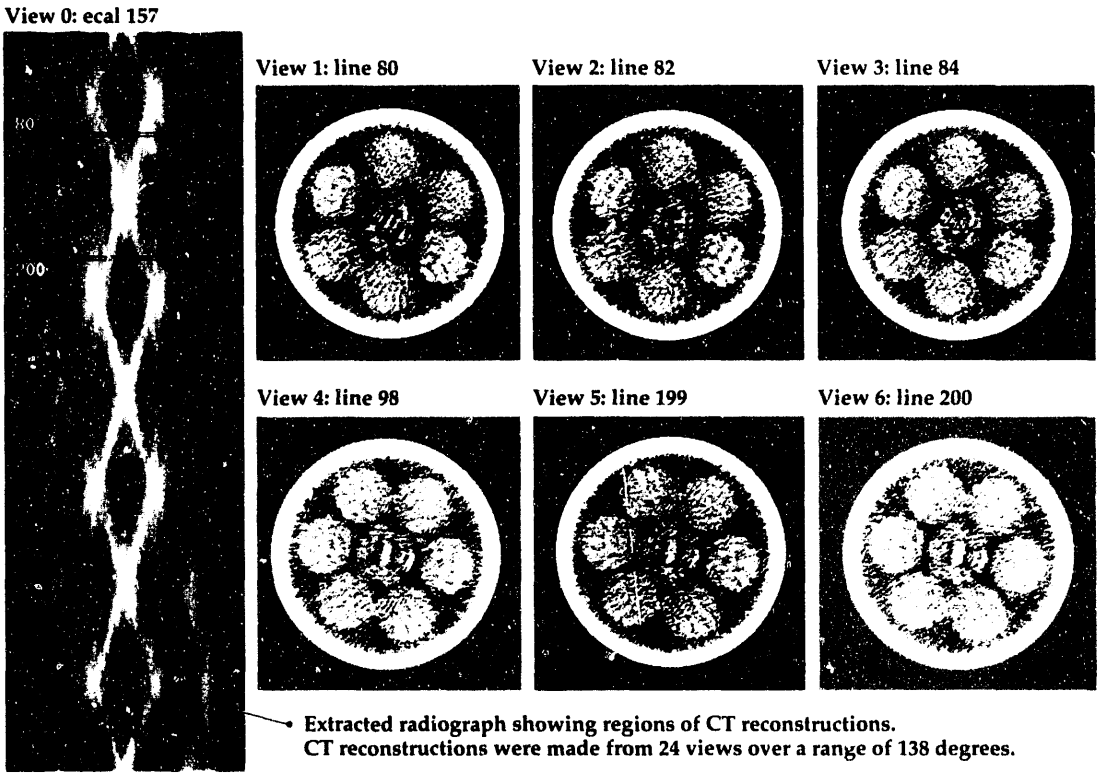


Figure 26. Representative radiograph and reconstructed images of a bridge cable. The CT data consisted of 24 views over a range of 138°. (Radiographs courtesy of HESCO, Woodside, California.)

The two artifacts are show in **Fig. 27** and have the geometry of a crescent moon and a ring. MECAT was used with an ^{192}Ir radioactive source and an energy-discriminating, high-purity, germanium detector system. The CT data were collected, with the ring placed within the concave edge of the crescent moon. A representative reconstructed image of a plane through the approximate midplane of the two artifacts is shown in **Fig. 28**. The data used for reconstruction is the measurement of x ray photon transmission from a narrow 317-keV peak of the ^{192}Ir source. The reconstruction represents a spatially resolved image of photon absorption at this specific energy.

The CT images of these artifacts have helped the archaeologists determine the extent of corrosion sustained over the years. CT has provided information that can help determine what material the artifacts are made of without destructively removing the encrusted substance around them. Also, the original geometry of the artifacts can be extrapolated from the CT images, without removing the encrusted material.

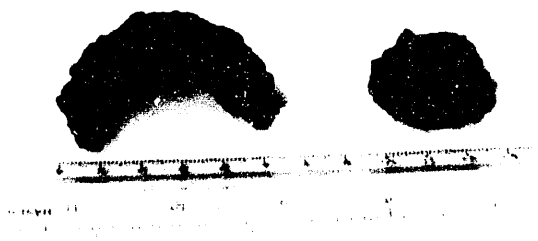


Figure 27. Ancient artifacts found in Nineveh, dated to 612 BC.

Future Work

Over the last few years, we have used our various CT systems on diverse experimental problems. In each case, CT has been demonstrated to be an effective tool for nondestructive inspection and/or characterization. We will continue to improve CT scanners built by our CT group and software tools extending the range of applications for CT technology. We expect to continue to participate in the transfer of our technology within LLNL, around the Department of Energy nuclear weapons complex, and to commercial areas of investigation to assist in scaling up to fieldable industrial uses.

Acknowledgements

The authors would like to thank Christine Robert-Coutant from LETI (France) for helping in the development of the A&PCT simulation software and in the analysis of the experimental A&PCT data. We appreciate the many helpful technical discussions with Bruce Hasegawa and Keenan

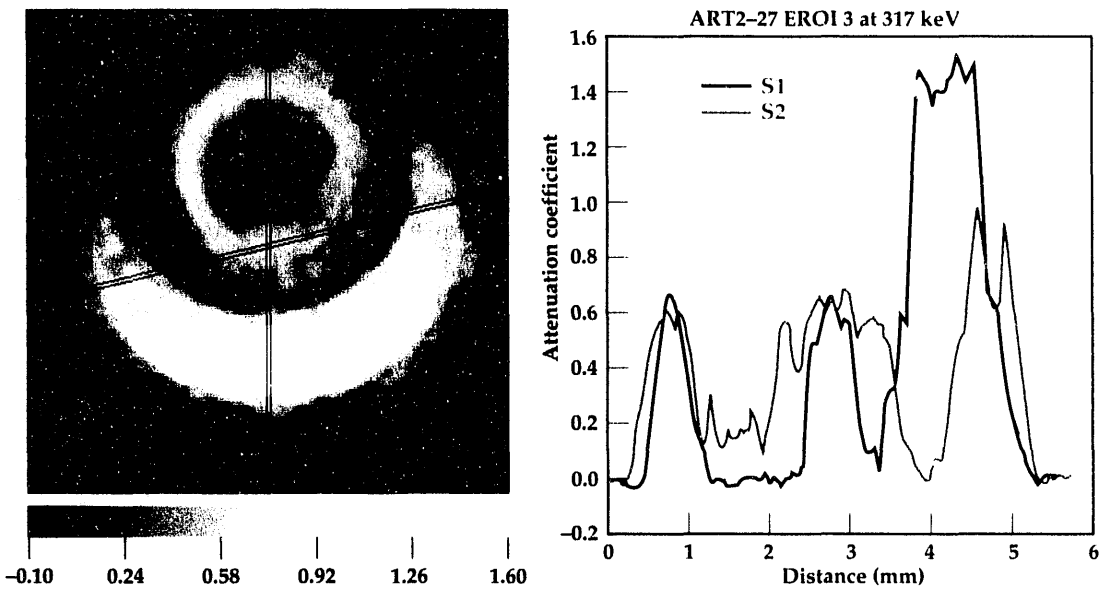


Figure 28. Representative reconstructed CT image of a plane through the approximate midplane of the ancient artifacts shown in Fig. 27. Two 1-D profiles are plotted to the right for further analysis of the CT image data.

Brown at UCSF on A&PCT simulation and experimental studies. Our discussions with Fred Cheu, General Motors, on the application of CT in the manufacturing environment have been very valuable in putting our research and development efforts in perspective. We thank Dick Albert at DigiRay for providing reverse-geometry cone-beam data for us to reconstruct and analyze. We are grateful to Rich Whipple for filling the shape charge with a mock plastic explosive. Thanks to Mark Cutler and Willie Johnson at HESCO for providing the film radiographs of the bridge cable. We also would like to acknowledge the valuable contributions to this project by Linwood Hester, Earl Updike, Derrill Rikard, Jerry Haskins, Dave Camp, Zach Koenig, and Clint Logan. Many thanks go to the Engineering Research Committee for their continued funding and support of this project.

1. H.E. Martz, D.J. Schneberk, G.P. Roberson, and S.G. Azevedo, "Computed Tomography," *Engineering Research, Development, and Technology*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-91, 7-6 (1992).

2. H.E. Martz, D.W. Perkins, S.G. Azevedo, D.J. Schneberk, M.E. Skeate, and G.P. Roberson, "Computerized Tomography," *Engineering Research and Development Thrust Area Report FY90*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-90, 7-14 (1991).

3. S. Azevedo, *Model-Based Computed Tomography for Nondestructive Evaluation*, Ph.D. Thesis, University of California Davis, Davis, California (1991).

4. H.E. Martz, S.G. Azevedo, G.P. Roberson, D.J. Schneberk, and M.E. Skeate, "Computerized Tomography," *Engineering Research and Development Thrust Area Report FY89*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-89, 7-15 (1990).

5. H.E. Martz, G.P. Roberson, and D.J. Schneberk, "Progess in Computerized Tomography Scanners," *Energy and Technology Review*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-52000-90-11,12 (November/December 1990).

6. H.E. Martz, D.J. Schneberk, and G.P. Roberson, "Applications of Computerized Tomography," *Energy and Technology Review*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-52000-90-11,12 (November/December 1990).

7. S.G. Azevedo, H.E. Martz, I.M. Brase, and K.F. Waltjen, "Computed Tomography," *Engineering Research and Development Thrust Area Report FY88*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-88, 7-1 (1989).

8. I.M. Brase, K.F. Waltjen, and H.E. Martz, "Computed Tomography," *Engineering Research and Development Thrust Area Report FY87*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-87, 9-14 (1988).

9. I.M. Brase, H.E. Martz, K.F. Waltjen, R.L. Hurd, and M.G. Wieting, *Quantitative Radiography*, Lawrence Livermore National Laboratory, Livermore, California, UCID 19323-86 (1986).

10. D.C. Camp, L-E. Wang, and H.E. Martz, "Preliminary Minimum Detectable Limit Measurements in 208-L Drums for Selected Actinide Isotopes in Mock-Waste Matrices," *Proc. Transuranic Waste Characterization Conf.*, Idaho State University (Pocatello, Idaho), (August 10–12, 1992); also Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-111345 (1992).

11. H.E. Martz, G.P. Roberson, D.J. Schneberk, and S.G. Azevedo, *IEEE Trans. Nucl. Sci.* 38, 623 (1991).

12. G.P. Roberson, H.E. Martz, D.J. Schneberk, and C.L. Logan, "Nuclear-Spectroscopy Computerized Tomography Scanners," *Proc. 1991 ASNT Spring Conf.* (Oakland, California), 107 (March 18–22, 1991).

13. H.E. Martz, G.P Roberson, C. Robert-Coutant, and D.C. Camp, "Experimental A&PCT Research and Development Efforts To Characterize Mixed Waste Forms," *Proc. Transuranic Waste Characterization Conf.*, Idaho State University (Pocatello, Idaho), (August 10–12, 1992); also Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-110826 (1992).

14. J.K. Brown, S.M. Reilly, B.H. Hasegawa, E.L. Gingold, T.F. Lang, and S.C. Liew, "Computer Simulation of an Emission-Transmission CT System," submitted to *Med. Phys.* (1992).

15. C. Robert-Coutant, H.E. Martz, and S.G. Azevedo, "Simulated A&PCT Data To Study the Mixed Waste Forms Characterization Problem," *Proc. Transuranic Waste Characterization Conf.*, Idaho State University (Pocatello, Idaho), (August 10–12, 1992); also Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-110827 (1992).

16. R.C. Placious, D. Polansky, H. Berger, C. Bueno, C.L. Vosberg, R.A. Betz, and D.J. Rogerson, *Mats. Eval.*, 1419 (November 1991).

17. R.C. Placious, D. Polansky, E.S. Gaynor, H. Berger, C. Bueno, R.A. Buchanan, C.L. Vosberg, and R.A. Betz, "An Improved Glass X Ray Scintillator," Final Report submitted to Naval Weapons Center, China Lake, California (1990).

18. A.H. Rodgers, Private communication, Synergistic Dector Designs, Mountain View, California (1992).

19. A.A. Harms and A. Zeilinger, *Phys. Med. Bio.* **22** (1), 70 (1977).

20. M. Barker, Private communication, Lockheed Missile and Space, Palo Alto, California (1992).

21. L.A. Feldkamp, L.C. Davis, and J.W. Kress, *JOSA A*, 612 (1984).

22. P. Grangeat, *Analyse d'un Système d'Imagerie 3D par Reconstruction à Partir de Radiographies X en Géométrie Conique*, Ph.D. Thesis, l'Ecole Nationale Superieure des Telecommunications, Grenoble, France (1987).

23. B.D. Smith, IEEE. *Trans. Med. Imaging*, **MI-4** (1), 14 (1985).

24. H. Kudo and T. Saito, *JOSA A* **7**, 2169 (1990).

25. Ph. Rizo, P. Grangeat, P. Sire, P. LeMasson, and P. Melennec, *JOSA A* **8** (10), 1639 (1991).

26. S. Azevedo, P. Grangeat, and Ph. Rizo, "Procede de Reconstruction d'Images Tridimensionelles d'une Region d'Interet d'un Objet, Comprenant la Combinaison de Mesures sur l'Ensemble de l'Objet a des Mesures sur une Region d'Interet de l'Objet, et Installation Appropriée," French Patent Application 92-11148, September 1992.

27. S. Azevedo, Ph. Rizo, and P. Grangeat, "Region-of-Interest Cone-beam Computed Tomography," submitted to JOSA A (1992).

28. R. Albert, Private communication, Digiray, San Ramon, California (1992).

29. S.G. Azevedo, H.E. Martz, and G.P. Roberson, "Computerized Tomography Reconstruction Technologies," *Energy and Technology Review*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-52000-90-11-12 (November/December 1990).

30. D.J. Schneberk *et al.*, *Limited Angle Radiography-Based Computed Tomography for In-Situ Inspections of Bridge Cables*, to be published (1993).

Laser Generation and Detection of Ultrasonic Energy

Graham H. Thomas
Engineering Sciences
Mechanical Engineering

We have developed a facility to generate and detect ultrasonic energy with lasers. Laser-generated ultrasonics is an attractive alternative to traditional ultrasonic nondestructive evaluation (NDE), because it allows remote, noncontacting, ultrasonic NDE. We are developing NDE applications for use on contamination-sensitive components and in hostile environments. Laser ultrasonics has several other advantages, such as broadband excitation, multimode acoustic energy generation, and adaptability to scanning complex shapes.

Introduction

Ultrasonic nondestructive evaluation (NDE) is a valuable technology for material characterization and defect classification.[1] Laser-based ultrasonics allows us to explore many new applications. For example, laser ultrasonics can be performed in hostile environments, such as in a furnace or a glove box. We are also pursuing laser ultrasonic techniques for providing feedback control for processes such as welding, composite curing, and solid-state bonding.[2]

Progress

We have acquired ultrasonic data on a variety of specimens to test the system's capabilities. We have demonstrated the feasibility of laser ultrasonics to perform feedback control for directing a welding operation. We performed an experiment to show the ability of laser ultrasonics to measure the distance to a weld seam. The application we are considering is for laser welding, where the welding laser must precisely track the joint. Our approach is to rigidly fix the laser acoustic system to the welding beam. The laser acoustic system can accurately measure the distance between the location of acoustic generation and the weld seam. If the welding laser wanders off the seam, the ultrasonic path length will change (see **Fig. 1**). The path length change will be fed to the welding laser alignment controller to adjust the laser location.

We demonstrated the ability of laser ultrasonics to measure the distance between the ultrasonic surface-wave-generation location and the weld seam. We generated ultrasonic surface waves in surrogate specimens. **Figure 2** displays an example of the ultrasonic signals, where the time between the two pulses is a function of the distance to the seam. If these pulses move relative to each

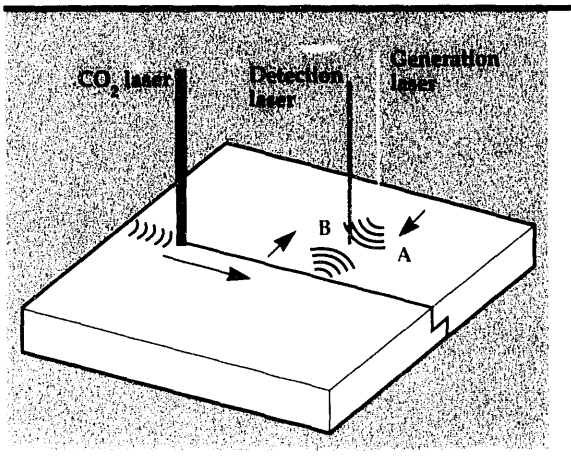


Figure 1. Diagram of laser ultrasonic feedback system for controlling welding CO_2 laser alignment. The ultrasonic surface wave is first sensed by the detection laser as it passes under the beam. The surface wave then reflects off seam ahead of the welder and is sensed again by the detection laser on the return.

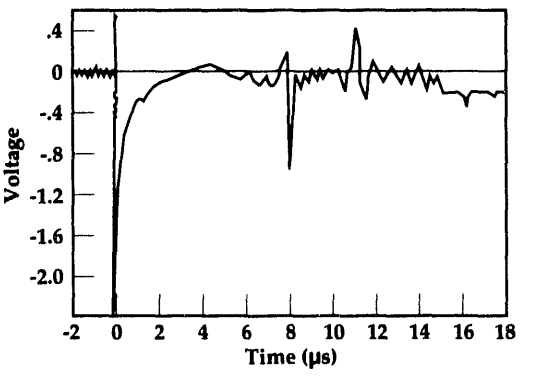


Figure 2. Sample results of laser ultrasonic signal from weld seam. Time between pulses provides feedback to control welding laser tracking.

other, the welding laser has moved off the seam. The timing between pulses should allow us to calculate position accuracy to .001 in.

Future Work

We are increasing our knowledge of laser generation and detection of acoustic energy, and improving our laser acoustic facility. We are investigating applications of laser acoustics to NDE problems simultaneously at Lawrence Livermore National Laboratory and within U.S. industry. Specifically, we will continue to explore applications for laser acoustics to control selected manufacturing processes, such as welder alignment, composite curing, and plutonium processing. Since each application of laser-generated ultrasonics entails a customized system, we need to have a thorough understanding of the fundamental capabilities and limitations of the technology to design the optimal inspection facility.

1. J. Krautkramer and H. Krautkramer, *Ulltrasonic Testing of Materials*, Springer-Verlag New York, Inc. (New York, New York), 1977.

2. N.M. Carlson and J.A. Johnson, "Laser Generation in a Weld Pool," *Review of Progress in Quantitative Nondestructive Evaluation* **7B**, Plenum Press (New York, New York), 1485 (1988).

Remote Sensing, Imaging, and Signal Engineering

Signal and image processing have always been important support for existing programs at Lawrence Livermore National Laboratory (LLNL), but now these technologies are becoming central to the formation of new programs. Exciting new applications such as high-resolution telescopes, radar remote sensing, and advanced medical imaging are allowing us to participate in the development of new programs.

The Remote Sensing, Imaging, and Signal Engineering (RISE) thrust area has been very active in working to define new directions.

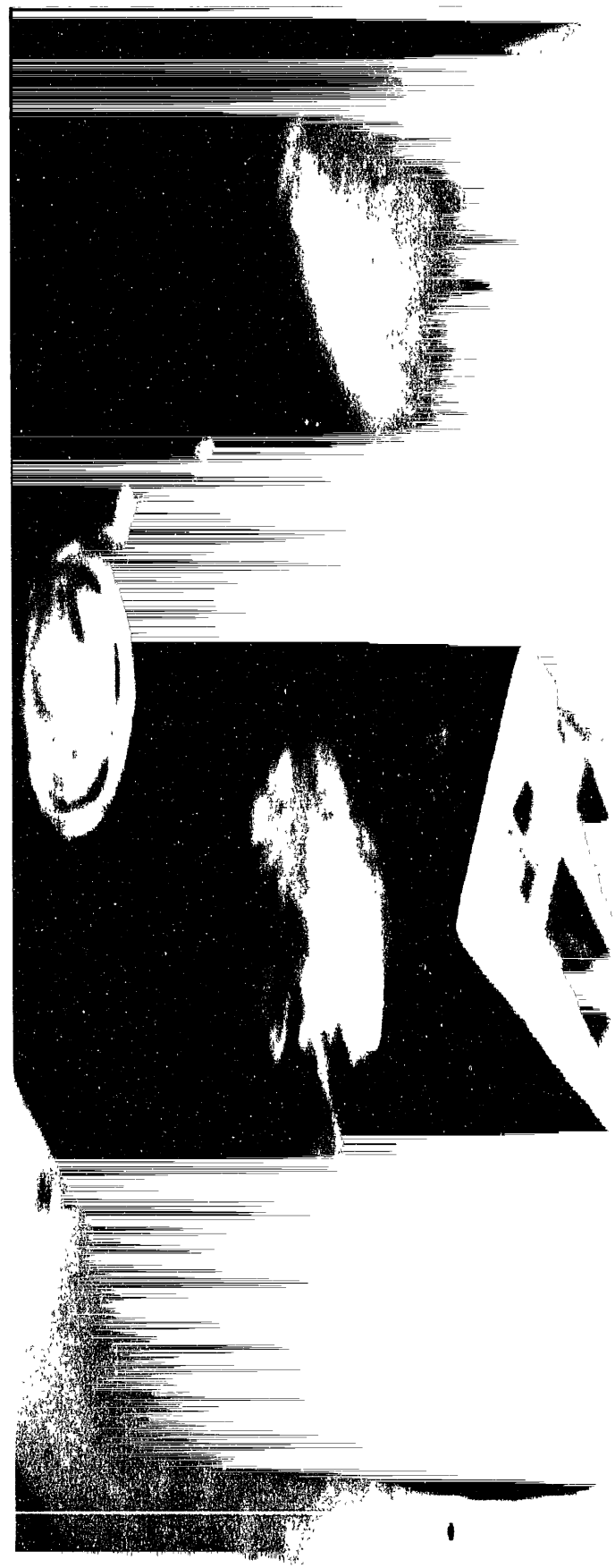
We also maintain and continue to build our technical base in signal and image processing in support of existing programs, through such applications as diagnostic image processing and seismic signal processing.

Over the past several years, RISE has developed a series of computer software systems for signal and image processing. These systems provide portability among the many computer systems used at LLNL and give us a platform for transferring the results of specific research and development projects to application areas. Our major signal- and image-processing systems, VIEW and VISION, are used by several major LLNL programs and have been distributed to many university, industry, and government sites.

Work in RISE involves a diverse set of sciences and technologies ranging from optical physics to microbiology to advanced computer architectures. Collaboration with other thrust areas, such as Nondestructive Evaluation and Computational Electronics and Electromagnetics, and with other LLNL organizations, such as the Physics Department and the Biomedical Sciences Division, is central to our continuing work in innovative imaging and signal-processing applications.

James M. Brase
Thrust Area Leader

Section 9



9. Remote Sensing, Imaging, and Signal Engineering

Overview

James M. Brase, Thrust Area Leader

Vision-Based Grasping for Autonomous Sorting of Unknown Objects

Shin-yee Lu, Robert K. Johnson, and Jose E. Hernandez **9-1**

Image-Restoration and Image-Recovery Algorithms

Dennis M. Goodman **9-7**

View: A Signal- and Image-Processing System

James M. Brase, Sean K. Lehman, Melvin G. Wieting, Joseph P. Phillips, and Hanna Szoke **9-11**

VISION: An Object-Oriented Environment for Computer Vision and Pattern Recognition

Jose E. Hernandez and Michael R. Buhl **9-15**

Biomedical Image Processing

Laura N. Mascio **9-21**

Multisensor Data Fusion Using Fuzzy Logic

Donald T. Gavel **9-23**

Adaptive Optics for Laser Guide Stars

James M. Brase, Kenneth Avicola, Donald T. Gavel, Kenneth E. Waltjen, and Horst D. Bissinger **9-27**

Vision-Based Grasping for Autonomous Sorting of Unknown Objects

Shin-yee Lu,
Robert K. Johnson, and
Jose E. Hernandez
Engineering Research Division
Electronics Engineering

The Department of Energy has a need for a method of treating existing nuclear waste. Hazardous waste stored in drums and boxes in old warehouses needs to be sorted and treated by the new standards of environmental regulations. At Lawrence Livermore National Laboratory, we are developing a vision-based grasping capability that can be used to pick and place unknown objects autonomously. Some preliminary results are described in this paper.

Introduction

In our experiment, we lay several objects on a table at arbitrary locations, simulating a conveyor belt. The objects are wrapped in plastic bags to simulate articles that are likely to be found in the waste containers. Two cameras are mounted above the table to create a stereo view of the work cell. The cameras are mounted approximately 2 m above the table, and have a field of view of approximately 2 m by 2 m. The images are captured and processed on a SUN SPARCstation-2, with image resolution of 510 × 480.

The stereo images are registered pixel-by-pixel using an efficient stereo-vision algorithm. A dense, three-dimensional (3-D) range map is generated by triangulating the registered pixels. Potential grasp locations for each object are generated by analyzing the shape of a two-dimensional (2-D) projection of the top view of the object. Locations around the handle or near the center of mass of the object are considered suitable for grasping, using a parallel gripper. The result of this analysis is used to generate information such as position, height, width, and orientation for executing the grasping task.

Progress

The experimental result shows high accuracy in the range estimation. We videotaped the experiment and studied the performance. The overall accuracy on the plane perpendicular to the cameras' lines of sight is within 2 mm, and along the line of sight is 5 mm. The total CPU time required for generating the grasping information is approximately 70 s for four objects. The computation time is proportional to the number of objects to be handled.

This experiment is an integration of camera calibration, stereo registration, shape analysis, and grasp planning. Algorithms used for camera calibration and image segmentation follow existing methods; however, our approach to stereo registration is different from most of the existing methods. It is efficient and highly parallelizable. Grasp planning at this point is a simple decision tree that matches the dynamic range of grippers to the size of the objects. Each of the different tasks is explained below in more detail, with an emphasis on the stereo registration algorithm.

General Approach

A set of transformation matrices for epipolar geometry correction are obtained through a camera calibration procedure. The images are segmented into regions, using a thresholding technique that separates the objects from the background. Since we assume that the objects are not touching each other, each region segment is assumed to correspond to an object in the scene. Correspondence of regions from the left image with those from the right image is then established, using features such as location and size of

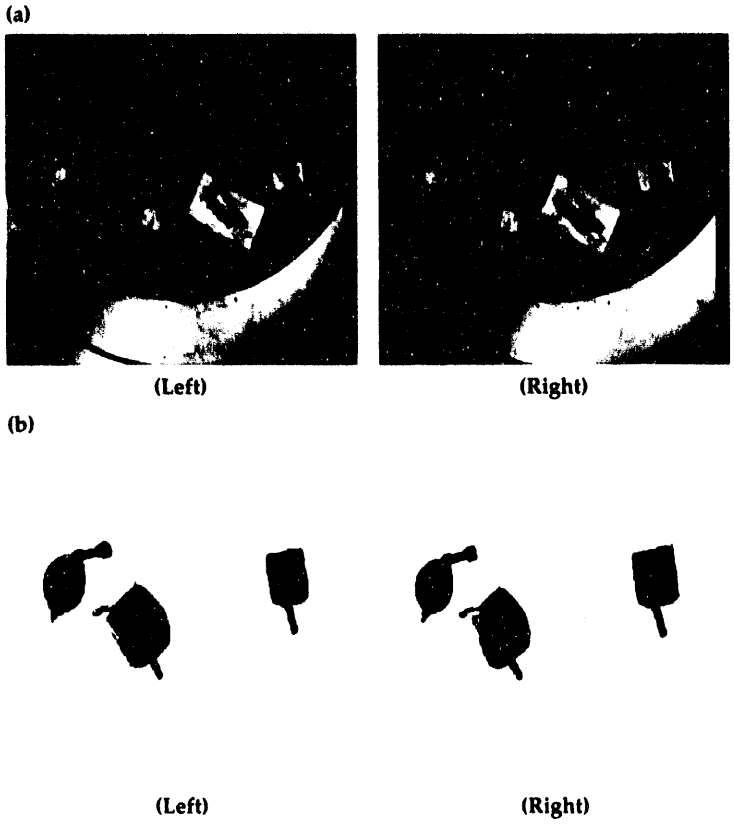


Figure 1. (a) Left and right images of a typical scene with four objects on a table; (b) segmented images with corresponded regions.

the regions. Corresponding regions are then transformed to the proper alignment according to epipolar geometry obtained from the camera calibration procedure. Pixel-to-pixel correspondence along epipolar lines of an object is accomplished using dynamic programming techniques. The registered pixels are triangulated using the camera transformation matrices. The result is dense range data. Assume that the line of the sight is along the z-axis. The range data thus obtained is a depth map on an irregular grid. We filter the range data and generate the final regular depth map for an object. A 2-D skeletonization technique is used on the resulting image of the depth map, for grasp analysis.

Camera Calibration

The accuracy of the 3-D range data obtained from stereo vision is directly dependent on the accurate modeling of the camera transformations from 3-D world coordinates to 2-D-image pixel coordinates. We are currently using a simple linear model[1] in which each of the two cameras is modeled by a 3-×-4 perspective transformation matrix, **B**:

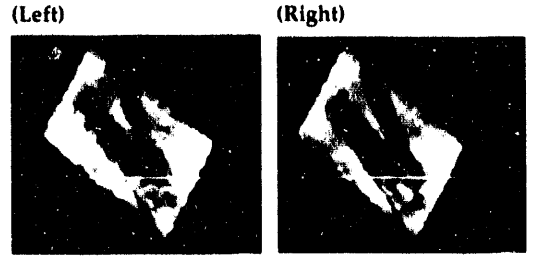


Figure 2. Corresponding epipolar lines (shown in white) for the epipolar-aligned box object.

$$\begin{aligned} c &= \mathbf{B}w \\ c &= [c_1, c_2, c_3]^t \\ w &= [x, y, z, 1]^t \\ p_x &= \frac{c_1}{c_3}, \quad p_y = \frac{c_2}{c_3}, \end{aligned} \tag{1}$$

where (x, y, z) is a world coordinate and (p_x, p_y) is the corresponding image pixel coordinate. The camera matrices, $\mathbf{B}_a$ and $\mathbf{B}_b$, are solved by the standard least-squares method, using calibration data consisting of world coordinates and their corresponding pixel coordinates. Each sample generates two equations in 12 unknowns for each camera. Therefore, a minimum of six samples is necessary to solve for the camera matrices. We currently use from 12 to 18 calibration points, by mounting fiducial markers in the robot work cell. The robot (Puma 560) is used to obtain the world coordinates, and centroiding is used to obtain the pixel coordinates. We are currently implementing an automated camera-calibration procedure, in which the robot will move a circular light source around the work cell to specified calibration points, for imaging and centroiding. This will allow for a much more efficient and thorough calibration of the vision system, in which hundreds of calibration points can be used. We expect that this will further reduce the observed positional errors. We are aware of more sophisticated calibration procedures,[2] in which camera distortion and image acquisition anomalies are modeled. However, for the current application, this additional complexity does not seem necessary.

Image Segmentation

A scene without objects is taken from each camera as a reference for background removal. Objects are segmented from the background by first subtracting the reference image from the im-

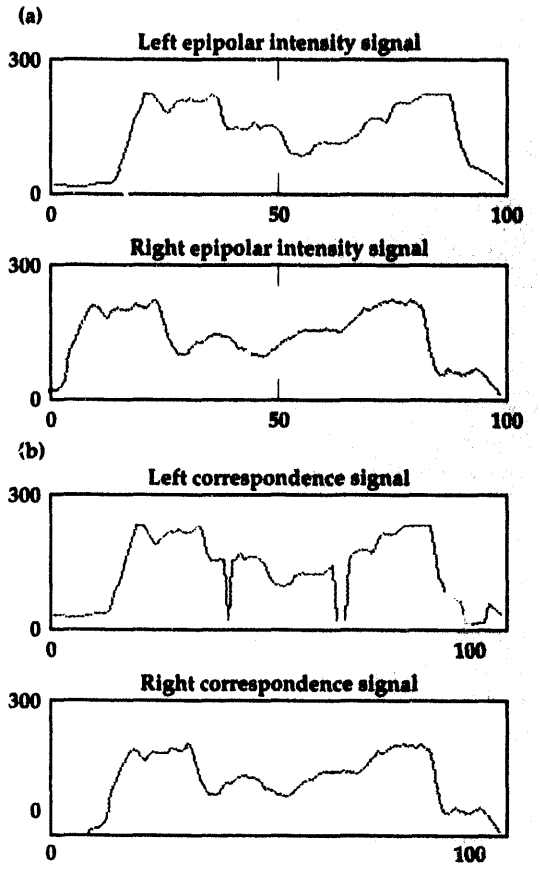


Figure 3. (a) Intensity profiles of epipolar lines from Fig. 2; (b) realigned correspondence profiles.

age with the objects, and then thresholding the difference image. Size constraints are used to eliminate small, noisy, background-region segments. Regions in one image are matched with regions in the other image, using simple heuristics based on the size and location of the segmented regions. Two stereo images of the scene and resulting corresponded regions are depicted in **Fig. 1**.

Epipolar Line Registration

Various dynamic programming techniques have previously been applied to matching edges in stereo images.[3,4] The result is a coarse disparity map for which a complex-surface-reconstruction algorithm is required to generate the final range map. In our approach, pixel-to-pixel registration is done by matching the intensity profiles of two corresponding epipolar lines, using a special dynamic programming technique called dynamic correlation.[5] Dynamic correlation is a method that optimally aligns data points, based on a similarity measure, preserving a defined ordering relation. When this technique is applied to stereo registration, we assume that pixels along epipolar lines have the same left-right relation. This relation can be represented mathematically by a linear ordering relation.[6] The ordering constraint is generally true for stereo registration, but can be violated. However, a strict linear ordering relation is obeyed by image pixels that pertain to the surface of an opaque object, that is, if pixels a and b are one matching pair and pixels a′ and b′ are another matching pair, then if a is to the left of a′ on one

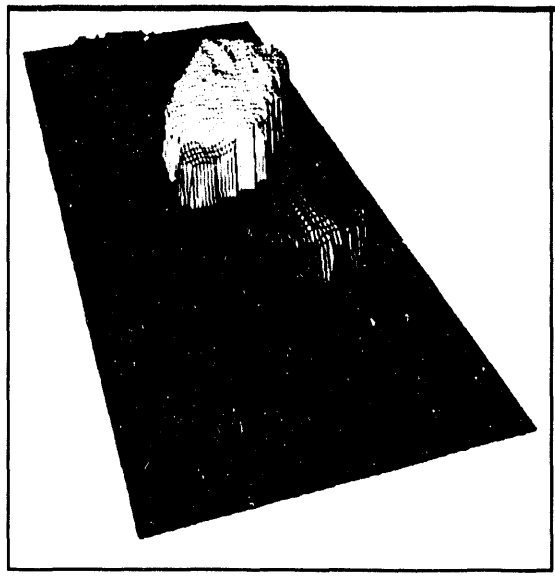


Figure 4. 3-D reconstruction of scene from Fig. 1.

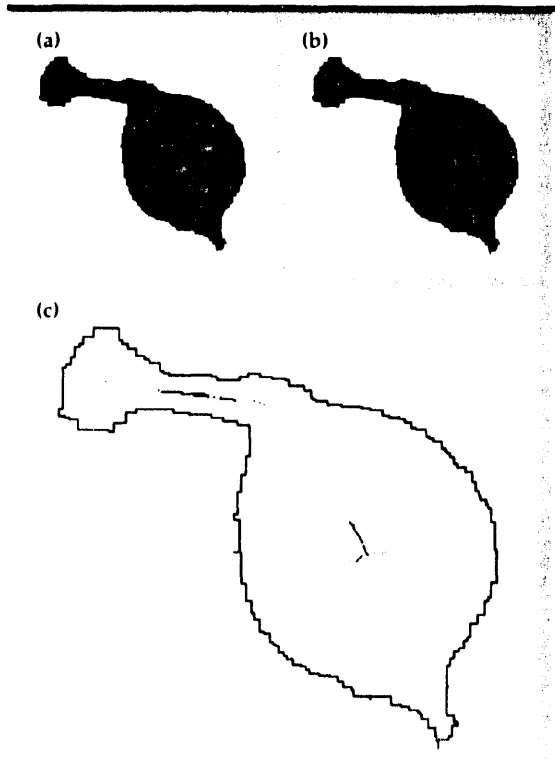


Figure 5. (a) 2-D gray-scale image of depth map; (b) 2-D Euclidean distance map; (c) 2-D symmetric skeleton.

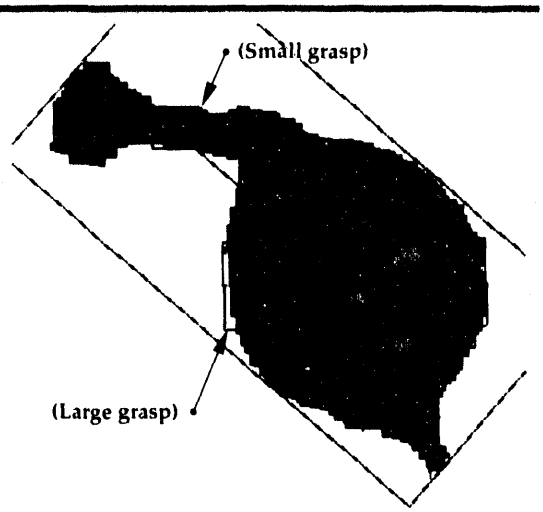


Figure 6. Two possible parallel-grasp orientations.

image, then it is necessary for b to be to the left of b' on the other image.

Algorithms for pixel-to-pixel registration have to be effective in handling (1) difference in scaling, (2) occlusion, and (3) variation in light reflectance. The deletion (or insertion) operation in dynamic correlation is used to handle both scaling and occlusion. The substitution operation represents a match, but allows variation in brightness. Let L_a and L_b be two corresponding epipolar lines, and let a_i, i = 1,2,...,n represent pixels on L_a, and b_j, j = 1,2,...,m represent pixels on L_b. The dynamic correlation algorithm calculates the cost of matching a_1, a_2,...,a_i and b_1, b_2,...,b_j, denoted C(i,j)):

$$C(i,j) = \min \begin{cases} C(i-1,j-1) + S(i,j) \\ C(i,j-1) + \alpha \\ C(i-1,j) + \alpha \end{cases}, \quad (2)$$

where

$$S(i,j) = 1 - \left\{ \frac{2R_{ab}(i,j)}{R_{aa}(i) + R_{bb}(j)} \right\}. \quad (3)$$

Here, $R_{ab}(i,j)$ is the windowed cross-correlation of L_a and L_b centered about a_i and b_j respectively, $R_{aa}(i)$ [$R_{bb}(j)$] is the windowed autocorrelation of $L_a(L_b)$ centered about $a_i(b_j)$, and α is a fixed cost of deleting an element of L_a or L_b. The normalized substitution cost, S(i,j), varies between 0 and 1, and the deletion cost, α, is fixed between 0 and 1. We have achieved good results with α between 0.2 and 0.5. The operation in **Eq. 2** defines a minimum cost matrix for i = 1,2,...,n and j = 1,2,...,m. C(0,j) and C(i,0) are given by $i\alpha$ and $j\alpha$, respectively. The minimum cost alignment can be traced back from C(n,m) to C(0,0) on the minimum cost matrix. If C(i,j) is derived from C(i–1,j–1), then pixel a_i matches b_j; if C(i,j) is derived from C(i,j–1), then pixel b_j on L_b does not have a match (a deletion); similarly, if C(i,j) is derived from C(i–1,j), then pixel a_i on L_a does not have a match.

Two corresponding epipolar lines are highlighted in **Fig. 2** for the epipolar-aligned box object from **Fig. 1a**. The result of pixel-to-pixel registration of these lines is illustrated in **Fig. 3**. The intensity profiles of the two epipolar lines are shown in **Fig. 3a**. These two intensity profiles are realigned after using the correlation algorithm. The realigned intensity profiles are shown in **Fig. 3b**. The matching pixels (substitutions) are aligned. When pixels on one intensity profile do not have a match (deletions), then a blank (shown as a zero value) is filled in on the opposite intensity profile. The occluded portions, i.e., the right-hand side of the box on the right image and the left-hand side of the box on the left image, are successfully deleted by the algorithm. The algorithm handles the slight difference in size (the box is at a larger skew angle to the left camera, therefore it is shown smaller on the left image than on the right image) by deleting four pixels from the right image at scattered locations. The 3-D reconstruction of all four objects from **Fig. 1** is shown in **Fig. 4**.

Grasp Analysis

Since the objects in this experiment are small, not too tall, and the bag handles are always placed parallel to the table, a simple shape analysis of the 2-D image of the depth map can be used to determine an 'optimum' grasp location with the gripper-oriented parallel table. First, we compute a Euclidean distance map from the 2-D projection of the depth map (see **Fig. 5**), using the fast raster scan algorithm.[7] A skeleton is then generated by locating generalized local maxima in the distance map.[8] Associated with each skeleton point is a vector pointing to the closest image point not contained in the object region. This orientation information is used to identify symmetric skeleton points (see **Fig. 5**) with respect to the object region boundary. Using information about the current available grippers, a grasp feature vector is then computed for each symmetric skeleton point, consisting of positional information, a grasp size, a parallel-boundary deviation measurement, and the distance from the object region centroid. Finally, a specified optimality criterion is used to

choose an optimal grasp. Our current criterion consists of first only considering grasps within the range of the current grippers and with parallel deviations less than a specified maximum. Of those, the grasp that minimizes a weighted average of parallel deviation and distance from centroid is chosen as the optimal grasp. In practice, we often divide the current grippers into two groups, (1) small grippers and (2) large grippers. We then find a grasp for each group. For the current objects, this often gives a large grasp about the center of mass and a small grasp about the bag handle, as shown in **Fig. 6**.

Future Work

Preliminary results of applying stereo vision and shape analysis to robot autonomous grasping of unknown objects show that stereo vision can provide fast and reliable range information. So far, we have used a relatively simple approach for grasp planning. We expect to deal with more complex objects as the waste sorting project progresses. We are also interested in the proper mating of object geometry and manipulator geometry, and plan to use special hardware such as transputers to speed up the process for real-time applications.

Acknowledgements

The authors would like to thank Maynard Holliday and the Advanced Processing Technology Program of Lawrence Livermore National Laboratory for their support and for the use of the robotic facilities in the Interactive Controls Laboratory.

1. D.H. Ballard and C.M. Brown, *Computer Vision*, Prentice-Hall, (Englewood Cliffs, New Jersey) 1982.
2. R.Y. Tsai, *IEEE J. Robotics and Automation* **RA-3**, 323 (1987).
3. Y. Ohta and T. Kanade, *IEEE Trans. Pattern Anal. and Mach. Intell.* **PAMI-7**, 139 (1985).
4. S.A. Lloyd, E.R. Haddow, and J.F. Boyce, *Computer Vision, Graphics, and Image Processing* **39**, 202 (1987).
5. S.Y. Lu, "A String-to-String Correlation Algorithm for Image Skeletonization," *Proc. 6th Int. Joint Conf. Pattern Recognition* (Munich, Germany), 178 (October 1982).
6. A.V. Aho, and J.D. Ullman, *The Theory of Parsing, Translation, and Compiling*, Prentice-Hall (Englewood Cliffs, New Jersey), 1972.
7. F. Leymarie and M.D. Levine, *CVGIP: Image Understanding* **55**, 84 (1992).
8. U. Montanari, *J. Assoc. Computing Machinery* **15**, 600 (1968).

Image-Restoration and Image-Recovery Algorithms

Dennis M. Goodman
Laser Engineering Division
Electronics Engineering

We have written computer codes for solving various image-restoration and image-recovery problems. These codes are based on a variant of the conjugate gradient algorithm that permits the imposition of constraints. Although the codes are essentially spatial-domain methods, most of the computation is done in the frequency domain. The result is that the flexibility of spatial-domain methods is preserved, but computation time is closer to that of conventional frequency-domain methods.

Introduction

A crucial tradeoff in applying image-processing algorithms to restoring a blurred image or recovering an image from data is accuracy vs time. Standard algorithms are non-iterative and operate in the frequency domain. Suppose the image is an array of N-×-N pixels. The number of floating point operations (FLOPS) required by these algorithms is typically of order $N^2 \log N$, the same order required for computing an N-×-N, two-dimensional, fast Fourier transform. Consequently, frequency-domain methods are reasonably fast; unfortunately, they are not flexible enough to impose non-negativity constraints, to handle nonlinear problems, or to deal with 'ringing' effects that occur when the blurred image is not zero at its boundaries.

The solution is to use spatial-domain methods, but the price paid in computer time is very high. Because direct inversion methods involve the storage and inversion of an N^2-×-N^2 matrix, these methods are impractical for all but very small images. Instead, iterative methods are used. A typical iterative method computes one convolution and one correlation per iteration. If these are implemented in the spatial domain, each requires on the order of N^4 FLOPS, so the total number of FLOPS required by an iterative method is on the order of MN^4 where M is the number of iterations. For small M, iterative methods are much faster than direct inversion methods, but are still much slower than frequency-domain methods. In fact, they are not practical for images larger than 128 × 128 pixels.

We have developed a new method that is basically a spatial-domain technique, but implements the iterations in the frequency domain. This reduces the cost per iteration to the order of $MN^2 \log N$ FLOPS. Our particular iterative technique is also new. It is based on the conjugate gradient algorithm and uses a 'bending' line search strategy, a special implementation of the active set strategy and the Hestenes–Stiefel formula.

Progress

In FY-91, we applied this algorithm to the standard, linear, least-squares image-restoration problem. We were able to demonstrate that the imposition of positivity constraints and the ability to properly handle boundary effects greatly enhanced image quality. This year, we performed Monte-Carlo experiments on small data problems, which demonstrated that the estimates obtained with our technique[1] were at least as good as those obtained with more conventional methods, such as constrained regularization and maximum entropy. As noted above, the conventional spatial-domain methods are too slow to apply to large data problems.

Many image-restoration and image-recovery problems are inherently nonlinear. For example, the algorithm we developed for the standard, linear, least-squares restoration problem is inappropriate when imaging at very low light levels. This is because the quantum nature of light must be accounted for, and the noise must be modeled as Poisson, rather than Gaussian. The result

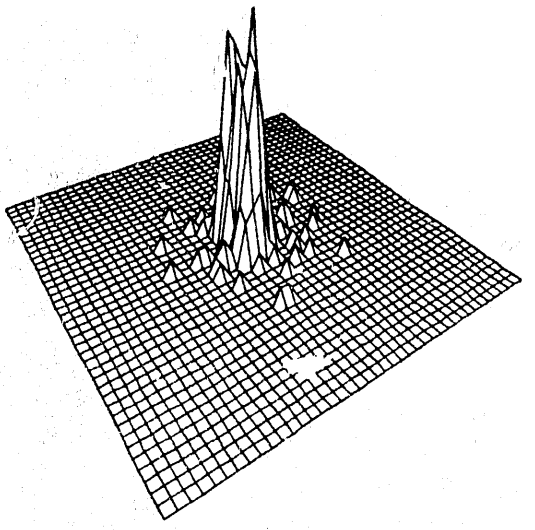


Figure 1. Blurred image with Poisson noise.

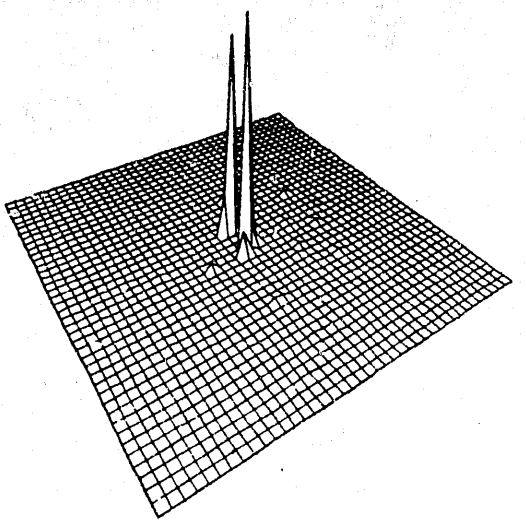


Figure 2. Estimate obtained using a least-squares criterion.

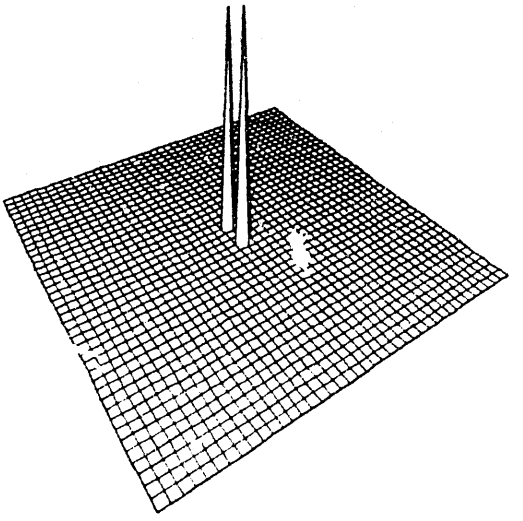


Figure 3. Estimate obtained by maximizing the Poisson likelihood function.

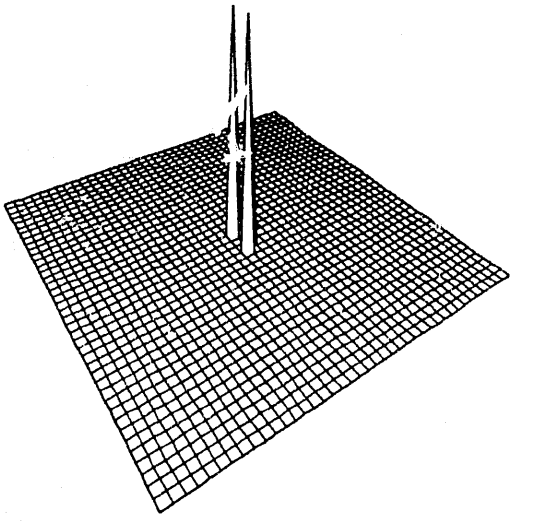


Figure 4. True noise-free image.

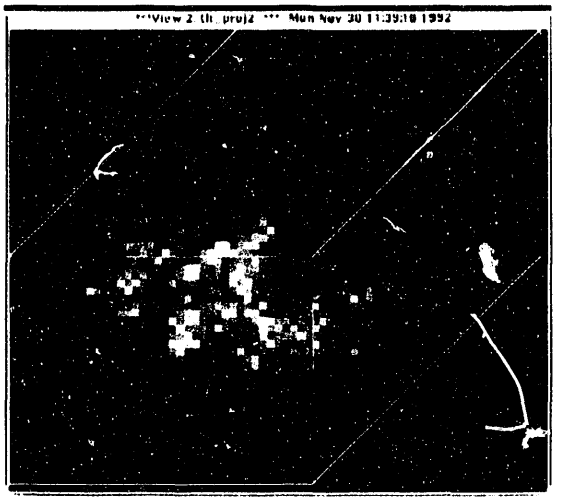


Figure 5. Reconstructed, three-dimensional, unit crystal cell for the protein thaumatin.

is a highly nonlinear likelihood function that must be maximized. In FY-92, we extended our algorithm to permit minimizing or maximizing general nonlinear functions.[2] **Figure 1** is a simulation of the result of imaging two closely spaced point sources through a circular aperture at a low light level. The result of deblurring using the least-squares criterion is shown in **Fig. 2**; the result of deblurring by maximizing the likelihood function for Poisson noise is shown in **Fig. 3**. The noise-free image is shown in **Fig. 4**. The estimate obtained by using the proper noise model is clearly superior.

We have also applied our algorithm to several other nonlinear imaging problems. These include speckle interferometry,[3] holography, and crystallography. A crystallographic example is shown in

Fig. 5. This image is a three-dimensional reconstruction of the protein thaumatin. The reconstruction is obtained using the Eden algorithm, which uses our algorithm as an inner iteration to repeatedly solve a non-negative least-squares problem consisting of 28,000 equations in 36,000 unknowns. The solution shows excellent clustering of the residual scatterers, since the reconstructed information occupied only 16% of the available grid positions.

Future Work

We plan to continue work on the crystallography problem in FY-93.

1. D.M. Goodman, "Deconvolution for Positive Signals," *Fundamentals of Discrete-Time Systems*, M. Jamshidi (Ed.), 1993.
2. D.M. Goodman, E.M. Johansson, and T.W. Lawrence, "On Applying the Conjugate Gradient Algorithm to Image Processing Problems," *Multivariate Analysis: Future Directions*, C.R. Rao (Ed.), North Holland, 1993.
3. D.M. Goodman, T.W. Lawrence, E.M. Johansson, and J.P. Fitch, "Bispectral Speckle Interferometry To Reconstruct Extended Objects from Turbulence Degraded Telescope Images," *Handbook of Statistics, Vol. 10: Signal Processing and its Applications*, N.K. Bose and C.R. Rao (Eds.), North Holland, 1993.

View: A Signal- and Image-Processing System

James M. Brase, Sean K. Lehman, and Melvin G. Wieting
Laser Engineering Division
Electronics Engineering

Joseph P. Phillips and Hanna Szoke
Scientific Software Division
Computation Directorate

View is an interactive signal- and image-processing environment for UNIX workstations with the X11 window system. View provides tools for image enhancement and general signal analysis. The system is used in programs at Lawrence Livermore National Laboratory for experimental data analysis and has been distributed to university, government, and industrial users.

In FY-92, we developed a capability to handle very large signal databases containing large numbers of signals or images with accompanying descriptive information; we continued to enhance the base View language, algorithms, and tools, and we demonstrated a prototype tool for distributed signal processing on a workstation network.

Introduction

A project to develop View,[1] an interactive signal- and image-processing environment for UNIX workstations with the X11 window system, was started at Lawrence Livermore National Laboratory (LLNL) in 1986. Early work focused on tools for image enhancement and analysis for nondestructive testing applications. View has been used extensively for image analysis for radiography and computed tomography as well as signal processing for ultrasonic imaging. Our development has continued with applications centered on radar imaging, remote sensing applications, and high-resolution astronomical imaging, including speckle interferometry and adaptive optics.

View provides most of the tools commonly required for signal and image analysis. Interactive capabilities include color map manipulation, line-out and region extraction, data value display, and image annotation. View's command interpreter provides a general purpose signal-manipulation language with looping and conditional constructs. The signal- and image-processing capabilities include spectral analysis, smoothing and sharpening filters, and adaptive noise-reduction techniques.

View continues to be used in the applications described above as well as in others at LLNL. It is the principal system used for reconstruction of x ray holograms and microscopy and as a diag-

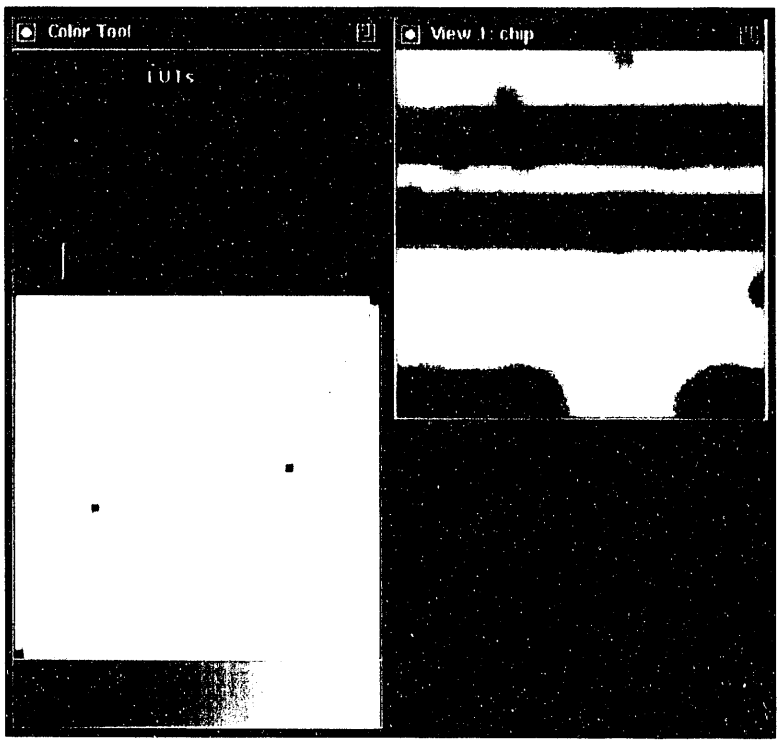


Figure 1. Image processing results from our new tool for interactive color map manipulation, allowing piecewise linear color maps.

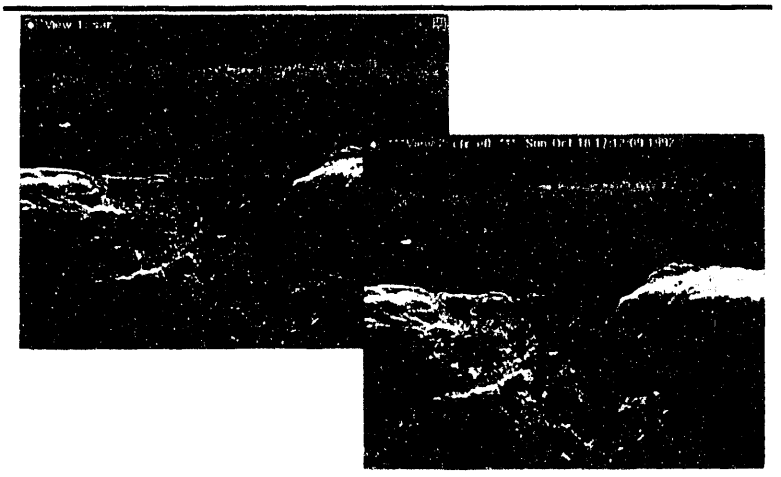


Figure 2. Image sharpening from a new two-channel technique added to View: (a) the original image; (b) recombined component images.

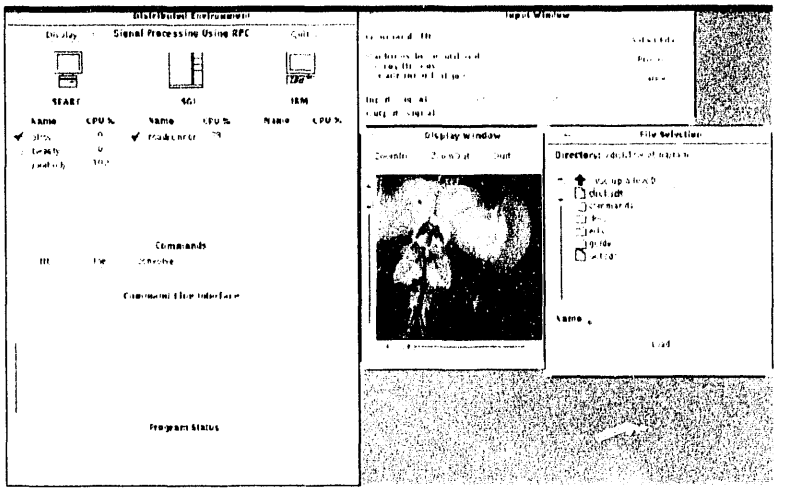


Figure 3. Graphical user interface from a prototype system for distributed signal processing. The user can assign parts of computations to different workstations on the network.

nostic tool for magnetic fusion experiments. Outside LLNL, View has been used for a wide variety of applications, including coding research, industrial tomography, and silicon wafer inspection. It has also served as a tool to teach basic signal and image processing at several universities.

Progress

Our development work on View in the past year has focused on three main areas: (1) the development of tools for handling large signal databases; (2) general enhancement of the available signal-processing algorithms and tools; and (3) the demonstration of a prototype distributed-signal-processing system.

The signal database is the heart of the View system. It comprises all the data known to the system along with associated information on its organization and characteristics. The original signal database had a flat structure; it was essentially a list of signals that could be accessed. This structure imposed the organization of analysis work on the user, either by using consistently meaningful names or simply by memory.

A better design is a hierarchy of signal groups or directories, similar to the way that most computer file systems are organized. With this structure, signals that are logically connected can be grouped together in the database structure. Signal names can be kept shorter, but the directory structure can still maintain a large amount of contextual information.

Our implementation of a hierarchical signal database is based on the UNIX file system. The database is a tree structure of signal directories that the user can move through and manipulate using a set of View commands. The commands have the same names as those in UNIX: **mkdir** to create a new subdirectory; **cd** to change the working directory; and **pwd** to print the current working directory. All other View operations can be executed normally and use the current working directory for signal storage, unless a specific path is given.

We have also developed a new capability to link the signal database to the existing Unix file system. If View is started with a "-u" option, a new signal database will be created that matches the UNIX directory structure starting at the current directory. All View data files in those directories will show up in the View internal-signal database, automatically. As the user moves through the signal database, the working UNIX directory changes to match it.

This linkage can greatly ease working with large amounts of data created by an external program and stored in a hierarchical file structure. We can immediately match the View database structure to the application.

Our second focus, continued development of new tools and algorithms for signal and image processing, has resulted in a new tool for interactively manipulating color maps (**Fig. 1**). This development allows piecewise linear mapping from data values to colors (shown in shades of gray), a valuable capability for images with multiple histogram peaks. The new color tool also allows the design of completely new color tables through interactive specification of red, green, and blue color maps.

Algorithm development is driven largely by the applications of View. Some of the new techniques added recently include adaptive smoothing filters for additive and multiplicative noise, improved spectrogram estimation for nonstationary signals, and basic capabilities for wavelet analysis. A new two-channel technique for nonlinear image sharpening is shown

in **Fig. 2**. On the left is the original image. We separate it into low- and high-frequency components and apply a nonlinear sharpening operation to each component.[2] The component images are then recombined to get the result on the right. This algorithm has proven useful in enhancing synthetic aperture radar imagery.

Our third development area in FY-92 was the demonstration of a prototype tool for distributed signal processing. The system allows operations to be distributed over a network of UNIX workstations. An interactive graphical user interface (**Fig. 3**) allows the user control over which machines run each command. Signal communication is through the network file system. These capabilities will form the base for future enhancements of View.

Future Work

Network-based distributed workstations will continue to be the foundation for our efforts in high-performance signal and image processing. View will be further developed to support both databases and computations that span the network. We are currently redesigning the signal-processing language and its interpreter to support these capabilities. We also plan continued development of user interface enhancements to conform to emerging standards in graphical user interfaces. Algorithm development will continue to be driven by ongoing applications.

1. J. Brase, V. Miller, M. Wieting, H. Szoke, and J. Phillips, *The View Signal and Image Processing System*, Lawrence Livermore National Laboratory, Livermore, California, UCID-21368 (1988).

2. S.K. Mitra, H. Li, I. Lin, and T. Yu, "A New Class of Nonlinear Filters for Image Enhancement," *Proc. Int. Conf. Acoustics, Speech, and Signal Processing* (Toronto, Canada), (1992).

VISION: An Object-Oriented Environment for Computer Vision and Pattern Recognition

Jose E. Hernandez and
Michael R. Buhl
Engineering Research Division
Electronics Engineering

VISION is a flexible and extensible object-oriented programming environment for prototyping solutions to problems requiring computer vision and pattern recognition techniques. VISION integrates signal/image processing, statistical pattern recognition, neural networks, low- and mid-level computer vision, and graphics into a cohesive framework useful for a wide variety of applications at Lawrence Livermore National Laboratory.

Introduction

During the past two years, we have been developing an object-oriented programming environment known as VISION, for computer vision and pattern recognition. VISION is a hybrid system consisting of C, Lisp/CLOS,[1,2,3] and some FORTRAN code. CLOS, the Common Lisp Object System, defines the new standard for object-oriented programming in the Common Lisp language.

The VISION system was developed with several goals in mind: (1) to provide a technology base at Lawrence Livermore National Laboratory (LLNL) in computer vision and pattern recognition; (2) to provide support to programs at LLNL requiring this technology; and (3) to provide a software package capable of being extended and customized directly by the end users.

During FY-91, most of the object-oriented framework was developed, including basic classes of data structures for signal/image processing, mid-level two-dimensional (2-D) computer vision, and unsupervised and supervised learning algorithms including several neural networks.[4,5] Some of the capabilities in VISION were applied to several projects sponsored by LLNL's Earth Sciences Department. Also, VISION was used as a development environment for the temperature-evaluated mine position survey (TEMPS) project for locating buried mines.[6] This preliminary work resulted in a 1.5-million-dollar project currently funded by the Defense Advanced Research Projects Agency, and a possible licensing agreement with a private company.

During FY-92, VISION was used as the development environment for a research project in stereo vision and grasp planning for robotics. This effort resulted in a demonstration system currently being used at LLNL's Interactive Controls Laboratory managed by the Advanced Processing Technology Program. Some of the capabilities in VISION have also been integrated into LLNL's Seismic Expert System,[7] sponsored by the Treaty Verification Program.

In FY-93, VISION will be used to prototype pattern recognition algorithms for LLNL's INSENS project, Broken Heart Valve project, and wake detection project, and for developing more advanced capabilities in computer vision for robotics.

Overview of VISION

VISION consists of two major parts: the programming environment, and the computer vision and pattern recognition capabilities. The programming environment is primarily provided by the Common Lisp environment itself. Some of its features are listed below.

(1) Interactive programming: eliminates the need to write a command-driven user interface and encourages incremental development;

(2) Run-time linking: C, FORTRAN, and compiled Lisp code can be loaded and linked dynamically at run-time;

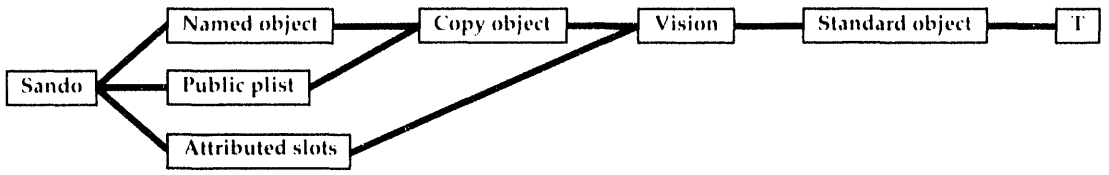


Figure 1. An interactive class browser, useful for understanding the VISION class system.

(3) Automatic memory management: Lisp handles the allocation and de-allocation of memory, hence code can be developed faster;

(4) Loosely typed language: since there is no need to declare data types, algorithms can be prototyped faster;

(5) Functional programming: we have the ability to dynamically define functions to be passed as arguments to other functions, which is essential to our framework for pattern recognition;

(6) Object-oriented programming: Lisp supports the object-oriented programming paradigm, which is essential for our extensible framework, through CLOS;

(7) Emacs interface: expressions, regions, and buffers within Emacs can be submitted to the Lisp interpreter directly, which improves productivity;

(8) Artificial intelligence (AI) software: Lisp software is available in the public domain for supporting many of the AI paradigms for high-level reasoning; and

(9) Class browser: an interactive class browser based on GARNET is available for browsing the VISION class system (see **Fig. 1**). (GARNET is a Lisp-based graphical user interface environment developed at Carnegie Mellon University.[8])

The computer vision and pattern recognition capabilities in VISION consist of an integrated set of data structures and algorithms within an object-oriented framework for (1) representing, processing, and segmenting one-dimensional (1-D), 2-D, or three-dimensional (3-D) data; (2) calculating and evaluating features for statistical pattern recognition; and (3) several paradigms for object recognition and classification including neural networks and an Assumption Truth Maintenance System.[9] In summary, VISION is an extension to the Common Lisp environment to make it more useful for signal/image processing, pattern recognition, and computer vision.

Progress

In FY-92, there were a substantial number of improvements and developments in the areas of computer vision and pattern recognition made to the FY-91 VISION release.

Collection Objects

The class Collection-Object is one of the most fundamental building blocks in VISION. This class unifies many of the data structures in VISION for storing collections of other data structures. Perhaps the most important aspect of this class is that it provides many methods for implementing generic higher-order functions (GHOF).[10] These are generic functions that accept other functions as arguments to

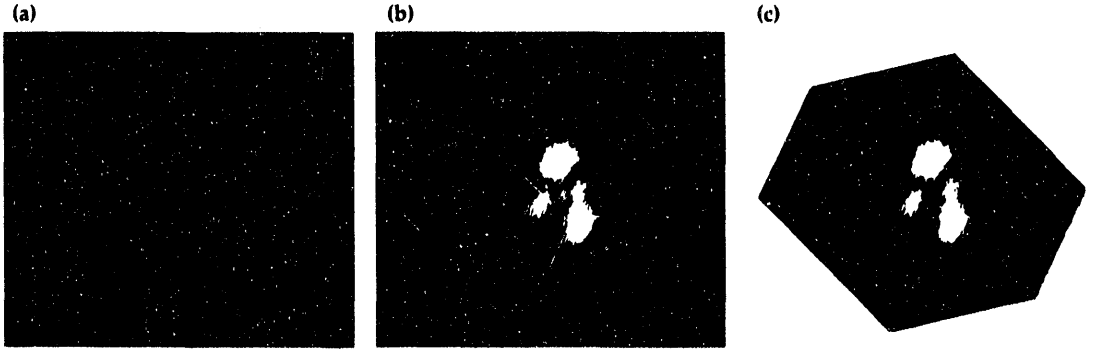


Figure 2. (a) 3-D data representing a cell, (b) thresholded data. Each voxel in the volume is assigned to one of three intensity bins represented by the three different gray levels. (c) Regions identified within the voxels in the third intensity bin. Three chromosomes were found.

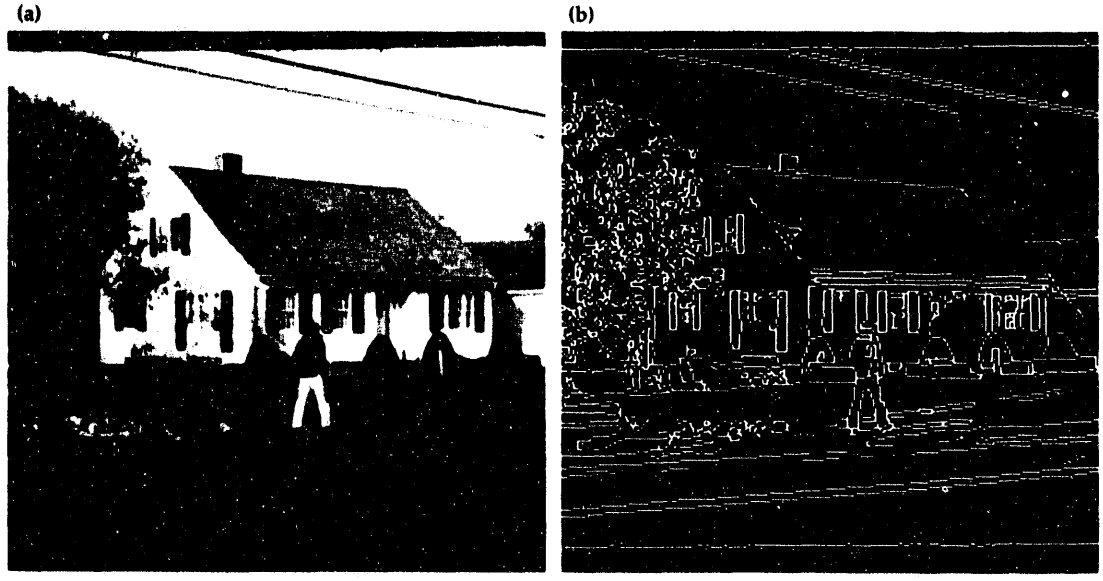


Figure 3. Example of the use of an automatic thresholding algorithm on the gradient of an image. Pictured are (a) a house scene and (b) edges of the image.

be applied to the individual objects in the collection. Generic functions[2,3] are functions for which methods can be defined to provide the appropriate functionality for different classes of objects. GHOF provide a powerful mechanism for solving problems without the explicit use of recursion or iteration. Furthermore, they hide the internal representation of the collection object, since the iteration process is hidden. For example, consider the GHOF **gcount-if**, which counts the number of objects in the collection that satisfy a predicate (test). In the context of computer vision, we could use this function to count all the round objects in a segmented image,

```
(gcount-if #'roundp segments)
=> 52
```

In general, algorithms can be prototyped faster, since we only need to develop primitives that deal with the individual objects in the collection. We can then use lambda expressions to combine these primitives and form more complex expressions that can be applied to the individual objects in the collection. Lambda expressions[1] are anonymous functions typically defined to be passed as arguments to other functions. For example, we could use a lambda expression to count all the signals from a collection of time series that have a positive mean,

```
(gcount-if #'(lambda (x)
               (plusp (mean x))) signals)
=> 11
```

Some examples of subclasses of the class Collection-Object are: Signals, for storing a collection of 1-D waveforms; Monochrome-Images, for storing a collection of 2-D images; and Segmented-Image, for storing a collection of region segments in an image.

Volume Segmentation

An object-oriented framework for 3-D data segmentation has been developed. The new capabilities are very similar to the computer-vision capabilities that were developed last year for 2-D data. In fact, the original, 2-D, computer-vision framework was redone so that it could be extended to any N-dimensional space. Due to our limited resources, there are no plans at this point to develop 3-D graphics capabilities in VISION. However, an interface was developed to write the different classes of 3-D data objects in VISION to disk in SUNVISION format for data visualization. SUNVISION is a 3-D, interactive visualization program available for the SUN workstations. LLNL currently has a site license for SUNVISION. An example of a segmented volume showing the nucleus and chromosomes of a cell is shown in **Fig. 2**.

In summary, the new 3-D capabilities include:

(1) multi-level thresholding of 3-D data,
(2) representation and processing capabilities for any arbitrary set of voxels in the volume (connected or non-connected),
(3) 3-D grouping algorithm for identifying 3-D 'regions' in the volume,
(4) basic-shape-analysis capabilities, and
(5) interface to SUNVISION for 3-D data visualization.

Automatic Thresholding

Several algorithms were developed for automatic thresholding of data.[11] Some of these algorithms are useful for separating background from non-back-

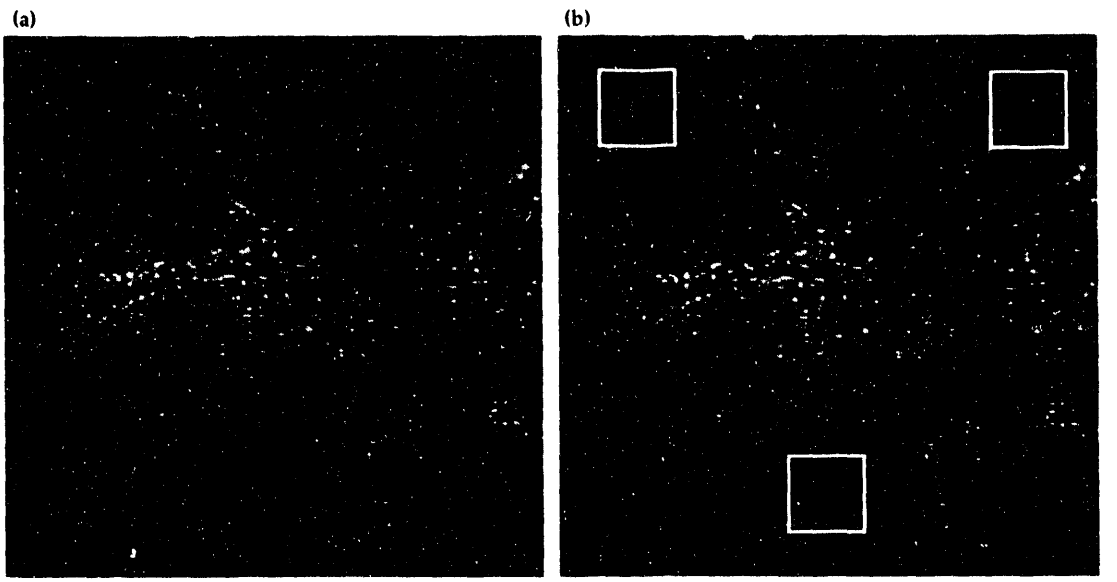


Figure 4. (a) An image from an infrared sensor showing several objects, (b) three sections in the image, classified as buried mines using a neural network.

ground pixels. They are also very useful for automatic thresholding of gradient images for edge detection (see **Fig. 3**).

A multi-thresholding algorithm was also developed, based on a K-means algorithm that clusters the data values directly from a histogram and therefore it is very fast. The algorithm also features the ability to find the best number of threshold values based on the ratio of the scatter-matrices.[12] This algorithm is currently being evaluated for segmenting x rays of suitcases taken at airports, for detecting explosives.

Feature-based Object Recognition

Many algorithms were developed for extracting features for object recognition. Some of them are listed below.

(1) Histogram features: also known as first-order features, used to extract features from the probability density function of the data;

(2) Central moments: can be used to extract shape information and are invariant to translation;

(3) Hu moments: similar to the central moments, but they are also invariant to rotation; and

(4) Texture features: also known as second-order features, used to extract texture features from an image.

Many other algorithms are also available for extracting information that could be used as features for object recognition. These features can be used as input to a rule-based system or to a neural network for object recognition.

Supervised Learning

An object-oriented framework for supervised learning using statistical pattern recognition techniques and neural networks was formalized this year. The framework consists of two classes for manipulating databases for supervised learning; feature-selection algorithms[13] for evaluating and selecting useful features for solving classification problems; and two new classifiers, a nearest-neighbor classifier and a probabilistic neural network.[14]

Most of the supervised learning algorithms in VISION operate on a few data structures referred to as training and patterns tables. A 'training table' is a data structure that associates a label (typically a symbol) meant to represent the name of a category of patterns, with a collection of objects (an instance of a Collection-Object class, in most cases). A 'patterns table' is a special kind of training table where the objects in the collection are constrained to be feature vectors, all of the same size. Several GHOF are provided for performing transformations on these data structures. For example, in a typical application, we might start by creating a training table that keeps track of file names with the original measurements associated with each category. We can use the GHOF **maptable** to create a patterns table with the actual feature vectors to be used by the learning algorithms. Consider the following example,

```
(setq files (make-training-table
              :male '("m1" "m2" ...)
              :female '("f1" "f2" ...)))
=> <training-table>
(setq patterns (maptable #'calc-features files
              :class 'patterns-table))
=> <patterns-table>
```

In this example, we assume that the function **calc-features** has already been defined, such that given a file name, it reads the file, calculates the appropriate features, and returns a feature vector. The function **maptable** takes care of applying this function to every file name in the original training table and producing a new table with the actual feature vectors. The lambda expressions are very useful for prototyping functions like **calc-features** in order to try different kinds of features for the learning algorithms. Also, there is no need to store all of the initial raw data read from disk (which could be a serious problem with large databases), since only the final results (the feature vectors) are kept in memory.

Once an initial set of features has been calculated, it is typically evaluated using one of the several feature-selection algorithms in VISION, in order to find the best set of features that separate the N-dimensional feature space. These features are then used to train one of the several classifiers in VISION, including a back-propagation neural network.[15] These techniques have been successfully used for detecting and locating buried mines using dual-band, infrared sensors (see **Fig. 4**).[16,17]

Future Work

The main goal for FY-93 is to complete the documentation for VISION[18] to make its capabilities more accessible to the LLNL community. We are also seeking technology transfer opportunities that will allow us to further expand our technology base in computer vision and pattern recognition. One organization from Pacific Gas and Electric Company is currently very interested in using VISION as its internal prototyping environment for applications in pattern recognition. We also expect current projects using VISION to contribute new algorithms and capabilities.

Acknowledgements

The authors want to acknowledge the contributions made to VISION during FY-92 by Robert K. Johnson, Sailes Sengupta, Robert J. Sherwood, Paul C. Schaich, and William J. Maurer.

1. G.L. Steele, Jr., *Common Lisp: The Language*, 2nd ed., Digital Press (Burlington, Massachusetts), 1990.

2. J.A. Lawless and N.M Miller, *Understanding CLOS, The Common Lisp Object System*, Digital Press (Burlington, Massachusetts), 1991.

3. S.A. Keene, *Object-Oriented Programming in Common Lisp*, Addison-Wesley (Reading, Massachusetts), 1989.

4. J.E. Hernandez, G.A. Clark, and S. Lu, "Computer Vision," *Engineering Research, Development, and Technology*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-91, 8-5 (1992).

5. J.E. Hernandez, S. Lu, R.J. Sherwood, G.A. Clark, and B.S. Lawver, *A Signal and Image Processing Object-Based System Using CLOS*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-108409 (1991).

6. N.K. Del Grande, G.A. Clark, P.F. Durbin, D.J. Fields, J.E. Hernandez, and R.J. Sherwood, "Buried Object Remote Detection Technology for Law Enforcement," *Proc. SPIE Orlando '91 Symposium* (Orlando, Florida), (April 1–5, 1991).

7. W.J. Maurer, F.U. Dowla, and S.P. Jarpe, *Seismic Event Interpretation Using Self-Organizing Neural Networks*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-108630 (1992).

8. B. Myers, D. Giuse, R. Dannenberg, B. Zaden, D. Kosbie, E. Pervin, A. Mickish, and P. Marchal, "GARNET," *IEEE Computer Magazine* **11**, 71 (1990).

9. R.J. Johnson, T.W. Canales, D.L. Lager, C.L. Mason, and R.M. Searfus, "Interpreting Signals with an Assumption-Based Truth Maintenance System," *Proc. SPIE—The International Society for Optical Engineering* **786**, 332 (May 1987).

10. J.E. Hernandez, *Higher-Order Generic Functions for CLOS*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-109776 (1992).

11. R. Haralick and L. Shapiro, *Computer and Robot Vision, Volume 1*, Addison-Wesley (Reading, Massachusetts), 1992.

12. G. Coleman and H. Andrews, "Image Segmentation by Clustering," *Proc. IEEE* **67** (5), (May 1979).

13. T. Young and K. Fu, *Handbook of Pattern Recognition and Image Processing*, Academic Press Inc. (San Diego, California), 1986.

14. D.F. Specht, "Probabilistic Neural Networks," *Neural Networks* **3**, 109 (1990).

15. E.M Johansson, F.U. Dowla, and D.M. Goodman, *Backpropagation Learning for Multi-Layered Feed-Forward Neural Networks Using the Conjugate Gradient Method*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-104850 (1991).

16. M.R. Buhl and J.E. Hernandez, *Dual-Band, Infrared Buried Mine Detection Using A Statistical Pattern Recognition Approach*, Lawrence Livermore National Laboratory, Livermore, California, in preparation.

17. J.E. Hernandez, M.R. Buhl, and S. Sengupta, *Detecting and Locating Buried Mines from Dual-Band IR Data: A Pattern Recognition Approach*, Lawrence Livermore National Laboratory, Livermore, California, in preparation.

18. J.E. Hernandez, *Using Vision*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-MA-112337-DRAFT (1992).

Biomedical Image Processing

Laura N. Masclo
Defense Sciences Engineering Division
Electronics Engineering

We have developed a bio-imaging application for a genetics study and have made advances in projects related to automated fluorescence, microscopy, and mammography.

Introduction

In FY-92, we used funds from a small grant to make contributions to several biomedical research projects, including (1) colony filter analysis for genetic studies; (2) the human genome project; and (3) the detection of microcalcifications in digitized mammography.

Progress

Colony Filter Analysis for Genetic Studies

We have made progress in the automation of quantitative colony filter analysis (CFA), an important and versatile tool used by biologists for a variety of research goals. One application is to pinpoint interesting regions in human DNA so that more highly detailed analyses, such as sequencing, can be applied directly to these regions. Another goal is to very precisely determine the expression patterns of a gene. Using these patterns for comparison can provide a measurement of the genetic differences between distinct groups, such as male vs female, diseased persons vs non-diseased persons, or young persons vs old persons. One of the many other designs for a CFA experiment can yield the location of a certain DNA sequence, or gene, along a chromosome.

Because of its versatility, the CFA is a powerful tool in today's genetics studies. Also because of its versatility, however, the analysis is highly complex, and automating this analysis is a technical challenge. One format for the data is an array of 18,000 radioactive data spots generated from a robotically prepared 20-cm-×-20-cm filter paper. Each of the 18,000 spots contains a signal of importance, although many signals may not be visible when imaged, and some are even difficult to detect computationally. When the filter paper has been imaged and digitized, it can form a data set up to 23 Mb in size.

To automate the quantitation and location of each of the 18,000 signals, we first developed an image-processing algorithm that locates the spots that are detectable, and then predicts the location of those that are not. This algorithm and its platform (SCIL-Image) are capable of handling 23 Mb of original data plus 4 to 5 times that for intermediate results.

Morphological image processing is the prominent methodology used in the automated CFA tool. The maximum (gray-scale dilation) and minimum (gray-scale erosion) operators are used in various combinations to provide background information, as well as texture or frequency information, for detecting the DNA colonies. These methods are documented thoroughly,[1] and outlined briefly in **Fig. 1**. Once the algorithm has detected all spots, the image may be rotated so the colony array is aligned with the image. Then, long, thin, maximum filters are used to 'smear' the dots, first horizontally, and then vertically. The intersection of the smearing lines predicts the location of undetected spots. The grid is then rotated to fit over the original data. The rotated data cannot be used, because we are interested in quantifying the colonies. The affine transform that performs the rotation uses interpolation methods to assign each pixel a new value in the rotated image.

Next, we use the smearing lines to form a dynamic grid (non-uniform) over the data set, so that each grid square contains only one DNA colony. This grid provides the framework by which each DNA colony can be assigned a coordinate position. That is, while it is trivial to know the pixel coordinates of a spot, it is much more useful and difficult to know its grid position. The assignment of coordinates to the grid squares is not as trivial as

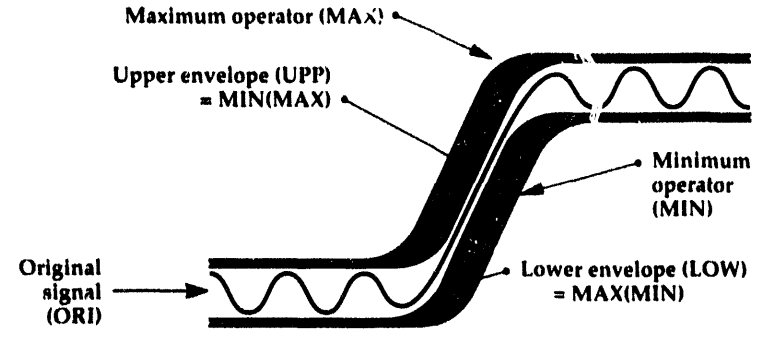
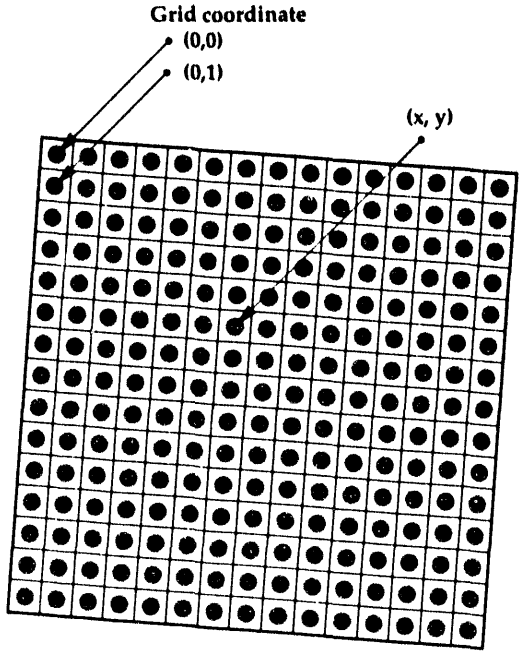


Figure 1. Maximum and minimum filters combined for powerful image processing. Shown is a one-dimensional signal and the effect of various max–min operator combinations. For the colony filter application, the lower envelope is used to estimate background values, so they can be subtracted before quantitation of the colonies. A low-detail image is computed, so that the image-intensity characteristics can be simplified before the high-detail image is computed. The high-detail image yields texture information, so the 'spots' formed by the colonies can be detected.

Figure 2. An example of a small filter with computer-generated grid-lines separating the objects. Each grid square must be assigned a grid coordinate, so that the results from each grid square can be related to the others.



the obvious solution, because the grid is not necessarily square with the image, and in addition, the grid is not uniform (see **Fig. 2**).

Once each spot is localized and assigned grid coordinates, the total integrated intensity of the signal in each grid square is summed. This intensity measurement must account for (subtract) global and local background intensities, which would act as a 'dc offset' to the signal. After subtracting the background, any intensity from a surrounding signal must also be removed. Intensity values for each grid square can be reported, or queried individually by the investigator.

Upon further interaction with biomedical scientists, a final application can be customized for use with their data-acquisition routine.

Automated Fluorescence Microscopy for DNA Mapping (Human Genome Project)

We have undertaken the reconstruction of an automated fluorescence microscope. The equipment needed to construct this instrument, i.e., microscope, cooled CCD camera, automated filter wheel, automated stage, and controlling computer software, has been assembled, and the controlling software is nearly complete. We will use this instrument and a software application previously developed to demonstrate the potential to automate DNA probe mapping in metaphase chromosomes as a part of our human genome project.

Detecting Microcalcifications in Digitized Mammograms

We have made significant contributions to a research effort for early detection of cancer-warning signals in mammograms. This project is one that makes use of existing, but under-used equipment from nondestructive evaluation of weapons parts, i.e., a high-resolution digitizer for radiographs. By digitizing radiographs of breast images, computational methods, such as the morphological methods described in **Fig. 1**, can be used to analyze mammograms and to highlight areas that may require a radiologist's attention. The aim of this project is to help reduce the number of missed diagnoses of cancer in cases where warning signs exist on a mammogram.

1. P.W. Verbeek, H.A. Vrooman, and L.J. Van Vliet, "Low-Level Image Processing by Max-Min Filters," *Signal Process.* **15**, 249 (1988).

Multisensor Data Fusion Using Fuzzy Logic

Donald T. Gavel
Laser Engineering Division
Electronics Engineering

We have developed an expert system based on fuzzy logic theory to fuse the data from multiple sensors and make classification decisions for objects in a waste reprocessing stream. Fuzzy set theory has found successful application in a number of decision and control applications in recent years. We have found that a fuzzy logic system is rather easy to design and train, and that with proper training, classification accuracy is quite high. We performed several tests sorting radioactive test samples using a gamma spectrometer to compare fuzzy logic to more conventional schemes.

Introduction

The Department of Energy (DOE) has an urgent need for the development of waste processing and cleanup technologies. Over the past few years, the Advanced Process Technology Program at Lawrence Livermore National Laboratory has been developing robotics and automation technology to support cleanup and reclamation efforts. In our Interactive Controls Laboratory, we have developed a sensor-based robot system for material sorting tasks.

Robotic sorting of materials in a waste stream has been largely motivated by the DOE cleanup needs. A large fraction of the buried radioactive waste must be dug up and repackaged because contaminants are leaching into underground water tables. Hazardous waste stored in barrels at local sites must be resorted, according to federal guidelines, into categories of high-level, low-level, transuranic, and mixed waste, and disposed of accordingly. Certain materials, such as lead and stainless steel, can be reclaimed after being cleaned of radioactive contamination. Low-level and mixed radioactive waste must be sorted into categories, such as burnable or vitrifiable, for later volume reduction and storage. Using robots instead of radiation-suited workers reduces the risk to humans, and also improves the reliability and speed of operation.

Weapons dismantlement is now another important issue. Technologies need to be developed to handle the waste materials derived from dismantlement. In particular, recycling of depleted uranium alloys has historically not been done, but new federal guidelines will require it. The recycling of alloys U-Ti and U-Nb will require segmentation and tracking to prevent cross-contamination.

Robot Sorting System

We have assembled a demonstration robotic waste sorting and classification system (**Fig. 1**). This automated workcell consists of a PUMA 560 articulating robot arm, a machine vision system, a conveyor system, a suite of remote sensors, and a hierarchical computer control system that coordinates the activities within the workcell. A network of computers located within the laboratory allows real-time control

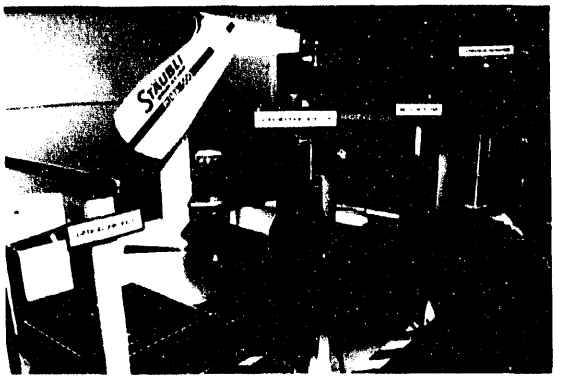


Figure 1. Interactive Controls Laboratory at Lawrence Livermore National Laboratory. The scrap conveyor is shown in the foreground along with the sensors used for material characterization. The PUMA robot arm with its wrist force/torque sensor is in the background. Not shown is a stereo camera pair mounted on the ceiling.

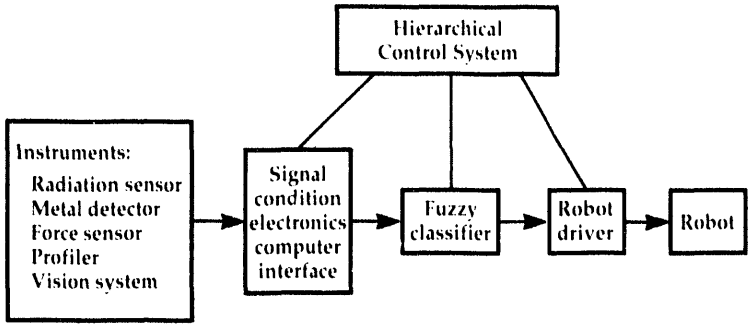


Figure 2. Block diagram of the robot control system. Input signals from various sensors are processed to determine the robot's next course of action. The fuzzy classifier categorizes the incoming material and provides this information to the robot driver.

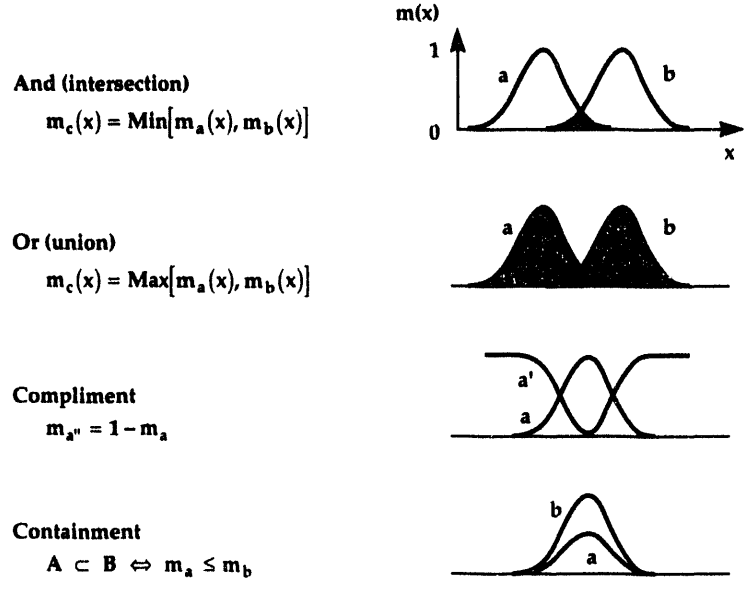


Figure 3. Operations on fuzzy sets. x is a measured quantity. The membership function $m_{(}(x)$ defines the boundaries of the fuzzy set $($. Membership in ordinary sets take on values in {0,1}, whereas fuzzy sets allow membership in the continuous range [0,1].

of the workcell through a three-dimensional graphical interface. Remote teleoperation of the robot arm is supported, with 'telepresence' information such as contact forces and stereo vision fed back to the operator.[1,2] Autonomous sensor-directed operation is being developed to handle difficult unstructured problems such as the scrap-sorting application described in this report.

Fuzzy Set Classifier

Robotic sorting of materials having unknown size, weight, and composition is a very different task from that of the usual picking and placing of known components in typical assembly line operations. The scrap materials must first be scanned by sensors located upstream in the conveyor system. The robot control system must then decide what action to take based on the sensory inputs. We have developed an expert system based on fuzzy logic theory to fuse the data from the multiple sensors and make classification decisions. **Figure 2** shows a block diagram of the robot control system. Sensing instruments include an optical profiler, gamma spectrometer, metal detector, force sensor (for weight), and machine vision system. From these measurements, we can infer size, density, metal content, and radiation content. Then, using the rule base, we can determine if the object is lead, steel, plastic or other light materials, or radiation-contaminated versions of these. The rule base consists of linguistic inference statements such as:

(if weight is heavy and size is small then density is high)

(if density is high and metal content is high then object is lead).

These rules are derived from experience by human operators using detectors and a set of known test items. These rules can be coded into a computer so long as the computer has an exact definition of terms like "heavy" and "high." As humans, we associate vaguely defined properties with these terms. To use this same idea of vagueness in a rigorous mathematical context, a theory of fuzzy sets has been developed.[3] This theory introduces a new type of set, called a fuzzy set, which generalizes the concept of an ordinary set. Fuzzy set operations, analogous to the logical operations on ordinary sets, are defined accordingly (see **Fig. 3**). It is then a simple matter to implement a fuzzy logic inference engine on a computer. A fuzzy logic system takes input data from an experiment, applies the rules in its rule base, and comes up with conclusions. The output is a list of possible conclusions along with their degrees of confidence. If a hard decision must be made, as is the case in the waste-sorting application (where the robot arm is instructed to place objects in various waste streams), then the conclusion with maximum confidence is taken.

Experimental Results

We performed experiments using gamma spectrometer data to compare the fuzzy logic inference technique to a detection threshold method in categorizing radiation sources. In this experiment, four spectral bands resolvable by the sodium iodide detector were selected to correspond to strong emission lines of americium 241 (50 to 70 keV), thorium 232 (210 to 270 keV), cesium 137 (600 to 730 keV), and cobalt 60 (1270 to 1430 keV), respectively. However, since the emission spectra for these sources overlap, there is a

Table 1. Results from an isotope identification experiment.

	Band 1	Band 2	Band 3	Band 4		
Am	1789	-50	-43	4	1 μCi sample	155 mm
bg	10	24	100	2		
Cs	15	5	2020	23	1 μCi sample	
bg	15	-48	33	-16		
Am	1719	48	20	6		
bg	6	66	23	-21		
Co	24	112	3	-16	1 μCi sample	
bg	11	58	3	-16		
Am	2043	28	-88	-24	smoke detector	
bg	24	-16	11	-2		
Th	-243	1298	-377	-37	Coleman lantern mantels	
bg	**17**	**161**	**-59**	**7**		
Th	88	120	-368	40	welding rods (under box)	
Th	303	1187	-1699	-52	welding rods (on top of box)	
bg	-6	72	2	-9		
Th	-520	13309	-7058	-146	lens	80 mm
Th	-891	3111	-1071	86	mantels	
Co	-16	120	-26	953	1 μCi sample	

The highlighted area indicates where the threshold algorithm with a 100-count threshold produces a false alarm: background (bg) misidentified as thorium 232 (Th). The fuzzy logic system, relying more on pattern recognition, correctly indicates this as background radiation, i.e., no source present. Some 'counts' are negative because of data preconditioning and normalization.

possibility that information concerning the presence of thorium, for example, is present in the cesium band, and so on. Therefore, information useful in the categorization of radioisotopes is contained in the pattern of the counts, not just in the counts per individual band. The rules in the fuzzy rule base were set accordingly. For example:

(if band 4 is high or (band 4 is high and band 2 is high), then source is cobalt 60).

Results from one test are shown in **Table 1**. We tested sample sources located from 80 to 150 mm from the detector, and also made an equal number of tests with no source present. The fuzzy logic system correctly identified the isotopes with 100% accuracy, while the threshold system had only 89% accuracy (one miss and one incorrect classification) with a 200-count threshold, and 72% accuracy (three incorrect classifications and two false alarms) with a 100-count threshold. As the sources are separated from the detector by larger distances, the signal becomes weaker, so the sensor and fuzzy logic system begin to fail to detect the radiation. However, even with weak signals, the isotope signature is often still present. We set up the fuzzy logic system to guess the isotope, even if the radiation count was low. At average separations of roughly 300 mm, the threshold system was failing nearly 100% of the time, while the fuzzy classifier was making correct guesses with about 50% accuracy.

Future Work

Sorting and classification of materials will be a crucial task in the weapons dismantlement process. Special nuclear materials resulting from dismantlement need to be identified and tracked by an automated system to prevent unauthorized diversion from the recycle stream. Depleted uranium alloys should be segmented from each other to prevent cross-contamination. We see this as a future growth area for multisensor fusion and fuzzy classification systems such as the one we have developed.

1. D.T. Gavel and S.-y. Lu, "Telerobotics and Machine Vision," *Engineering Research, Development, and Technology*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-53868-91,9–6 (1992).
2. A. Dougan, D.T. Gavel, D. Gustaveson, M. Holliday, R. Hurd, R. Johnson, B. Kettering, and K. Wilhelmsen, *"Demonstration of Automated Robotic Workcell for Hazardous Waste Characterization,"* submitted to 1993 IEEE International Conf. Robotics and Automation (Atlanta, Georgia), (May 1993).
3. L.A. Zadeh, *Inf. Control* 8, 338 (1965).

Adaptive Optics for Laser Guide Stars

James M. Brase,
Kenneth Avicola,
Donald T. Gavel, and
Kenneth E. Waltjen
Laser Engineering Division
Electronics Engineering

Horst D. Bissinger
Energy Systems Engineering
Mechanical Engineering

We are investigating advanced concepts in adaptive optics (AO) systems and developing a comprehensive analysis and modeling capability to predict the performance of AO systems. In FY-92, we demonstrated the generation of a Na guide star and verified our models of its formation. We have made the first Hartmann-sensor wavefront measurement from a Na guide star, and evaluated its potential as a reference for a closed-loop AO system.

Introduction

Turbulence in the atmosphere blurs images seen in ground-based telescopes and places a severe limit on their angular resolution. Typical atmospheric blurring is so severe that even a 10-m telescope has no better resolution than a small 8-in. telescope, despite the fact that the larger instrument gathers far more light.

There are two methods for gaining dramatically improved resolution. The first is to go above the atmosphere, as did the Hubble Space Telescope. This approach has the additional advantage that regions of the spectrum such as the ultraviolet, which cannot penetrate the atmosphere, are accessible. However, going into space is expensive and inherently less flexible than observing from the ground. The second alternative is to use a technique called 'adaptive optics' (AO) to improve resolution for ground-based telescopes. We are investigating advanced concepts in AO systems and developing a comprehensive analysis and modeling capability to predict the performance of AO systems.

AO systems have been demonstrated for astronomical applications.[1] These systems use a bright natural star as a reference to correct the dimmer astronomical object. One of the major problems with applying AO to astronomy is the scarcity of natural stars close enough and bright enough to serve as references. Our approach to solving this problem is shown in **Fig. 1**. We will use the copper-vapor pumped dye laser system, developed for laser isotope separation at Lawrence Livermore National Laboratory (LLNL), to illuminate a small circular area of the atmospheric sodium lay-

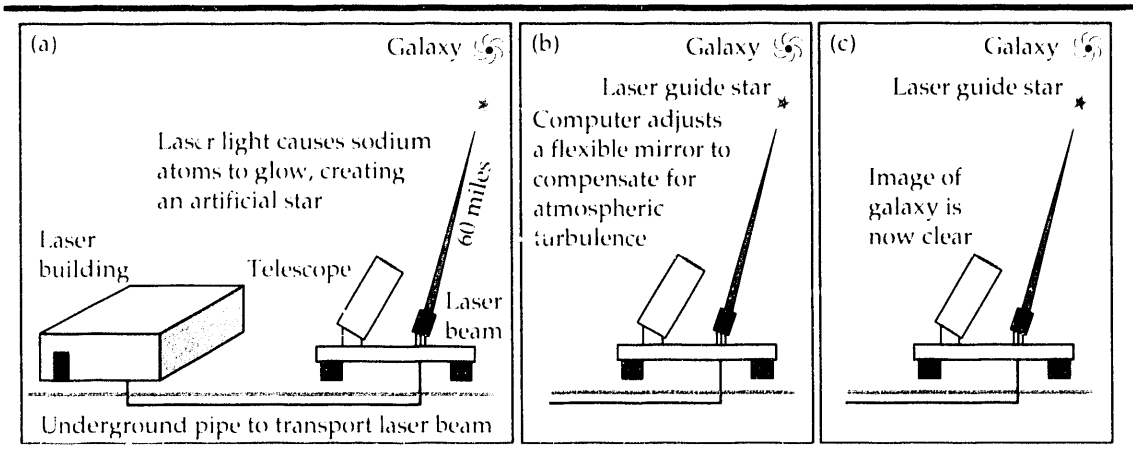


Figure 1. Use of the laser guide star system to remove atmospheric distortion and improve the resolution of ground-based telescopes. (a) Laser guide star is created. (b) Adaptive optics correct for atmospheric distortion using the reference. (c) Corrected astronomical image is formed.

Figure 2. Na guide star. The laser guide star is the small round spot on the right end. The long streak leading to it is Rayleigh scatter from a point lower in the atmosphere.

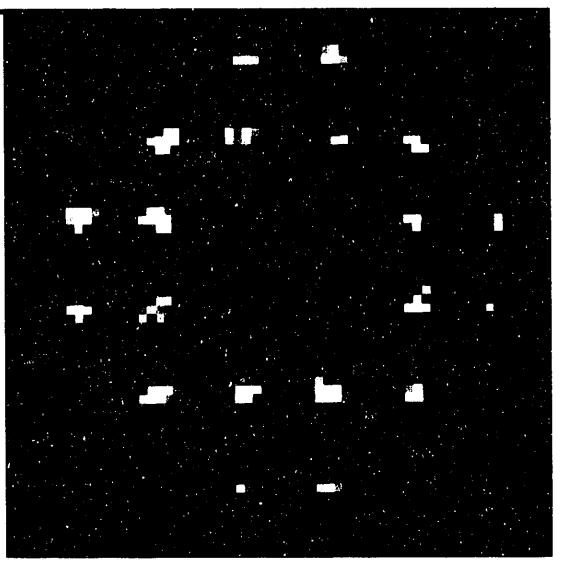


Figure 3. A Hartmann sensor image from the laser guide star, in which each small spot corresponds to a part of the telescope aperture. By analyzing the motion of the spots, we can reconstruct the atmospheric turbulence.

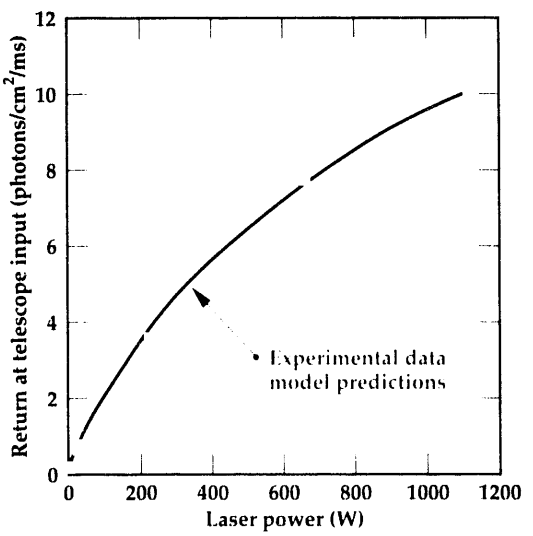


Figure 4. The predicted sodium emission intensity (solid line) vs the experimental measurements (squares) from the Na guide star at several laser power levels.

er at a height of about 100 km. When the laser is tuned to the proper wavelength, the sodium will glow and produce a point-like reference source. An observing telescope on the ground measures in detail the light coming from this 'laser guide star' and, with the aid of a computer, deduces what distortions have been placed on the wavefront by atmospheric turbulence. The computer then calculates the corrections to be applied to a deformable mirror in the optical train of the telescope to correct for the turbulence. The light from a nearby astronomical object is also corrected by the deformable mirror so that an improved image is formed.

The basic technologies for laser-guide-star AO systems have been demonstrated over the past ten years.[2,3,4] Success with Na guide stars has been limited by the lack of an appropriate laser. However, the LLNL copper-vapor pumped dye laser is well-suited for the demonstration of astronomical laser guide stars. It has more than enough power at the Na wavelength (1.5 kW at 589 nm), excellent reliability, and high beam quality. We are performing a series of feasibility experiments on laser guide stars using this laser. From the data obtained in these experiments, we will be able to design a smaller and more economical system optimized for use at an astronomical observatory. Our long-term goal is to establish a technology base in AO that will allow us to implement a system for a large astronomical instrument such as the 10-m Keck Observatory telescope.

The laser guide star experiments at LLNL are being done in two phases. In the first, which began in July 1992, we have generated a Na guide star (**Fig. 2**) and have measured its intensity and motion.[5] In the second phase, currently underway, we are developing an AO system to demonstrate closed-loop correction of an astronomical object with a Na guide star.[6]

Progress

In FY-92, we demonstrated the generation of a Na guide star and verified our models of its formation. We have made the first Hartmann sensor wavefront measurement from a Na guide star, and evaluated its potential as a reference for a closed-loop AO system.

Wavefront Sensing

An important part of an AO system is the sensor that analyzes the laser guide star wavefront in real time. Over the past year, we have developed a new high-speed Hartmann wavefront sensor[7] ca-

pable of measuring local wavefront slopes at one thousand frames per second. In a recent series of experiments, we made the first Hartmann-sensor wavefront measurements of a Na laser guide star. A typical Hartmann image is shown in **Fig. 3**. The motion of individual spots in these images is analyzed to estimate local wavefront slopes. These slopes are ultimately integrated into the wavefront phase distribution, which is used to control the deformable mirror. We are in the process of analyzing this preliminary data and performing more experiments to characterize the performance of this system.

We have also performed a series of wavefront sensing experiments using natural stars, to determine requirements for AO systems at LLNL. These experiments will be expanded to include the University of California's Lick Observatory on Mt. Hamilton as the first step towards implementation of an AO system there.

Analysis and Modeling

Our long-term goal is to develop laser guide star systems for 10-m-class telescopes like that of the Keck Observatory. The initial development, however, will take place on smaller telescopes both at LLNL and at Lick Observatory. It is vital that we use computer simulations to understand the scaling of the results from our demonstration experiments, to what we should expect from large astronomical telescopes. The initial experiments will allow us to validate our simulations, so that we can have a greater degree of confidence in the results for 10-m telescopes.

Some problems that will arise on large telescopes will not be evident in our smaller systems. For example, as the telescope gets large, a single laser guide star can no longer be used to correct the entire aperture, because of the finite height of the laser guide star. Multiple laser guide stars must be generated to accurately correct the images. A complete simulation will allow us to develop these techniques even before we have access to a large telescope.

To date, implementation of astronomical AO systems has been devoted mainly to system development. Very little actual astronomy with adaptively corrected telescopes has yet been done anywhere in the world. Because of the change in quality of the correction across the field of view and with changes in atmospheric conditions, standard astronomical measurements such as photometry and spectroscopy may become more complicated. Our simulation tools will allow us to explore these problems before large-scale AO systems are designed.

One of our first tasks in model validation has been to compare the results of our initial laser-guide-star experiments with the predictions of our simulations. The comparison of predicted Na emission intensity with the experimental measurements is shown in **Fig. 4**. The excellent agreement increases our confidence in other simulation results.

Future Work

In FY-93, we will demonstrate closed-loop AO correction of a small telescope at LLNL using a Na guide star. This demonstration experiment will require the wavefront sensing technology that we have developed. It will also allow us to verify our analysis and simulation tools. We are beginning to apply these techniques to a variety of new problems in high-resolution imaging and beam control.

1. G. Rousset, J.C. Fontanella, P. Kern, D. Gigan, F. Rigaut, P. Lena, C. Boyer, P. Jagourel, J.P. Gaffard, and F. Merkle, *Astron. Astrophys.* **230**, L29 (1990).
2. R. Fugate, D. Fried, G. Ameer, B. Boeke, S. Browne, P. Roberts, R. Ruane, G. Tyler, and L. Wopat, *Nature* **353** (September 12, 1991).
3. C. Primmerman, D. Murphy, D. Page, B. Zollars, and H. Barclay, *Nature* **353** (September 12, 1991).
4. C. Gardner and L. Thompson, *Proc. IEEE* **78** (11), 1721 (1990).
5. K. Avicola, J.M. Brase, J.R. Morris, H.D. Bissinger, H.W. Friedman, D.T. Gavel, C.E. Max, S.S. Olivier, R.W. Presta, D.A. Rapp, J.T. Salmon, and K.E. Waltjen, *Sodium-Layer Laser Guide Star Experimental Results*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-111896 (1992).
6. C.E. Max, H.W. Friedman, J.M. Brase, K. Avicola, H.D. Bissinger, D.T. Gavel, J.A. Horton, J.R. Morris, S.S. Olivier, R.W. Presta, D.A. Rapp, J.T. Salmon, and K.E. Waltjen, *Design, Layout, and Early Results of a Feasibility Experiment for Sodium-Layer Laser Guide Star Adaptive Optics*, Lawrence Livermore National Laboratory, Livermore, California, UCRL-JC-112162 (1992).
7. R.K. Tyson, *Principles of Adaptive Optics*, Academic Press (Boston, Massachusetts), 1991.

Alesso, H.P. 4-27
Angel, S.M. 6-17
Avalle, C.A. 1-21, 7-23
Avicola, K. 9-27
Azevedo, S.G. 8-5

Balch, J.W. 3-21
Belak, J. 5-7
Biltoft, P.J. 5-5
Bissinger, H.D. 9-27
Boercker, D.B. 5-1, 5-7
Branscomb, E.W. 4-29
Brase, J.M. 9-11, 9-27
Brinkmann, R.P. 7-13
Brown, A.E. 6-11
Bryan, Jr., S.R. 5-5
Buettner, H.M. 4-31
Buhl, M.R. 9-15

Caplan, M. 1-13
Chow, R. 3-1
Christon, M.A. 2-19
Ciarlo, D.R. 3-1, 3-15
Colvin, M.E. 4-29
Cooper, G.A. 3-1
Cravey, W.R. 7-19

Daily, W.D. 4-31
Davidson, J.C. 3-21
DeFord, J.F. 1-13
De Groot, A.J. 2-11
DeMartini, D.C. 4-17
DeTeresa, S.J. 6-11
Dijaili, S.P. 3-1, 3-5
Donich, T.R. 4-13
Douglass, B. 7-5

Engelmann, B.E. 2-1

Falabella, S. 5-1, 5-5
Faux, D.R. 4-21
Feng, W.W. 6-11
Foiles, L. 7-19

Gavel, D.T. 9-23, 9-27
Glass, R.S. 3-13
Goodman, D.M. 9-7
Govindjee, S. 2-35
Grant, J.B. 1-25
Groves, S.E. 6-11

Harris, D.B. 4-17
Hawkins, R.J. 1-7
Hawley-Fedder, R.A. 7-19
Hernandez, J.E. 9-1, 9-15
Hernandez, J.M. 7-5
Heuze, F.E. 2-27
Hofer, W.W. 7-13
Hoover, C.G. 2-11
Hui, W.C. 3-19
Hutchings, L.J. 2-27

Jarpe, S.P. 2-27, 4-17
Johansson, E.M. 7-5
Johnson, R.K. 9-1
Johnson, R.R. 4-9
Joshi, R. 7-13
Judson, R.S. 4-29

Kallman, J.S. 1-7
Kania, D.R. 7-13
Kay, G.J. 2-35
Khanaka, G.H. 3-15
Kirbie, H.C. 7-27
Koo, J.C. 3-1, 3-21

Landram, C.S. 4-13
Lauer, E. 7-27
Laursen, T.A. 2-7
Lee, H. 7-5
Lehman, S.K. 9-11
Lesuer, D.R. 6-1, 6-23
Liliental-Weber, Z. 3-1
Lloyd, W.R. 4-27
Logan, R.W. 4-1
Lu, S. 4-29, 9-1
Luedtka, W.R. 7-19
Lyon, R.E. 6-11, 6-17

Madsen, N.K. 1-1
Maker, B.N. 2-7
Maltby, J.D. 2-11, 2-23
Mariella, Jr., R.P. 3-1
Martz, H.E. 8-5
Mascio, L.N. 9-21
McAllister, S.W. 4-13
McCallen, D.B. 2-27
McConaghy, C.F. 3-5
McKinley, B.J. 8-1
Milanovich, F.P. 8-1
Morse, J.D. 3-5, 3-9
Myrick, M.L. 6-17

Nelson, S.D. 1-21, 7-5

Olsen, B.L. 5-5

Payne, A.N. 7-27

Pearson, J.S. 4-27
Phillips, J.P. 9-11
Pombo, R.F. 5-5
Preuss, C.S. 6-23
Prosnitz, D. 7-27

Raboin, P.J. 6-23
Randich, E. 3-15
Roberson, G.P. 8-5
Rosinsky, R.W. 4-21

Sampayan, S.E. 7-27
Sanchez, R.J. 6-11
Sanders, D.M. 5-1
Schneberk, D.J. 8-5
Schoenbach, K.H. 7-13
Shang, C.C. 1-13
Shapiro, A.B. 6-7
Sherby, O.D. 6-1
Sherwood, R.J. 4-17
Sinz, K.H. 4-23
Stowers, I.F. 5-7
Syn, C.K. 6-1, 6-23
Szoke, H. 9-11

Thomas, G.H. 8-23

Vess, T.M. 6-17
Vogtlin, G.E. 7-1

Waltjen, K.E. 9-27
Warhus, J.P. 7-5
Whirley, R.G. 2-1, 2-11
Wieting, M.G. 9-11

Yee, J.H. 3-15
Yu, C.M. 3-13

Zacharias, R.A. 7-23
Ziolkowski, R.W. 1-7
Zywicz, E. 2-15

END

DATE FILMED

9/30/93