

NASA Conference Publication 3263

Dual-Use Space Technology Transfer Conference and Exhibition

Volume I

*Proceedings of a conference and exhibition held at
Lyndon B. Johnson Space Center
Houston, Texas
February 1-3, 1994*



NASA Conference Publication 3263

Dual-Use Space Technology Transfer Conference and Exhibition

Volume I

*Compiled by
Kumar Krishen
Lyndon B. Johnson Space Center
Houston, Texas*

Proceedings of a conference and exhibition held at
Lyndon B. Johnson Space Center
Houston, Texas
February 1-3, 1994



National Aeronautics and
Space Administration

1994

CONTENTS

INTRODUCTION

PLENARY SESSION

KEYNOTE ADDRESS

Session C1: INNOVATIVE MICROWAVE AND OPTICAL APPLICATIONS

Session Chair: G. D. Arndt

Microwave Sensor for Ice Detection.....	1 - 1
Determination of the Residence Time of Food Particles During Aseptic Sterilization	7 - 2
Electromagnetic Probe Technique for Fluid Flow Measurements.....	15 - 3
Dual Use of Image Based Tracking Techniques: Laser Eye Surgery and Low Vision Prosthesis.....	23 - 4

Session C2: COMMUNICATIONS AND DATA SYSTEMS

Session Chair: Andrew Benjamin

ESTL High Rate Optical Communications System (EHROCS).....	31
A Mobile Communications Space Link Between the Space Shuttle Orbiter and the Advanced Communications Technology Satellite	42
TDMA Algorithm for the Space to Space Communications System	48
Transition from NASA Space Communication Systems to Commercial Communication Products.....	60

Session C3: COMMUNICATIONS SIGNAL PROCESSING AND ANALYSIS

Session Chair: Robert Panneton

Scattering Effects of Solar Panels on Space Station Antenna Performance.....	67
Dynamic Environment Communications Analysis Testbed (DECAT) and Its Applications to Dual Use Space Technology	74
Software-Implemented Fault Tolerance in Communications Systems	82
Signal Processor Developments by Personnel of the JSC Signal Processing Section.....	90

Session C4:	APPLICATIONS DERIVED FROM CONTROL CENTER DATA SYSTEMS	
Session Chair:	Marvin LeBlanc	
	The DASH Data Sharing System	103
	Adaptation of a Control Center Development Environment for Industrial Process Control	111
	Adaptation of Control Center Software to Commercial Real-Time Display Applications.....	119
Session G3:	INTEGRATED VEHICLE HEALTH MANAGEMENT	
Session Chair:	Wayne McCandless	
	Dual-Use Aspects of System Health Management.....	127
	Reaction Jet Drive Electronics	147
	A Multichannel Data Acquisition Module with State-Based Feature Recognition for System Health Management	153
Session G4:	ADVANCED AVIONICS	
Session Chair:	John Ruppert	
	Multi-flight-phase GPS Navigation Filter Applications to Terrestrial Vehicle Navigation and Positioning.....	159
	Automated Launch Operations	160
Session H1:	HUMAN FACTORS TECHNOLOGY DUAL USE	
Session Chair:	John Schuessler	
	Human Factors Engineering: Current and Emerging Dual-Use Applications.....	167
	Research on Personal Protective Equipment for Dual-Use Technology and Technology Transfer	174
Session H2:	HUMAN PERFORMANCE EVALUATION	
Session Chair:	Barbara Woolford	
	Advanced Video Analysis Needs for Human Performance Evaluation	187
	Use of Video Analysis System for Working Posture Evaluations.....	195
	Application Reuse Library for Software, Requirements, and Guidelines.....	203
	NASA's Man-Systems Integration Standards: A Human Factors Engineering Standard for Everyone in the Nineties.....	211

Session H3:	SYSTEMS/PROCESSES IN HUMAN SUPPORT TECHNOLOGY	
Session Chair:	John Schuessler	
	Solar Photovoltaic Powered Heat Pump.....	221
	Kinetic Study of Methyl Acetate Oxidation in a Pt/A1 ₂ O ₃ Fixed-Bed Reactor.....	226
	Solid State Oxygen Sensor Development.....	236
	Performance Enhancement of Heat and Mass Transfer Devices With Electrohydrodynamics (EHD).....	242
Session L1:	HUMAN FACTORS AND HABITATION	
Session Chair:	R. Bond	
	Human-Computer Interaction with Medical Decision Support Systems.....	255
	Development of Shelf Stable Tortillas for Space Missions.....	263
	Large-Scale Numerical Simulations of Human Motion.....	270
	Advanced Life Support Systems: Opportunities for Technology Transfer.....	283
Session L3:	MEDICAL CARE	
Session Chair:	J. Homick	
	A Feasibility Study for Perioperative Ventricular Tachycardia Prognosis and Detection and Noise Detection Using a Neural Network and Predictive Linear Operators.....	295
Session M1:	MATERIALS AND STRUCTURES I	
Session Chair:	David Hamilton	
	Vibration Control of Deployable ASTROMAST Boom - Preliminary Experiments.....	305
	Structural Health Monitoring of Large Structures	314
	An Analytical Tool for Fracture Control.....	319
	Design and Fabrication of NDE Standards for Fracture Control.....	326
	Modal Test Technology as Non-Destructive Evaluation of Space Shuttle Structures	329
	Development of Advanced Alloys Using Fullerenes	335

Session M2: MATERIALS AND STRUCTURES II
Session Chair: Lubert Leger

Ilmenite as a Dual-Use Material347

The Materials Chemistry of Atomic Oxygen With Applications to
Anisotropic Etching of Submicron Structures in Microelectronics
and the Surface Chemistry Engineering of Porous Solids354

Application of Chlorine-Assisted Chemical Vapor Deposition of Diamond at
Low Temperatures.....368

Adaptation of an Ammonia Detecting Paint to Commercial Applications.....374

Development and Testing of Textile Fibers in the JSC Crew and Thermal
Systems Division.....383

Session P2: POWER I
Session Chair: Dave Belanger

Electric Auxiliary Power Unit (EAPU) for the Space Shuttle.....389

Session P3: POWER II
Session Chair: Dave Belanger

A PC Based Time Domain Reflectometer for Space Station Cable Fault
Isolation..... 395

Space Batteries and Their Role in Dual Use Technology404

Session P4: POWER III
Session Chair: Bill Boyd

Proton Exchange Membrane Fuel Cells for Space and Electric Vehicle
Applications—From Basic Research to Technology Development409

Ozone as a Laundry Agent on Orbit and on the Ground.....421

Session R1: PERCEPTION AND VISION TECHNOLOGIES
Session Chair: Thomas W. Pendleton

Object Recognition and Pose Estimation of Planar Objects from Range Data.....427

Microwave Imaging of Metal Objects.....435

Robotic Vision Techniques for Space Operations.....441

Perception for Mobile Robot Navigation: A Survey of the State of the Art446

Session R2: AUTOMATION TECHNOLOGIES
Session Chair: Kathleen Jurica

Computer-Aided Operations Engineering with Integrated Models of Systems
and Operations455

A Prototype Supervised Intelligent Robot for Helping Astronauts462

Failure Environment Analysis Tool Applications472

NASA Scheduling Technologies482

Config - Integrated Engineering of Systems and Operations498

Session R3: ROBOTICS TECHNOLOGIES
Session Chair: Reginald Berka

Lyndon B. Johnson Space Center (JSC) Proposed Dual-Use Technology
Investment Program in Intelligent Robotics505

Dual-Use Technologies for the Mining, Processing, and Energy Industries513

Virtual Reality Applications in Robotic Simulations519

A Distributed Telerobotics Construction Set523

Dual Use Display Systems for Telerobotics529

Session S1: INFORMATION ACCESS
Session Chair: Chris Culbert

A New Machine Classification Method Applied to Human Peripheral
Blood Leukocytes535

Spatial Data Management System (SDMS)546

A Method for Automatically Abstracting Visual Documents552

Session S2: INTELLIGENT SYSTEMS
Session Chair: Wade Webster

CLIPS: The C Language Integrated Production System569

A Feasibility Study for Long-Path Multiple Detection Using a Neural
Network 577

Intelligent Computer-Aided Training Authoring Environment584

Session S3:	SOFTWARE ENGINEERING I	
Session Chair:	Susan Gerhart	
	Reengineering Legacy Software to Object-Oriented Systems.....	595
	A Measurement System for Large, Complex Software Programs.....	601
Session S4:	SOFTWARE ENGINEERING II	
Session Chair:	Ernest Fridge	
	Understanding Software Faults and Their Role in Software Reliability Modeling.....	609
	Experiences in Improving the State of the Practice in Verification and Validation of Knowledge-Based Systems.....	617
Session S5:	MATHEMATICS, MODELING AND SIMULATION	
Session Chair:	Sam Veerasamy	
	Recent Advances in Wavelet Technology.....	625
	A Method to Compute SEU Fault Probabilities in Memory Arrays with Error Correction.....	633
	Computer Simulation: A Modern Day Crystal Ball?.....	638
Session S6:	NETWORKS, CONTROL CENTERS, AND DISTRIBUTED SYSTEMS	
Session Chair:	Zafar Tacqvi	
	Average Waiting Time in FDDI Networks with Local-Priorities.....	643
	The Electronic Documentation Project in the NASA Mission Control Center Environment.....	649
Session T1:	MARKETING AND BARRIERS	
Session Chair:	Robert Brown	
	Invention Driven Marketing.....	657
Session T2:	BUSINESS PROCESS AND TECHNOLOGY TRANSFER	
Session Chair:	Russ Cargo	
	Weapons to Widgets: Organic Systems and Public Policy for Tech Transfer.....	667
	Business Process Re-engineering—A Precursor to Technology Transfer.....	675
	FuzzyCLIPS from Research to Product.....	679

Session T3:	NEW WAYS OF DOING BUSINESS	
Session Chair:	John P. Van Blois	
	Improved Ways of Doing Business Between Industry and Government.....	681
	National Aero-Space Plane Program.....	689
Session T4:	TECHNOLOGY TRANSFER METHODS	
Session Chair:	Ken Cox	
	Speeding the Payoff: New Paradigms in Technology Transfer	721
	Accelerating the Improvement of Software Practice: Squeezing 20 Years into an 8-Year Time Capsule.....	738
	The Meeting of Needs: Success Factors in Government/Industry Technology Transfer	750
	Unique Strategies for Technical Information Management at Johnson Space Center	756
Session T5:	TECHNOLOGY TRANSFER—EXAMPLE I	
Session Chair:	Jack Aldridge	
	The Retinal Funduscope Demonstration Project	769
	Depth Imaging from Eastman Kodak Company.....	777
Session T6:	TECHNOLOGY TRANSFER—EXAMPLE II	
Session Chair:	Kyle Fairchild	
	Artificial Bone from Space Shuttle Tile.....	781
	Developing Technology Commercialization Infrastructure.....	787
Session T7:	TECHNOLOGY TRANSFER “HOW TO”	
Session Chair:	George Ulrich	
	Licensing NASA Technology	803
	Technology as a Commercial Resource.....	809
	AUTHOR INDEX.....	819

INTRODUCTION

INTRODUCTION

Kumar Krishen, Ph.D.

On February 22, 1993, President Clinton and Vice President Gore issued their technology policy titled, "Technology for America's Economic Growth, A New Direction to Build Economic Strength." The critical theme in their policy deals with redirecting America's research and technology programs to realize economic growth and create an educational and training system that is challenging for workers. In November 1993, the Department of Commerce issued a draft report titled, "Commerce ACTS: Advanced Civilian Technology Strategy for Jobs and Economic Growth." A key component of this strategic plan deals with increasing the transfer of technology from Federal laboratories to private industry for commercialization and international competitiveness. In the numerous documents issued by White House officials and Federal agencies during the past 12 months, the message given is that technology creates new businesses, new markets, and new jobs—the fundamental engine of sustained growth to the economy.

In concert with the Administration's focus on technology policy and transfer of federally developed technology to the private sector, NASA has developed the Space Technology Enterprise as one of the five strategic enterprises that have been planned for the Agency. The objectives of the Space Technology Enterprise, as stated in a draft of NASA's Strategic Plan dated February 15, 1994, are to:

- Proactively transfer technology to aerospace and nonaerospace industries to enhance U.S. competitiveness.
- Develop new and innovative space technologies to improve the performance and lower the cost of future space missions.
- Develop technology to revitalize access to space.

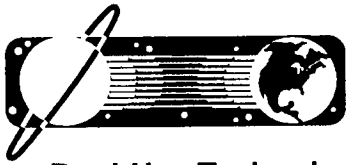
To date, space research and technology have resulted in more than 30,000 "spin-offs" or second-hand uses of technology. The commercial applications of space technology include health and medicine, Earth environment and resource management, transportation, industrial productivity, public safety and home systems, training and education, communications, and information systems. Today, as a result of numerous technological innovations and developments, NASA is in an excellent position to share its research and technology with the Nation's industrial community.

At Johnson Space Center, one of the activities in support of NASA's thrust in technology transfer and commercialization was the development and implementation of the Dual-Use Space Technology Transfer Conference and Exhibition with the focus on bridging the space community with industry. This conference was intended to inform industry of NASA's existing programs and available mechanisms for technology transfer and brought speakers from aerospace industry to describe their existing technologies and future plans. A number of key NASA and aerospace industry representatives provided information regarding important opportunities for space technology transfer for commercialization. Technical areas included were Software and Computer Technology, Human Support Technology, Avionics, Guidance and Control, Propulsion and Power, Communications and Data Processing, Robotics and Automation, Materials Technology, Medical and Life Sciences, and Technology Transfer. At this conference, more than 150 papers and talks were given, more than 35 technologies were exhibited, 325 attendees were registered, and more than 300 visited the exhibits. We believe that the conference and exhibition were a great success. These proceedings will provide documentation on key projects and ideas and help promote interaction between space researchers and technologists and their counterparts in other industries.

Table 1 lists the most prominent persons and organizations who made the Dual-Use Space Technology Transfer Conference and Exhibition possible. I wish to thank each of the individuals and organizations listed for their outstanding support. This conference was an effort to bring together government scientists, engineers, and managers with leaders in business. It is our hope that

we forged links which will strengthen America's economic competitiveness by sharing technology already developed and giving insight into future research trends.

Dr. Kumar Krishen
Program Co-Chairperson
Code HA
NASA Johnson Space Center
Houston, TX 77058



Dual-Use Technology Transfer Conference will include programmatic overviews, panel sessions, exhibits, and technical papers in the following areas:

- Software and Computer Technology
- Human Support Technology
- Avionics, Guidance and Control
- Propulsion and Power
- Communications and Data Processing
- Robotics and Automation
- Materials and Structures
- Medical and Life Support
- Technology for Technology Transfer

Exhibit Hours

Tuesday, February 1	11:00 am – 7:00 pm
Wednesday, February 2	8:00 am – 7:00 pm
Thursday, February 3	8:00 am – Noon

Plenary Session (February 1, 9:00 am–11:30 am)

Chairman:	Mr. Jim Kollaer <i>Opening Remarks & Welcome</i>
Invited Speakers:	Mr. Jim Kollaer Dr. Carolyn Huntoon Mr. Greg Reck Dr. Neil Pellis Dr. George Kozmetsky Dr. Kumar Krishen

Keynote Dinner Session (February 2, 6:30 pm – 9:30 pm)

Keynote Speaker:	Mr. William Huffstetler NASA/Johnson Space Center
------------------	--

Technical Area Track Chair

Software and Computer Technology
Dr. Troy Henson, Rice University

Human Support Technology
Mr. John Schuessler, McDonnell Douglas

Avionics, Guidance and Control
Mr. Don Brown, JSC
Mr. Ken Goodwin, Draper Lab

Propulsion and Power
Mr. John Griffin, JSC

Communications and Data Processing
Mr. Sidney Novosad, JSC

Robotics and Automation
Mr. Norman Chaffee, JSC

Materials Technology
Dr. Lubert Leger, JSC

Medical and Life Sciences
Dr. John Rummel, JSC

Technology for Technology Transfer
Mr. Robert Brown
Mr. George Ulrich, Consultant

Committee Members

Administrative Co-Chairpersons

Ms. Amy Reynolds, NASA/
Texas Aerospace Commission
Ms. Carla Colangelo, I-NET Inc.

Exhibit Co-Chairpersons

Mr. Chris Ortiz, NASA/JSC
Mr. Scot McCullough/ I-NET Inc.

General Co-Chairpersons

Dr. Aaron Cohen, Texas A&M University
Mr. Jim Kollaer, Greater Houston Partnership
Mr. William Huffstetler, NASA/JSC

Program Co-Chairpersons

Dr. Kumar Krishen, NASA/JSC
Dr. Hatem Nasr, Honeywell
Mr. Robert Savely, NASA/JSC

Publicity Co-Chairpersons

Mr. Stephen Nesbitt, NASA/JSC
Mrs. Joy Robertson, NASA/JSC

Planning and Coordinating

Co-Chairpersons

Dr. Debora Jones, NASA/JSC
Mr. Richard Rogers, Lockheed Engineering
& Sciences Co.

Arrangements Co-Chairpersons

Dr. Troy Henson, Rice University
Mr. Thomas Wolfmeyer, Rockwell
Ms. Lana Arnold, Lockheed Engineering
& Sciences Co.

Advisory Committee

Mr. Wayne Alexander,
Southwestern Bell Telephone/
Texas Aerospace Commission
Dr. George Kozmetsky, IC² Institute/
University of Texas
Dr. Roger Elliott, Texas Higher Education
Coordinating Board
Mr. C. J. Reinhartsen
Clear Lake Area Economic Foundation

Table 1. Key Persons Who Organized and Implemented the Conference and Exhibition

PLENARY SESSION



NASA JOHNSON SPACE CENTER TECHNOLOGY COMMERCIALIZATION STRATEGIES AND INFRASTRUCTURE

PLENARY SESSION

Dual-Use Space Technology Transfer
Conference and Exhibition
February 1-3, 1993

NASA/JOHNSON SPACE CENTER
Gilruth Center

Kumar Krishen, PH.D.
NASA/Johnson Space Center

94/1-10-0006



Outline

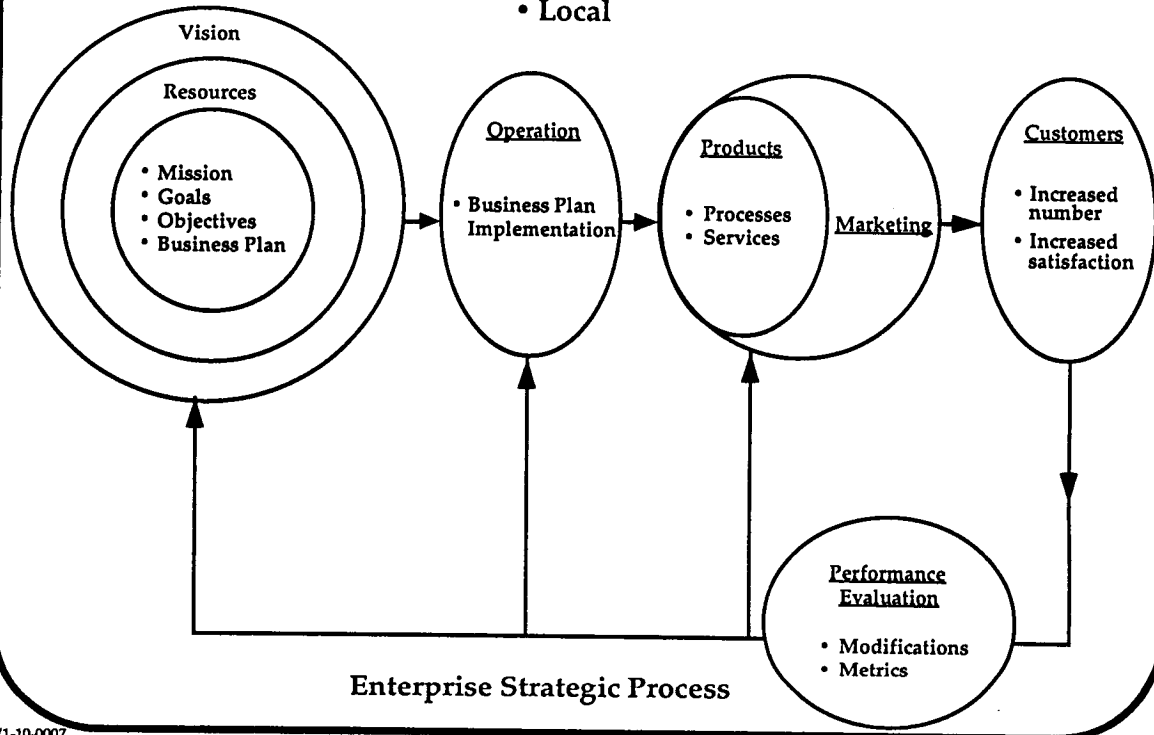
- The Challenge for Us
- Technology Interdependencies
 - STIG
 - SATWG
- Technology Transfer at JSC
- Space Technology Needs/Requirements
- JSC Areas of Research and Technology
- JSC Inventions and Discoveries
- Dual-Use Space Technology Conference and Exhibition Sponsors
- List of Exhibitions

94/1-10-0028



Political and Commercial Environment

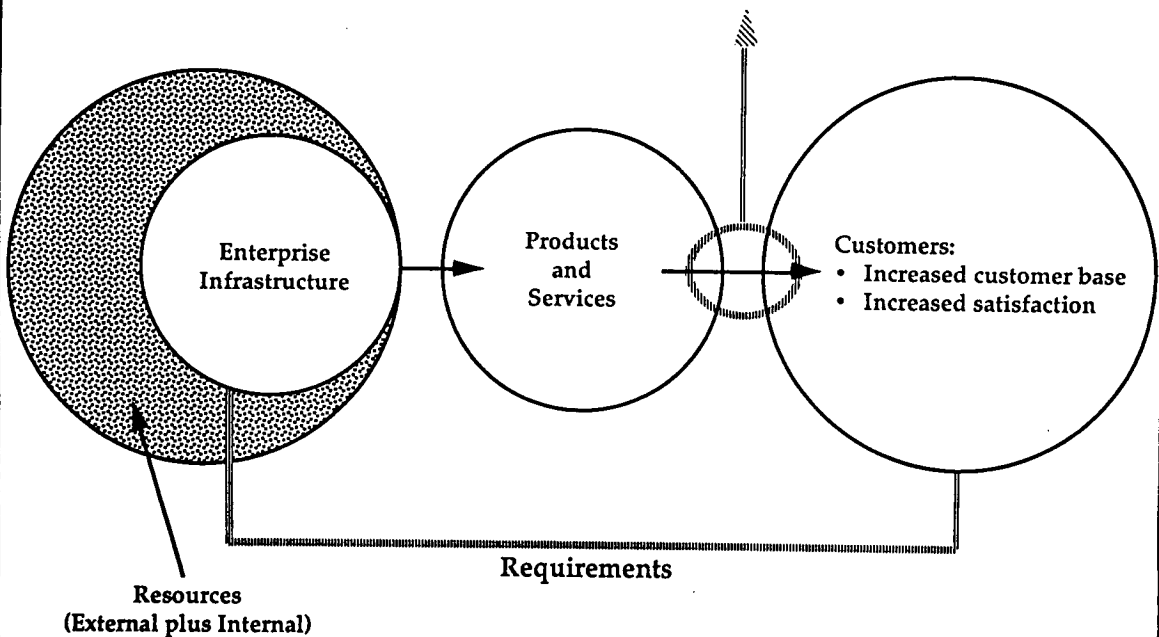
- Global
- Local



94/1-10-0007



Technology Transition



94/1-10-0008



The Challenge for Us

- **Develop efficient research and technology implementation processes/infrastructure**
 - Utilize experiences and lessons learned
- **Implement "processes" which make "transition" of technology possible**
 - Continuously improving
 - Part of program/project from inception
- **Utilize the Dual-Use Space Technology Conference and Exhibition as part of the "process"**
 - Provide feedback on areas of improvement

94/1-10-0009



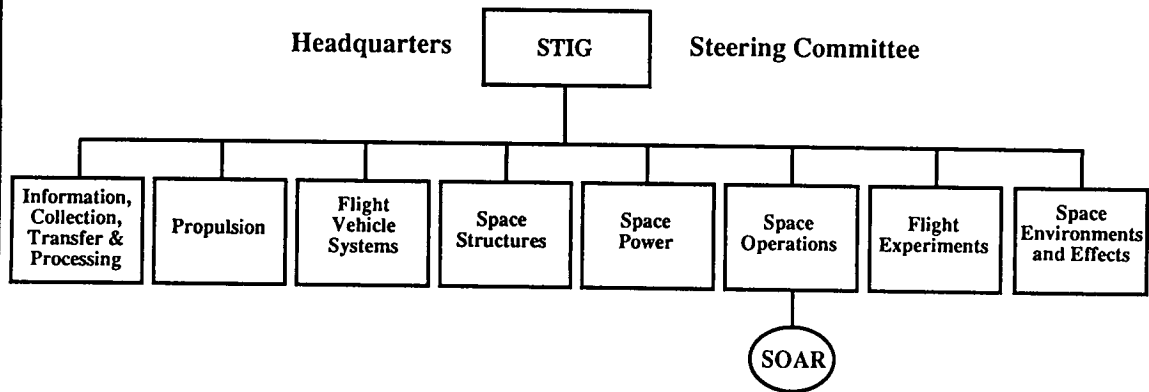
Space Technology Interdependency Group (STIG)

- **STIG was established in May 1982 to identify and promote the pursuit of new opportunities for cooperative relationships and monitor ongoing cooperative activities between the National Aeronautics and Space Administration (NASA) and the U. S. Air Force.**
- **Since then the U. S. Army and Navy, the Advanced Research Projects Agency, the Ballistic Missile Development Office, the Department of Energy, and the Department of Commerce have joined.**
- **The goal of STIG is to provide advocacy, oversight, and guidance to facilitate and encourage cooperative development programs and to avoid duplication of effort and resources on space technology activities.**

94/1-10-0010



STIG COMMITTEES



94/1-10-0011



Strategic Avionics Technology Working Group (SATWG*) Charter

- Provide a forum to develop a Space Advanced Technology Plan involving avionics strategies and associated roadmaps
- Include both flight element and ground/flight infrastructure support activities
- Facilitate cooperative and collaborative relationships between NASA, government agencies, industry, academia, and professional societies
- Connect and network national pockets of technology and advanced development expertise and knowledge

* Established in 1990

Dr. Kenneth J. Cox
July 1993

94/1-10-0013



Benefits of STI

- The benefits of STI include:
 - Increasing interagency communication at all levels
 - Creating national technology cohesiveness
 - Sharing of facilities and expertise across agencies, industry, and educational institutions
 - Avoiding undesired duplication and reinventing through sharing of lessons learned
 - Developing cost-effective approaches through interdependent programs
 - Facilitating the identification of technology requirements for specific applications
 - Creating an environment to gain a substantial edge in international competitiveness

94/1-10-0030



The Vision and the Challenge for Space Technology Interdependency (STI)

- The Vision
 - Create and promote STI infrastructures to encourage and coordinate cooperative projects in research and technology for mutual benefit to organizations involved
- The Challenge
 - Develop management infrastructure which will provide motivation to personnel to implement and promote cooperative efforts
 - Hybrid approach – top-down and bottom-up
 - Incorporating measures of success

94/1-10-0012a



ECONOMICALLY IMPORTANT TECHNOLOGY-INTENSIVE INDUSTRIAL SECTORS*

- APPLIED MOLECULAR BIOLOGY
- DISTRIBUTED COMPUTING AND TELECOMMUNICATIONS
- ELECTRICITY SUPPLY AND DISTRIBUTION
- FLEXIBLE INTEGRATED MANUFACTURING
- MATERIALS SYNTHESIS AND PROCESSING
- MICROELECTRONICS AND OPTOELECTRONICS
- POLLUTION MINIMIZATION AND REMEDIATION
- SOFTWARE
- TRANSPORTATION

*2nd Biennial Report - National Critical Technologies
Panel, January 1993

Kumar Krishen, Ph.D.
NASA Johnson Space Center
Houston, Texas 77058

JSC Technology Utilization Office

Federal Technology

- Federal technology is developed at over 700 laboratories and centers.
- These laboratories are funded at about 25 billion dollars annually
- Over 100,000 scientists and engineers work at these sites in virtually every area of science and technology



JOHNSON SPACE CENTER

Major Federal T² Programs

- The President's Technology Initiative is a broad program involving many agencies and designed to stimulate the economy
- The Federal Laboratory Consortium (FLC) is composed of all federal laboratories and centers and is chartered to promote federal tech transfer
- The Technology Reinvestment Program (TRP) is a major (\$0.5 B) program chaired by DOD (ARPA)
- Many federal laboratories are developing their own strong tech transfer programs, especially DOD, SDIO, and DOE



JOHNSON SPACE CENTER

Accessing Federal Resources

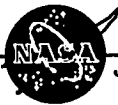
- Sharing information via visits, workshops, technical publications
- Exchanging personnel
- Finding technical assistance
- Using unique federal laboratory facilities & capabilities
- Licensing patents and technical Know-how
- Acquiring software
- Performing Cooperative R&D
- Cooperative Research & Development Agreements (CRADAs)
- forming consortia
- Using technology developed under government contract
- Working cooperatively with NASA (Unique among federal agencies)



JOHNSON SPACE CENTER

NASA Technology Transfer

- NASA Technology Transfer is conducted by the NASA Commercial Technology Network
- The Network offers the following services:
 - Industry-Led Partnerships
 - Dual-Development Technology
 - Regional Alliances
 - Small Business Development
 - Commercial Technology Acquisition
 - Fast-track Mechanisms



JOHNSON SPACE CENTER

NASA Commercial Technology Network

- For the purposes of this presentation, the Network will be discussed in 2 parts:
- First - The elements of the Network:
 - NASA Headquarters
 - NASA Field Centers
 - NASA Sponsored Organizations
- Second - Recent and on-going efforts to improve the Network and the Tech Transfer process



JOHNSON SPACE CENTER

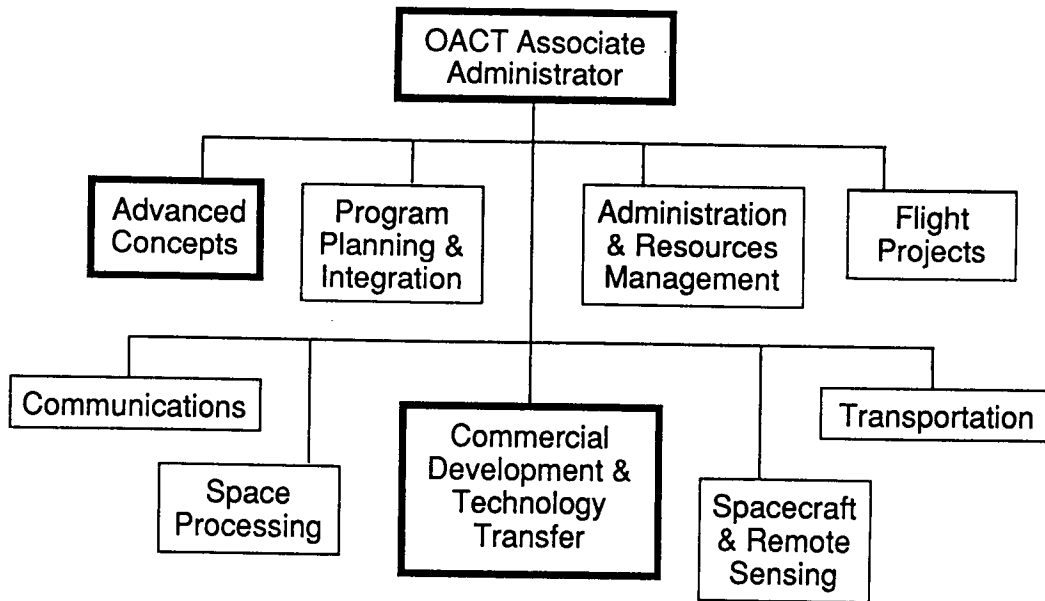
Headquarters T² Managers

- NASA Administrator - Daniel Goldin
- Office of Advanced Concepts and Technology (OACT or Code C)
 - Associate Administrator - Greg Reck
- OACT Advanced Concepts Division
 - Director - Robert Norwood
 - SBIR Program , NASA TRP participation, and new Industrial Technology Program
- OACT Commercial Development & Technology Transfer Division
 - Director - Greg Reck (acting)
 - NASA Technology Commercial Network and the Technology Utilization Program



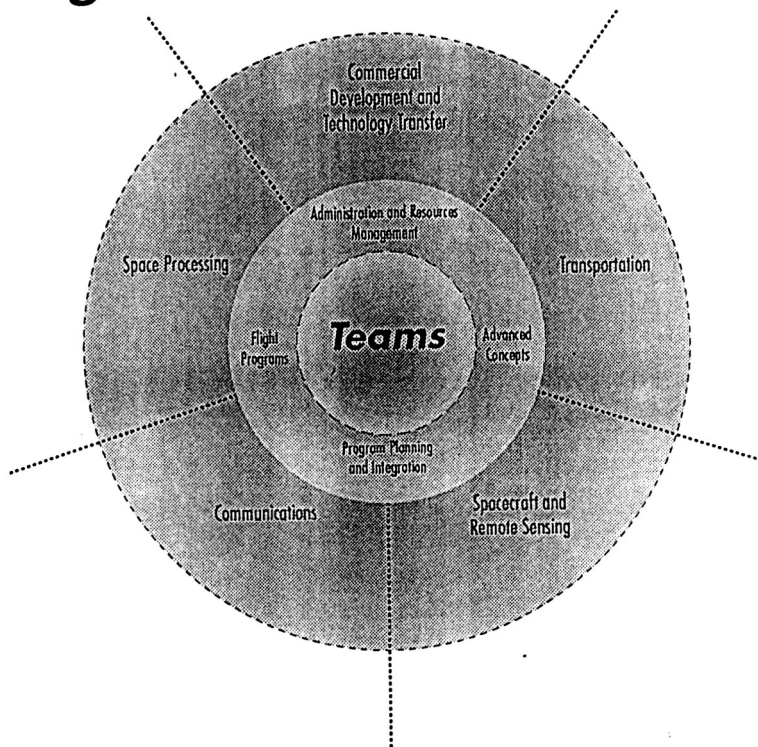
JOHNSON SPACE CENTER

OACT Organization



JOHNSON SPACE CENTER

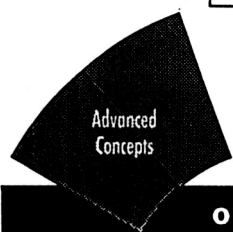
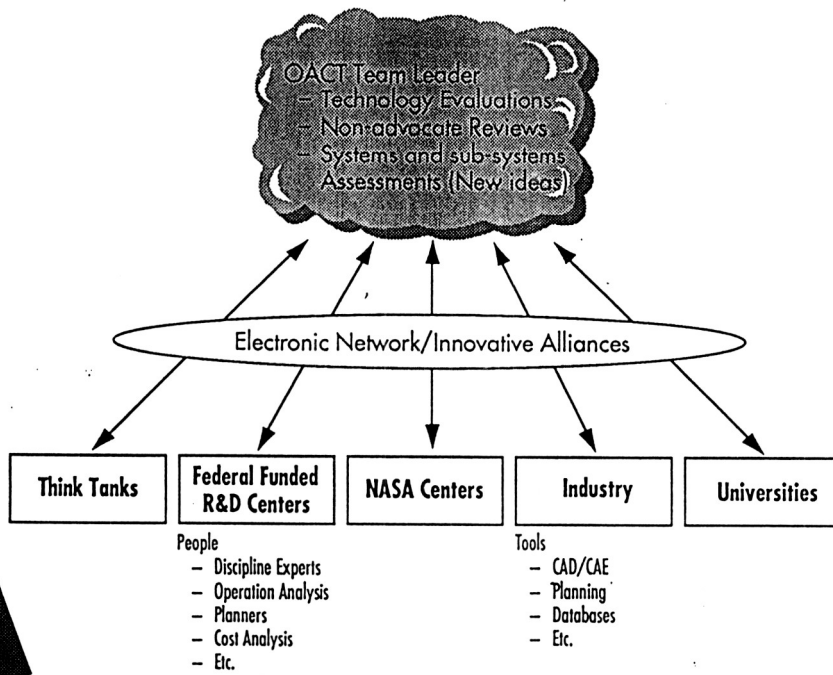
OACT Organization



OFFICE OF ADVANCED CONCEPTS AND TECHNOLOGY CA 6047-G 4/12/93

Advanced Concepts

A Borderless Organization



OFFICE OF ADVANCED CONCEPTS AND TECHNOLOGY CA 6047-D 4/12/93

NASA Field Centers

The NASA field centers and the network contacts are the following:

- | | |
|---|---------------------|
| <input type="checkbox"/> Ames Research Center | Syed Shariq |
| <input type="checkbox"/> Dryden | ... |
| <input type="checkbox"/> Goddard Space flight Center | George Alcorn |
| <input type="checkbox"/> Jet Propulsion Laboratory | William Spuck |
| <input type="checkbox"/> Johnson Space Center | William Milligan |
| <input type="checkbox"/> Kennedy Space Center | Walter Murphy |
| <input type="checkbox"/> Langley Research Center | Charlie Blankenship |
| <input type="checkbox"/> Lewis Research Center | Harvey Schwartz |
| <input type="checkbox"/> Marshall Space Flight Center | Gabriel Wallace |
| <input type="checkbox"/> Stennis Space Center | Jon Roth |



JOHNSON SPACE CENTER

Field Center Responsibilities

- Each center has an SBIR Office and a TU Office.
 - The office names and organizations may vary from center to center, but the functions are the same
- Small Business Innovation Research (SBIR)
 - The SBIR Office manager is responsible for managing the center's SBIR program
- Technology Utilization (TU) Office
 - The TU Office Officer or TUO is responsible for managing the center's TU program



JOHNSON SPACE CENTER

SBIR Manager Responsibilities

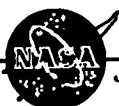
- The SBIR Office Managers are responsible for the implementation of the SBIR program at each field center which provides:
 - Phase I SBIR contracts for up to \$70K
 - Phase II SBIR contracts for up to \$600K
- This includes:
 - Development of technical interest in the form of subtopics to be included in the NASA solicitation document
 - Receipt, evaluation, and selection of proposals from the small business community
 - Procurement activity required to support development of the technology



JOHNSON SPACE CENTER

TUO Responsibilities

- Identifies innovations and technologies that have commercial potential
 - Operates the New Technology Report (NTR) system which identifies new technologies generated by the Civil Service and NASA contractors
- Promotes the transfer of these innovations and technologies to the private sector
 - Responds to requests for information
 - Publishes technology in NASA Tech Briefs and elsewhere
 - Supplies software for distribution by COSMIC
 - Develops commercial applications projects with partners in the private sector



JOHNSON SPACE CENTER

NASA Sponsored Organizations

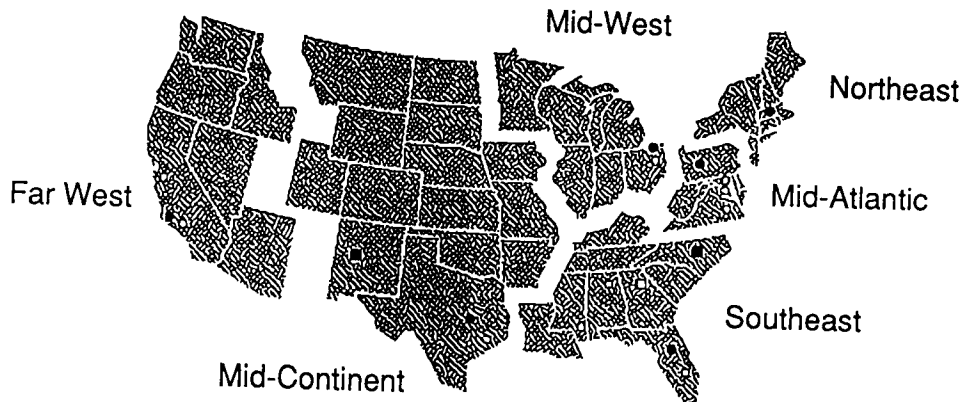
- 6 Regional Technology Transfer Centers (RTTCs)
- National Technology Transfer Center (NTTC) in Wheeling, WV
- Technology Applications Team at the Research Triangle Institute (RTI), Research Triangle Park, NC
- American Technology Initiative (AmTech) at ARC
- 11 Centers for Commercial Development of Space (CCDSs)
- Computer Software Management & Information Center (COSMIC) at the University of Georgia
- Technology Applications Center (TAC) in Albuquerque, NM
- 2 Technology Commercialization Centers (TCCs) at Ames and JSC
- Center for Aerospace Information (CASI) in Baltimore, MD
- The NASA Tech Briefs writers (ICT) and publishers (ABP) in NY.



JOHNSON SPACE CENTER

RTTC Regions

The NASA RTTCs are organized in common regions with the FLC. The RTTC system replaced the old NASA Industrial Application Center (IAC) system.



JOHNSON SPACE CENTER

Regional Technology Transfer Centers (RTTCs)

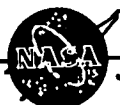
- The 6 Regional Technology Transfer Centers (RTTCs) provide information retrieval services and technical help to industrial and government Clients.
- The RTTCs transfer all federal technologies (not just NASA)
- The Mid-Continent Technology Transfer Center (MCTTC) is located at the Texas Engineering Experiment Station, Texas A&M University.
 - MCTTC Director at Texas A&M - Gary Sera
 - MCTTC Univ of Houston Clear Lake Office - David Wicks



JOHNSON SPACE CENTER

Technology Commercialization Centers (TCCs)

- The TCCs are an extension of the successful new business incubator program of the University of Texas (Austin), School of Business, Institute for Creative Capital (IC Squared).
- The TCC's purpose is to create and incubate new businesses based on NASA technologies.
- The first 2 TCCs are located at JSC and ARC
- The JSC TCC is located across from JSC and is often called the JTCC
 - JTCC Director - Jill Fabricant, Ph.D.



JOHNSON SPACE CENTER

Recent and On-Going Efforts

- The Creedon Report
 - One of the Red/Blue Teams
 - Generated numerous follow-on teams at Hq and Centers
- Administrator's Technology Transfer Directive
- Dual - Development
 - 10% - 20% of R&D to be used for partnerships with industry
 - Commercialization potential now required as part of technology project proposals (e.g. RTOPs)
- NASA Technology Commercialization Management Team



JOHNSON SPACE CENTER

The Creedon Report

In December, 1992, the NASA Administrator's NASA Institutional Team on Technology Transfer, chaired by Dr. Jeremiah Creedon of LaRC, issued its final report. It recommended, among other things, the creation of the following teams, all of which were formed and some of which are still meeting:

- Tech Brief Process
- Patent Applications and Licenses
- Execution of Secondary Targeted (Center lead teams)
- Software Distribution and Transfer
- Conversion of Non-targeted to Secondary Targeted
- Conversion of Primary to Secondary Targeted
- SBIR's
- CCDS's
- Employee Motivation /Incentives



JOHNSON SPACE CENTER



Dual-Development

- It is the Administration's policy that 10 percent to 20 percent of R&D funds should be used in partnerships with industry
 - The implementation details of this policy are in work
- Commercialization potential is now required as part of technology project proposals – for example the OACT RTOP now requires:
 - Customer
 - Industry involvement
 - Technology Transfer plans

94/1-10-0031.1



NASA Technology Commercialization Management Team

- This team is a major effort by OACT to review and improve the Network and the Technology Transfer process generally
- Greg Reck, the OACT Acting Associate Administrator, is leading the team
- The JSC representative is the Technology and Commercial Projects office

94/1-10-0031.2



Improving NASA's Technology for Space Science

- Committee on Space Science Technology Planning
- Aeronautics and Space Engineering Board Commission on Engineering and Technical Systems
- Space Studies Board Commission on Physical Sciences, Mathematics, and Applications
- National Research Council

National Academy Press
Washington, D. C. 1993

94/1-10-0014



RECOMMENDATIONS OF THE ASEB REPORT

- **Advanced propulsion**
 - Advanced Earth-to-orbit engines
 - Reusable cryogenic orbital transfer vehicles
 - High-performance orbital transfer systems for sending humans to Mars
 - New spacecraft propulsion systems for solar system exploration
- **Humans in space**
 - Radiation protection
 - Closed-cycle life support systems
 - Improved EVA equipment
 - Autonomous system and robotic augmentations for humans
 - Human factors research
- **Autonomous systems and robotics**
 - Lightweight, limber manipulators
 - Advanced sensing and control techniques
 - Teleoperators
- **Space power supplies**
 - 100 Kw nuclear power source

Kumar Krishen, Ph.D.
NASA Johnson Space Center
Houston, Texas 77058



SPACE TECHNOLOGY TO MEET FUTURE NEEDS

- **Materials and structures**
 - Advanced metallic materials based on alloy synthesis
 - "Hot" structures to counter reentry heating
 - "Trainable" control systems for large flexible structures
- **Information and control**
 - autonomous on-board computing system
 - High-speed, low-error rate digital transmission over long distances
 - Voice/video communications
 - Spaceborne tracking and data relay
 - Equipment monitoring technology
 - Ground data handling, storage, distribution and analysis
- **Advanced sensor technology**
 - Large aperture optical and quasi-optical systems
 - Detection devices and systems
 - Cryogenic systems
 - In-situ analysis and sample return
- **Supporting technologies**
 - Radiation insensitive computational systems
 - High-precision attitude sensors and axis transfer systems

Kumar Krishen, Ph.D.
 NASA Johnson Space Center
 Houston, Texas 77058



JSC Technology Development Activities

Emerging Technology/JSC Technologies

<u>Advanced Materials</u> <ul style="list-style-type: none"> • Ceramic composites • Space durable coatings • Entry thermal protection systems • Hypervelocity impact shield • Fracture mechanics analysis tools 	<u>Advanced Semiconductors</u> <ul style="list-style-type: none"> • Vertical stack devices • X-ray lithographic • Gallium arsenide • Etching • Ebeam, epitaxial bonding • Analog devices, DSP • Bio sensors 	<u>Digital Imaging</u> <ul style="list-style-type: none"> • Pattern recognition • Virtual reality • Video rate image warping • DSP • High definition television • Chromosome analysis • Digital image compression • Retinography • Electronic still camera
<u>Superconductors</u> <ul style="list-style-type: none"> • Antenna system applications • Magnetics systems applications 	<u>High Density Data Storage</u> <ul style="list-style-type: none"> • Ferroelectric • CD ROM • Gigabyte storage 	



JSC Technology Development Activities - Continued

Emerging Technology/JSC Technologies	
<u>Artificial Intelligence</u> <ul style="list-style-type: none"> • Expert systems • Fuzzy logic • Neural networks • Automation of operations • Automation of engineering analysis • Human-computer interaction 	<u>Flexible Computer-Integrated Manufacturing (robotics/teletrobotics)</u> <ul style="list-style-type: none"> • Telescience • Virtual reality • Remote operator interaction • Integrating SSF robots for assembly and maintenance • Dextrous end effector • Evaluating robotic technologies • Fault tolerant robots
<u>High Performance Computers</u> <ul style="list-style-type: none"> • Parallel processing • Neural networks 	<u>Sensor Technology</u> <ul style="list-style-type: none"> • Types/categories (infrared, nuclear magnetic resonance, magnetic, ultrasonic, temperature, chemometric) • Impedance imaging • Smart sensors • Plasma grid detector • Backwater Mossbauer spectrometer (BAMS) • Thermal analysis analyzer • Ultra sensitivity instrumental neutron activation analysis
	<u>Optoelectrics</u> <ul style="list-style-type: none"> • Spatial light modulators • Fast joint transformation correlators • Optimal filter theory

94/1-10-0018



JSC Technology Development Activities - Continued

Emerging Technology/JSC Technologies		
<u>Biotechnology</u> <ul style="list-style-type: none"> • Bioreactor • Tissue Equivalent Proportional Counter (TEPC) • Cell support systems • Understanding of osteoporosis, immunology, cell biochemistry 	<u>Medical Devices and Diagnosis</u> <ul style="list-style-type: none"> • Image warping and patterned sensors • Joint transformation correlation • Data reduction and analysis algorithms for sensors • Dried blood collection/analysis techniques • Behavioral performance devices/algorithms • Echocardiography • Electrical muscle stimulation • Robust exercise equipment • Medical equipment computer • Advanced life support (training, equipment, supplies, and layout) 	<u>Life Support</u> <ul style="list-style-type: none"> • Air revitalization • Water recovery • Food storage and preparation • Waste management and processing • Resource recovery • Thermal control • Long-duration testing (testbed) • Environmental quality monitoring technology • Closed environment requirements and microbiology and toxicology

94/1-10-0019



JSC Technology Development Activities - Concluded

Emerging Technology/JSC Technologies	
<u>EVA</u> <ul style="list-style-type: none">• Air revitalization• Thermal control• Power/batteries• Displays and control• Weight reduction• High mobility• In-suit monitoring• Mobility/ease of work• Protection from decompression sickness	<u>Human Factors</u> <ul style="list-style-type: none">• Anthropometry and biomechanics• Computerized human modeling (animation, lighting)• Human-computer interfaces• Virtual reality for EVA• Personal emergency provisions• Tools and diagnostic equipment• Escape systems• Crew health care systems<ul style="list-style-type: none">- Countermeasures• Display panels• Computer input devices• Instrument noise abatement• Instrument vibration isolation• Performance evaluation• User-friendly science monitoring workstations• Crew systems

94/1-10-0020



Discoveries and Innovations at NASA Johnson Space Center

Discoveries in Life Sciences

- Change in central venous pressure in space flight is related to space motion sickness.
- Simulated microgravity bed rest reduces bubble nucleation and this reduction may reduce the incidence of decompression sickness in microgravity.
- Man can survive and function well in the microgravity space environment leading to progressively extended flight durations and crew responsibilities, and turning attention to those physiologic responses to microgravity or return to Earth gravity that require countermeasure development.

94/1-10-0021



Discoveries and Innovations at NASA Johnson Space Center

Discoveries in Life Sciences - Concluded

- **Development of the OTOLITH Tile Translation Reinterpretation, the hypothesis leading to understanding of underlying causes of space motion sickness and to the development of preflight adaptation training.**
- **The progressive loss of bone mass during extended missions through bone density studies during Skylab and in cooperation with the Russians during their flights. This discovery has led to pursuit of exercise and other countermeasures.**

94/1-10-0022



Discoveries and Innovations at NASA Johnson Space Center

Discoveries in Lunar Science

- **The Moon is not a primordial object; it is an evolved terrestrial planet with internal zoning similar to that of Earth.**
- **The Moon is ancient and still preserves an early history (the first billion years) that must be common to all terrestrial planets.**
- **The youngest Moon rocks are virtually as old as the oldest Earth rocks. The earliest processes and events that probably affected both planetary bodies can now only be found on the Moon.**
- **The Moon and Earth are genetically related and formed from different proportions of a common reservoir of materials.**
- **The Moon is lifeless; it contains no living organisms, fossils, or native organic compounds.**

94/1-10-0023



Discoveries and Innovations at NASA Johnson Space Center

Discoveries in Lunar Science - Concluded

- All Moon rocks originated through high-temperature processes with little or no involvement with water. They are roughly divisible into three types: basalts, anorthosites, and breccias.
- Early in its history, the Moon was melted to great depths to form a "magma ocean." The lunar highlands contain the remnants of early, low-density rocks that floated to the surface of the magma ocean.
- The lunar magma ocean was followed by a series of huge asteroid impacts that created basins which were later filled by lava flows.
- The Moon is slightly asymmetrical in bulk form, possibly as a consequence of its evolution under Earth's gravitational influence. Its crust is thicker on the far side, while most volcanic basin – and unusual mass concentrations – occur on the near side.

94/1-10-0024



Discoveries and Innovations at NASA Johnson Space Center

Innovations

- Blunt body design to overcome reentry heating
- Couch body support to enhance crew tolerance to high "G: loads
- Concentric rendezvous technique
- Liquid-cooled garment
- Fly-by-wire control for spacecraft
- Navigation system for space-based establishment of a state vector
- Very lightweight, reusable thermal protection system
- Hydrogen/oxygen Bacon fuel cell as a major energy source
- NASA Structural Analysis System (NASTRAN)
- C-Language Integrated Production Systems (CLIPS)
- Intelligent Computer-Aided Training (ICAT)
- Programmable Image Remapper
- Magnetic End Effector
- A new dried-blood technique for storing samples in space

94/1-10-0025



Discoveries and Innovations at NASA Johnson Space Center

Innovations - Continued

- Optimal filter theory
- Hybrid vision's eye-tracking
- Network and Execution and Training Simulator (NETS)
- Radio frequency hyperthermia system for cancer treatment
- NASA/Baylor Axial Flow Pump for ventricular support
- Rotating Bioreactor Cell Culture apparatus
- C-Language Integrated Production System (CLIPS) - Version 6.0
- Intelligent Computer-Aided Training (ICAT)
- Real Time Data System (RTDS)
- Task Analysis Rep Generation Tool (TARGET)
- Beta cloth/fabric structures
- Targeting and Reflective Alignment Concept (TRAC)
- Computer-Assisted Schedule System (COMPASS)

94/1-10-0026



Discoveries and Innovations at NASA Johnson Space Center

Innovations - Concluded

- Cool suit/liquid-cooled garment
- Physiological monitoring instrumentation
- Improved firefighter's breathing system
- Earth observation technology
- First U. S. manned spacecraft (Mercury)
- First U. S. manned two-manned spacecraft and space walk (Gemini)
- First U. S. manned lunar spacecraft and lunar landing (Apollo)
- First reusable manned spacecraft (i. e., Shuttle)

94/1-10-0027



NASA Technology Areas of Commercialization

- Health and Medicine
- Training and Education
- Communications and Information Systems
- Earth Environment and Resource Management
- Industrial Productivity
- Public Safety and Home Systems
- Recreation and Quality of Life
- Transportation

94/1-10-0029



Dual-Use Space Technology Transfer Conference and Exhibition

Sponsors

NASA/Johnson Space Center (JSC)
Clear Lake Area Economic Development Foundation
Clear Lake Council of Technical Societies
IEEE – Galveston Bay Section
Houston Advanced Research Center
AIAA – Houston Section
Texas Innovation Network
Texas Space Commission
Texas Department of Commerce
Mid-Continent Technology Transfer Center
JSC Technology Commercialization Center
University of Houston - Clear Lake
University of Houston, Texas Product Development Center
Greater Houston Partnership
Engineer's Council of Houston

94/1-10-0031a

KEYNOTE ADDRESS

Dual-Use Space Technology Transfer: Challenges and Opportunities for Commercialization

by

George Kozmetsky

It is a pleasure to participate in the NASA/Johnson Space Center "Dual-Use Space Technology Transfer Conference." It is a special honor and privilege to be invited as one of the keynote speakers. I also wish to thank each of the co-sponsors for their active participation in this conference and recognize their day-to-day operations in NASA technology transfer. The organizations listed in your program truly represent an effective collaboration between NASA/Johnson, the higher educational units in Houston, the relevant State of Texas departments and commissions, the Houston-Clear Lake economic entities that comprise the Clear Lake communities commercialization networks, the Houston-based scientific and technical society chapters, and NASA's unique technology transfer organizations. In fact, these institutions and entities make up NASA/Johnson's regional Commercial Technology Network. In my opinion, this regional network is truly the beginnings of a virtual organization for dual-use space technology transfer.

Few federal, state, academic, and research entities have established such a regional virtual network. The future challenges and opportunities for the NASA/Johnson virtual network was recently stated by Daniel Goldin, NASA Administrator, as follows:

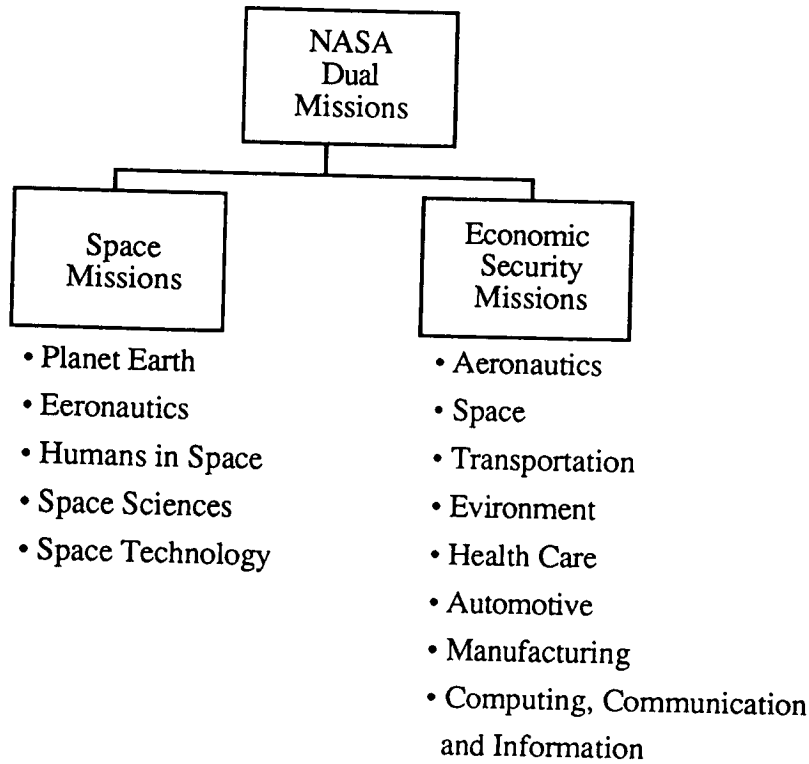
A significant part of this [network] will result from fostering technology developments in the general economy that serve our [NASA] needs as well as the needs of the nation. Technology transfer is a fundamental mission; it is as important as any NASA mission, and it must be pursued.

You will note that Goldin has broadened the scope of NASA's technology transfer precepts. In the past, technology transfer was too often focused on NASA spin-offs and NASA-generated knowledge transfer. Goldin has raised commercialization of technology to a mission status—a level above what is normally thought of as dual-use technologies. The commercialization mission concurrently with NASA's basic missions make way for dual technology transfer for comprehensive security. NASA can pioneer comprehensive security by successful accomplishments its basic space and aeronautical missions as well as contribute directly to missions that contribute to America's economic strength.

Under a cold-war society, national security was predominantly concerned with Federal laboratories' research and developments for security purposes. Many of NASA's past programs were concerned with staying ahead of the USSR space efforts. NASA was, in many respects, involved in a competitive space race. The transfer of R & D to the private commercial sector under these conditions was predominantly a "trickle-out" process. The overarching goals for U.S. federal research and development today must be one of "shared prosperity" at home and abroad. Shared prosperity in turn requires Goldin's concept of technology transfer as a fundamental mission; e.g., economic security. Figure 1 sets forth a concept for NASA's concurrent dual missions for which they can be held accountable.

Concurrent dual missions will lead NASA to a new way of doing business. The commercialization mission will require that each NASA program be carried out in ways that proactively seek private sector alliances or working combinations from the onset to develop new ways of doing business that both meet their respective missions, program goals, schedules, and budgets, as well as maximize specific commercial projects and required transfers. In short, NASA has many technologies that can contribute to its own technology transfer missions that meet the Clinton Administration national priorities, such as health care, clean cars, environmental monitoring, affordable housing, telecommunications and large-scale computer networks, and the knowledge highway.

Figure 1. NASA Concurrent Dual Mission Roles

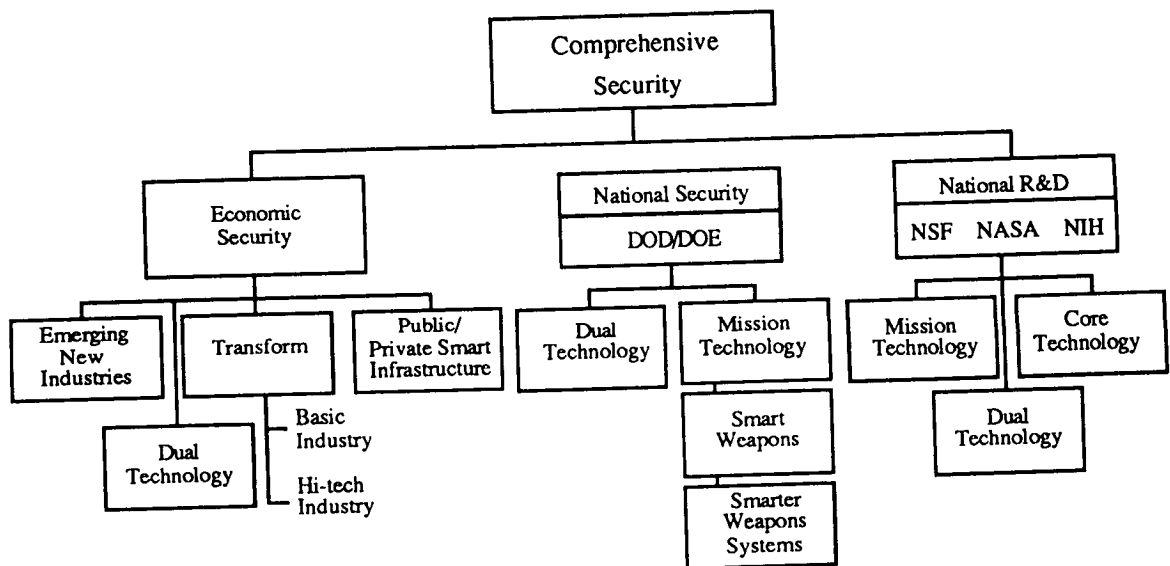


Dual mission technology commercialization, in my opinion, is the emerging role of all federal agencies for overall comprehensive security. See Figure 2. While national security missions are a continuing and abiding societal need, federal government agencies and laboratories have to link and leverage these missions to those required for economic security.

It is indeed a revolutionary idea that dual missions can result in technology transfer missions that first develop projects for domestic manufacturing and global distribution, and then used for security purposes. Such path-breaking and infrastructural technologies can be the core of successful technology commercialization for NASA centers and other federal

laboratories. U.S. security will depend on how federal R & D is transferred creatively and innovatively into jobs, exports, and productivity.

Figure 2. Comprehensive Security through an Emphasis on Dual Missions



The remainder of this paper is in three parts:

1. Background for Understanding Dual Missions
2. Definition of Technology Transfer
3. IC² Institute Johnson Technology Commercialization Center

Background for Understanding Dual Missions: The Case for Change

The past cold-war period is now over four years old. The past cold-war national policies, however, are still in the process of public debate for more formulations and specific programming. Nevertheless there are clear trends that involve how Federal laboratories, such as NASA, must change their way of doing business. These trends are:

1. Changing global environment;
2. National economy and dual deficits;
3. Newer national agenda;
4. Emerging fourth industrial revolution technologies; and
5. Directions for policy change.

I would like to briefly comment on each of these trends and their case for changing NASA into a dual mission-oriented center.

Changing Global Environment

The end of the cold war has changed our nation's economic structure. The change is from an emphasis on global security to global economic competition. Economic competition is going to depend more on our economic leadership than on the U.S.'s superpower military position. Currently there is a major restructuring and downsizing of defense, aerospace, and space industries as well as the Fortune 500 companies. The restructuring will involve downsizing of the major prime contractors as well as the various

tier suppliers. There needs to be a rapid redeployment of skilled and unskilled labor to peace-time, globally competitive economic pursuits.

For more than forty years, the U.S. science and technology policies were primarily focused on competition with its enemies in the Eastern European bloc. The current issues are those concerned with the viability of the U.S.'s current economic policies as well as the ability of U.S. industry and the effectiveness of the firms to compete technologically with its cold-war allies. From a NASA perspective, commercial space launch competitors are in France, Russia, and China. Japan is an emerging competitor. In aircraft manufacturing, competition is coming from not only Europe and Brazil but also from joint ventures between China and South Korea and alliances between U.S. aircraft manufacturers with Japanese firms. The fields of high speed computers, graphics, flat panels, and high-definition t.v. are also internationally competitive. The domestic knowledge highways need to be extended into global knowledge highways.

The days when the U.S. could depend on its peace-time economy based on a higher level of real income and a rapidly expanding domestic market are at best questionable if not over. The days when a few industrial sectors based on advanced technologies that could sustain a nation's general economy and provide a global economic prosperity are also ending.

National Economy and Dual Deficits

The U.S. is in a period of dual deficits. Deficit in the federal budget as well as in trade. Federal budget deficits cannot continue to become our growing way of life. Federal deficits in the decade of the 1950s were \$15 billion; in the 1960s, they were over \$63

billion; and in the 1970s, \$420 billion. In the 1980s, they were over a trillion dollars; so that today, the total Federal budget deficit is over \$4 trillion.

In the 1990s, the U.S. is faced with the problem of what to do with 2.6 million civilians who worked on cold war products and services. The problem is more compounded because the domestic non-defense industries are also resizing. It is difficult, even for Americans let alone other foreign nations, to comprehend that most of the U.S. employment in the past two decades was provided by small businesses with less than 500 employees. In the past, such firms were primarily suppliers for the Fortune 500-100 firms. It is still more astonishing to believe that women entrepreneurs have provided more jobs in the past decade than all of today's Fortune 500 firms. The 1990s is a period in which our domestic corporations resize, foreign companies globalize, military bases close, and post-cold war defense conversion takes place.

What is facing U.S. Congress and the President is a need for science and technology policies that could arrest the de-industrialization process of the last four decades from transfer of work abroad, growing dependency on imports of high tech industrial goods, continuing trend toward underdevelopment and undercapitalization of U.S. industries that are based on cold-war policies.

Mandatory spending has risen significantly since the mid-1960s. In the 1990s and subsequent period of growth of the U.S. economy has not provided sufficient tax funds for mandatory spending as well as maintaining defense and space budgets. Tax increases in 1993 have been made for non-aerospace and defense economic developments. The health care reform costs in 1994 will further impinge on defense and space expenditures.

Recent projections for Federal research development funding, done by the JFK School of Government at Harvard University, shows a drop from \$43.5 billion in 1992 to \$35 billion in 2000. National Lab portions would drop from \$13.5 billion in 1992 to \$8 billion in 2000. The result is that for the past year, OMB and OSTP have been pushing

DOD to move in the direction of considering partnerships with industry that would match their R & D budget by at least 20 percent and increasing to 25 percent. NASA has been asked to match their R& D by 10 to 20 percent with industry partnerships. There is no question that innovative R & D private industry collaborations and other commercialization revenue mechanisms are a major means for NASA centers to deal with downsizing of the Federal R & D budgets. These other means can be CRADA, work for other Federal agencies, income-generating measures such as licensing royalties, and joint R & D developments between NASA and the private sector.

Each of you is aware of the U.S. trade balance-of-payment problems. The continuing negative trade balance of payments since 1982 must be reversed. Our trade deficits can be attributed to higher rates of productivity increases in Europe and Japan than in the past.

The trade deficits are, to a large extent, in U.S. basic industries; namely, automobiles, apparel, telecommunications and sound reproduction equipment, metal manufacturers, iron and steel, footwear, and miscellaneous electrical machines and components. U.S. favorable trade balances are in aircraft and other transportation, plastic, scientific instruments, organic chemicals, military goods, office equipment and computers, and drugs and medicines.

What the trade figures disclose, in my opinion, is that the U.S. basic and high-tech industries have lost their classical advantage in adapting newer technologies for their economic renewal and for starting competitive, emerging industries. In this sense, we are losing the race in commercializing innovations generated in both the private and public sectors.

The trade race in the twenty-first century will be based on technology commercialization. First-to-market is important. Most of our industries are caught up in an apparent endless competition that requires each successive generation of products and

services represent a marked economic improvement on its predecessor or meet a newer customer requirement in the private or public market. For NASA, dual technology missions or uses means that R & D is intimately linked to dual markets. In addition, the development process must yield an economically viable product that can be manufactured competitively for a global marketplace. Research and development can no longer be just a patent game or one that places emphasis on Nobel prizes or citations in learned journals. The goals of R & D must be to generate successful market results under global competition that provides favorable trade balances, jobs, and justifies returns on investments.

New National Agenda

The post-cold war priorities have changed the national agenda to global economic competitiveness and job creation. This means using commercial R & D as a key instrument for the national agenda. Quality of the environment is also a priority, not only in terms of R & D but also in terms of rapidly creating a globally competitive domestic industry. Clean-ups of federal labs are a basis for emerging markets. Much of NASA technology as well as research facilities can be utilized for clean air, water, and pollution monitoring and control technology transfer.

Health care is another priority on Clinton's national agenda. Commercialization of the Johnson Space Center's bio-medicine research can play a major role in both better health care as well as job creation.

Education and training is a growing national priority. NASA can play an important role in computerized education and training as well as a newer delivery system.

The information or knowledge super-highway is also a national priority. NASA centers have been pioneers in utilizing such networks for research and knowledge exchange.

Technology Trends

The current changes in technology are dramatically different from those of previous periods. Business and economic historians refer to these as industrial revolutions. See Table 1.

Table 1. The Four Industrial Revolutions

Industrial Revolution	Key Technologies	Managerial Issues
First	Textile, iron from coke steam engine	One man or one family affair
Second	Railroads and steel making	Exploitation of scale and investments in production, marketing, and management
Third	Electricity, batch elements and combustion engine	Exploitation of R&D, mass production, and scope (more than one product)
Fourth	Microelectronics, biotechnology, genetic engineering, new materials robots, waste management technology	Global markets, niche markets, flexible manufacturing, and linking of science, innovation and management for rapid change

U.S. economic security will depend, to a large degree, on the fourth industrial revolution technologies. These are the technologies that interest NASA. The fourth industrial revolution technologies have special characteristics; namely,

- They are closely linked to the basic sciences that are in a state of a fast moving revolution. How closely the scientist, engineer, business person, and worker operate in partnership will determine the pace of economic security.
- The technologies will transform every sector of the economy; e.g., basic and high technology industries, agriculture, forestry, and services including education, medicine, and finance.
- The new technologies can be implemented by emerging nations as rapidly as highly developed nations.
- Each technology of the fourth industrial revolution is so diversified it is highly unlikely that any one nation will establish an unambiguous leadership.
- These technologies will change the nature of the work place, obsoleting old skills, and requiring new capabilities in those affected by them.

As a result, the fourth industrial technologies require creating new types of institutional alliances among academia, business, and government. Ownership issues are becoming more important. They will require greater intellectual property protection. NASA will need to develop its intellectual property assets and improve their means of transfer for commercialization that result in royalty and license fees.

The impact of the fourth industrial revolution on the technology transfer process is that it must be more closely allied with manufacturing, marketing and distribution, maintenance and service, and financing than ever before. The alliance must be more global

in context than before if traditional domestic investments are to be met. The line between manufacturing, service and financing are breaking down. Consumers are demanding more complete packages of hardware, software and the means to implement successful innovations without building larger and larger in-house service staff. Today, technology-based products are no longer simply a single product like a PC computer. PC computer manufacturers must provide services that end up in economically viable innovations that solve a company's problem in less time over a rapid pay-off period.

In summary, fourth industrial revolution product share the following characteristics:

1. Products are not the results from one organization; they are the results of an alliance of organizations or an enterprise network.
2. Small companies are becoming more important for job creation, export growth, as well as developing winning combinations with the third Industrial Revolution large companies, and developing growth companies.
3. Newer ways of doing business will be required. Among these are alliances, collaboration, and joint ventures.
4. Newer information infrastructures or knowledge highways for electronic data exchange, electronic funds transfer and investments, electronic mail, enterprise networks for joint design, and electronic marketplaces are needed.

Direction of Policy Changes

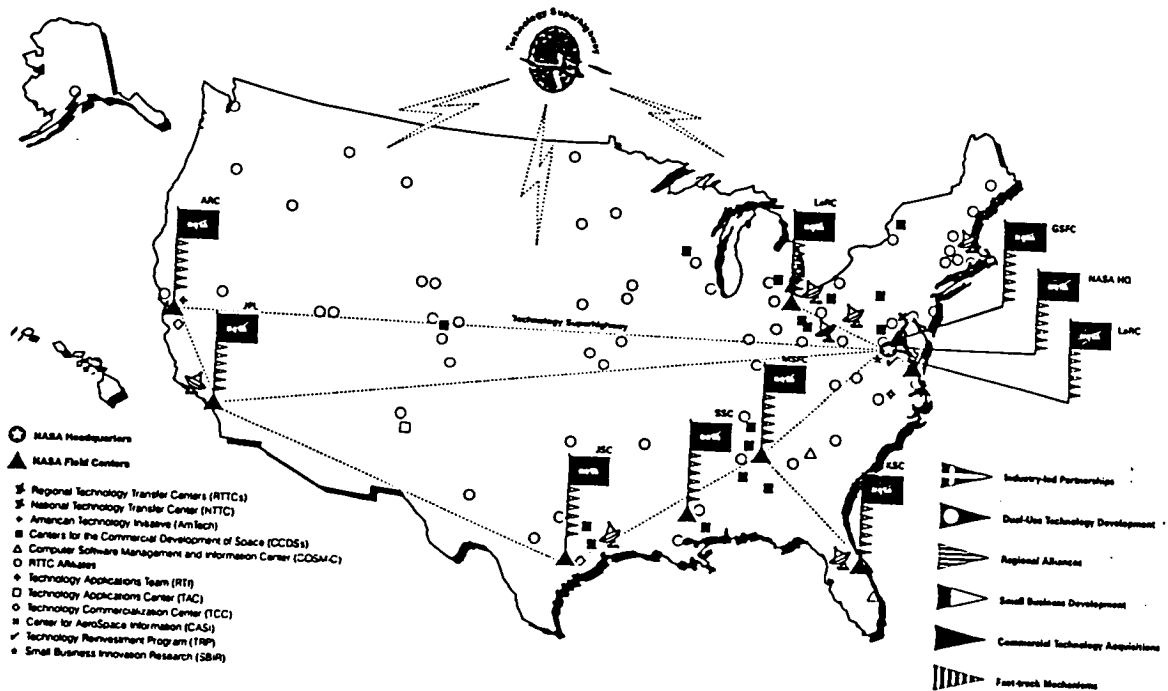
The direction of policy change can be enumerated as follows:

- Shifting priorities to develop enhanced industrial competitiveness through collaborations with private industry. The federal R & D funding will have priority when it is done in partnership with private firms.
- More emphasis on helping individual firms to obtain, adapt, and use technology for the marketplace.
- Civil technology and missions will require more and more federal laboratory alliances with private firms.
- Federal laboratories must become more dual missioned or face downsizing.

In summary, these trends point out that NASA's attitudes must change more than its technology transfer processes. NASA's commercial technology network needs to become better linked as well as more proactive in new ways of doing business to implement its dual missions. Currently there is a need to strengthen and focus virtual networks that tie NASA headquarters and centers through industry-led partnerships, regional alliances, small business creation, dual-use technology development, fast-track mechanisms, and commercial technology acquisitions. Figure 3 shows the existing NASA Commercial Technology Network. The challenge is how to more effectively link and exercise NASA's Regional Technology Centers (RTTC) and their affiliates, National Technology Transfer Center (NTTC), American Technology Initiative (AmTech), Centers for Commercial Development of Space (CCDS), Computer Software Management & Information Center (COSMIC), Technology Application Team (RTI), Technology Application Center (TAC), Technology Commercialization Center (TCC), Center for Aero-Space Information (CASI), Technology Reinvestment Program (TRP), and Small Business

Innovation Research (SBIR) to meet changing policy needs. This will require reexamining NASA technology transfer practices and developing ways to measure their effectiveness.

Figure 3. NASA Commercial Technology Network



Technology Transfer for Dual Missions

Dual-mission or dual-technology usage requires technology transfer. In other words, NASA's space missions require transfers to and from NASA's civil missions as well as to and from other commercially developed technologies. NASA's technology transfer missions require transfer to and from NASA space missions as well as to and from other commercially developed technologies. Technology transfers from the laboratory to

commercialization has been a continuing challenge in both the private, public, and academic sectors. Criticism focus on return on investment and timeliness. There have been few successful examples or case studies. Ultimately, technology transfers are being judged in terms of applications profitably used in the global marketplace. That is how many believe the NASA transfer process will be judged even though it is a new mandate.

In their forthcoming book titled *R&D Collaboration on Trial*, IC² Institute Fellows Gibson and Rogers set forth that there are three kinds of technology transfer that directly influence U.S. economic security. One is transferring out technologies into start-up companies. The other is transferring new technologies to established firms. The third is transferring technologies between different research organizations. See Figure 4.

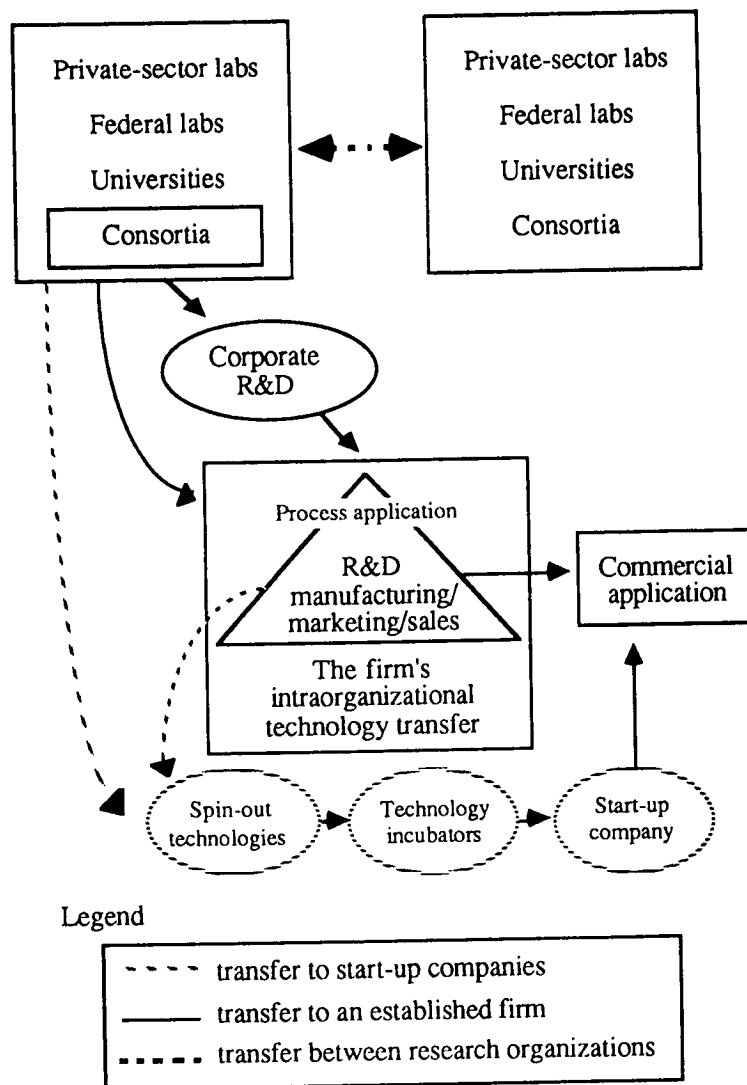
Gibson and Rogers have identified four levels of technology transfer. These levels can be used for understanding technology commercialization in a global context. See Figure 5.

Level I success is measured by the quantity and quality of research. Too often this is measured in terms of number of research reports and journal articles. Technology transfer plans and processes are not considered very important. *Research strength* and reputation has been considered most important. Traditionally, technology users have not been involved at this level of the transfer process. The belief is that good ideas sell themselves; pressures of the marketplace are all that is needed to drive technology use and commercialization. The "trickle out" method of technology transfer in the past was believed to be sufficient to sustain global competitiveness. Level I transfer processes will not provide NASA with the required successes it needs.

Level II transfer, technology acceptance, calls for the beginnings of shared responsibility between technology developers and users. Success occurs when a technology is transferred across personal, functional, or organizational boundaries, and it is accepted and understood by designated users. A Level II perspective encourages the belief

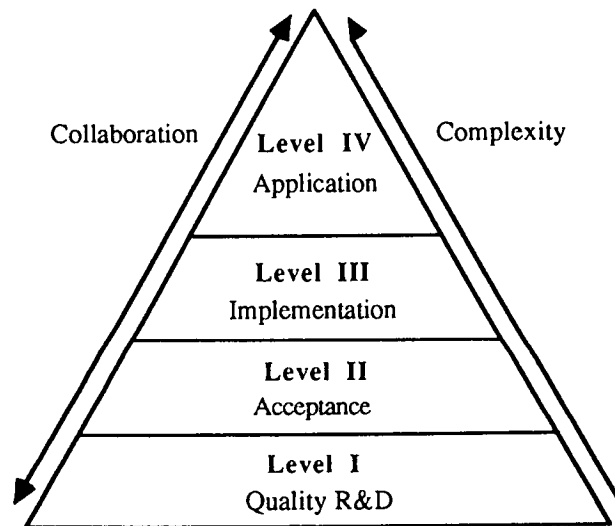
that successful technology transfer is simply a matter of getting the right information to the right people at the right time. Technology acceptance will require NASA to strengthen its virtual network if they are to implement successfully dual-mission transfer policies.

Figure 4. Three Basic Forms of Technology Transfer to Commercial Applications



In Level III transfer, technology implementation, success is marked by the timely and efficient employment of the technology. For Level III success to occur, technology users must have the knowledge and resources needed to implement, or "beta" test, the technology. Technology implementation can occur within the user organization in terms of manufacturing or other processes, or it can occur in terms of product development, such as building a prototype or proof of concept for commercial application. *Industrial strength* is required. It is at this stage where the user organization provides value added to the transferred technology. NASA must strengthen its organizational structure to provide for value-added transfer to both established firms and start-ups.

Figure 5. Technology Transfer At Four Levels Of Involvement



Level IV transfer, technology application, centers on product commercialization. Level IV builds cumulatively on the successes achieved in attaining the objectives of the three previous stages; but *market strength* is required. Feedback from technology users

drives the transfer process. Success is measured in terms of ROI or market share. Industry partners need to be identified early as well as work in concert with NASA.

Table 2 shows that moving from Level I to Level IV is *not* a linear, step-by-step process. Multidimensional collaboration is required. Complexity increases significantly as the technology, and perhaps the technology developers, move from Level I to Level IV. The success factors are also different at each level. Technology transfer needs to have an infrastructure developed as part of the transfer process. That is the major technology transfer gap for NASA as I see it.

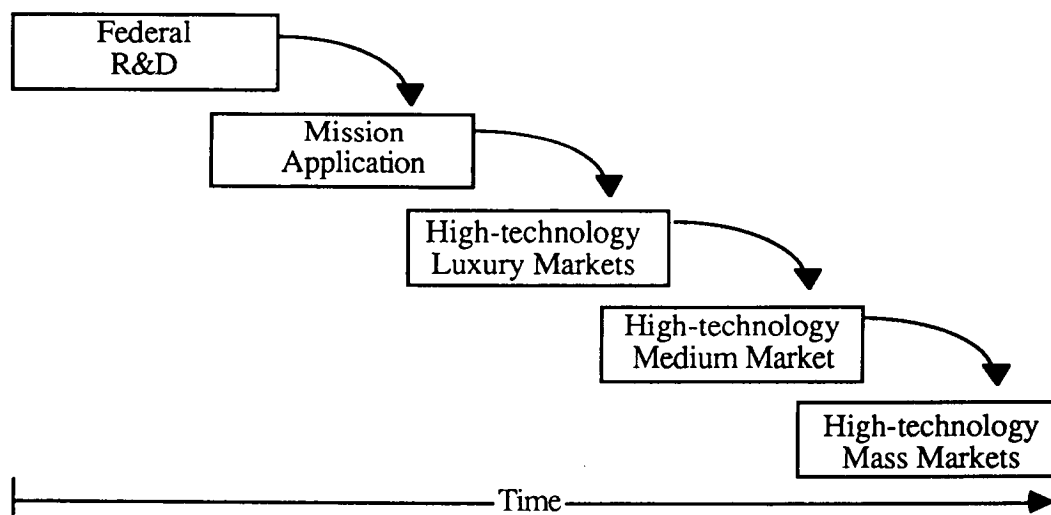
Table 2.
Technology Transfer Levels Collaboration and Complexity

	Level	Collaboration	Complexity
I	Quality of R&D	Technology users not involved	Trickle out over time
II	Acceptance	Shared responsibility between technology developers and users	Right information to the right people at right time
III	Implementation	Technology users must have the knowledge and resources needed to implement or test the technology	Industrial strength and value-added ability
IV	Application	Product commercialization	Market strength is required

The days when technology transfer from Federal Labs could concentrate on one market—namely, the government—are over. Nor can we rely on technology transfer to be a trickle-down transfer policy to position U.S. economic security. Let me elaborate a little

on the trickle-down policy. As shown in Figure 6, the Federal R & D trickled from mission applications to the technology luxury market applications, then to the technology medium-margin market applications, and finally to the technology low-margin, mass-production market. Economic security requires that the traditional process be changed to one which is more concurrent managed process. The concurrent process will require NASA to change its management styles. Under the traditional model, the return on investment was assumed to be adequate after mission applications. The newer trends make it imperative that NASA go further in justifying the ROI. This will require the generation of more jobs; more taxes—local, state, and federal; higher profits, and enhancement of U.S. competitiveness. The metrics of success will no longer be only measured by successful space mission accomplishments. Those who accept the challenge of concurrence will be the leaders for the new paradigm of dual missions.

Figure 6. Traditional Federal R&D Trickle Down Process



The IC² Institute Johnson Space Center TCC

The IC² Institute's Johnson Space Center TCC experiment was activated in February 1993. It is an experiment that deals with NASA technology transfer to start-up companies and their licensing to established firms. It is also an early stage concurrency experiment. Four milestone achievements were required for the first year of operations.

These were:

- Selection of qualified personnel on-site director, technology specialist, and marketing managers. Over 600 people were screened and interviewed for these positions.
- Selection of office space within two miles of the Johnson Space Center.
- Starting a "know-how" network within the Houston-Clear Lake region.
- Getting new companies started into the Johnson TTC incubator as well as searching for technologies that could be commercialized.
- Granting licenses to established companies.

All four milestones were achieved within the first nine months. Figure 7 summarizes the early results of the first four months of actual operations.

The IC² commercialization process developed consists of three phases:

Phase I	Technology Identification and Definition
Phase II	Opportunity Analyses
Phase III	Execution (actual linkage or tenant acceptance)

Figure 7. NASA Johnson TCC Status Review

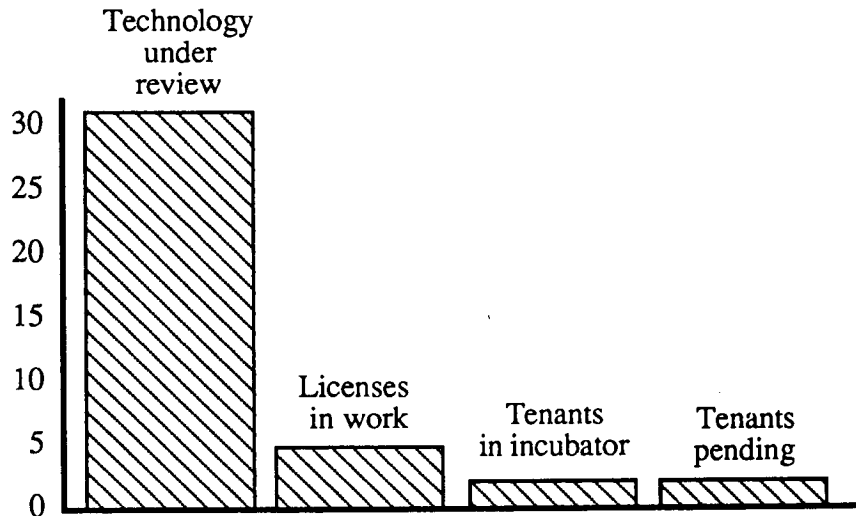
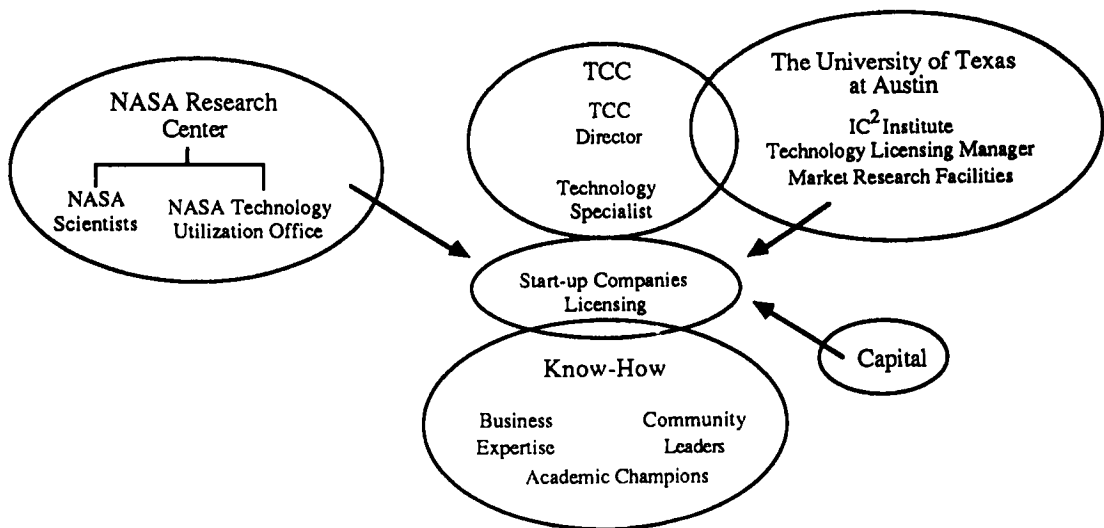


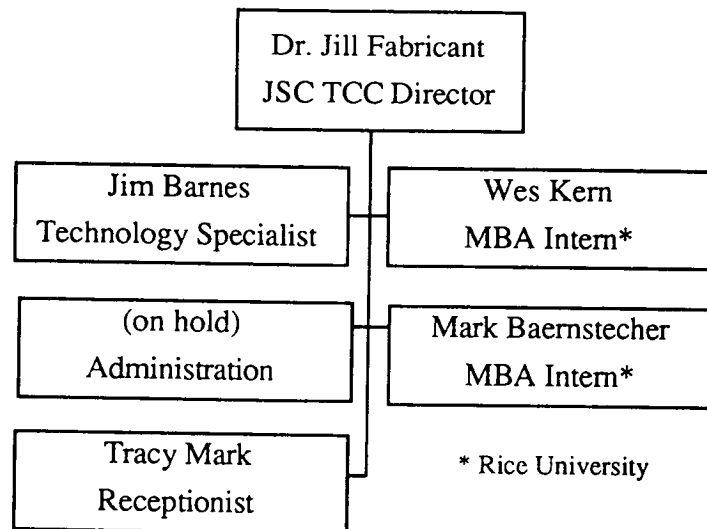
Figure 8 shows the roles of the NASA Johnson Research Center, the Johnson Technology Commercialization Center, and the IC² Institute.

Figure 8. Roles in Commercialization



The Johnson TCC is located three blocks from the main gate of JSC. It occupies about 7,800 square feet. Its current personnel are shown in Figure 9. The first two tenants are AdvanTex. Inc., whose graphics software products are based on three NASA software development tools. The Tenagra Corporation provides information-technology related software products. It already has a \$120,000 contract for commercial spacecraft mission control, developing operating reports and procedures for the control center. The NASA technology transferred was the NELS or NASA Electronic Library System.

Figure 9. TSC-TCC Organization



Both companies have a total of three employees. The primary source of funding was provided by the respective founders. Tenagra is expecting additional funds from an SBIR proposal to Hermann Hospital in Houston.

The other two tentative tenants have business plans in final review. Control Technologies, Inc., is licensing seven distinct software technologies. Laser Professor,

Inc., is licensing NASA software developed to train NASA astronauts, flight controllers, and engineers. Laser Professor has entered into an alliance with Texas Learning Technology Group, a not-for-profit consortium of school districts.

The technologies under execution for licensing is zeoionic soil. Currently, a group of UT-Austin Graduate School of Business students in the MOOT CORP® have prepared a business plan and raising are financing to start Zeoionic Inc.

There are nine technologies in the opportunity phase. These are Hypermen, shelf stable tortillas, heart pump, ice detection system for aircraft, shock absorbers, chemical stripper for anodized aluminum, multi-phase flowmeter, dried blood plasma storage, and measurement of urokinase (cancer cells).

There are eight technologies entering the Phase I identification stage. As would be expected, two technologies were found not ready for commercialization.

The Johnson commercialization project with SEMATECH for semiconductor etching was transferred by SEMATECH to AT&T, Hughes Research Labs, and Los Alamos Labs.

Table 3 shows the licensing effort of JSC-TCC by technology commercialization phases.

**Table 3. Summary of Technologies Evaluated
for Commercialization Growth***

	Total	Food	Software	Medical	Electronic Products	Materials	Mechanical Products	Instruments
Phase I								
Technology Identification	8			1	2	1	2	2
Phase II								
Opportunity Analysis	9	1	1	3		1	1	2
Phase III								
Execution	1					1		
Not Ready for Commercialization	2			2				
Total	20	1	1	6	2	3	3	4

*See Appendix for more details.

Conclusion

The IC² Institute JSC TTC experiment is off to a good beginning. In terms of metrics, there is little doubt that within the first 12-18 months the cash cost of a direct job will be about under the cost of one temporary job funded by the federal government. More specifically about \$20,000 for JSC-TTC v \$40,000 for the federal programs to expect the JSC cash cost per job to become lower over the next three years—somewhere between \$10-15,000. Of course the JSC-TCC direct job will also generate indirect jobs in the Houston-Clear Lake region, perhaps by a factor of 3-4. New capital generated in the region is currently about \$200,000 but that too will change in the next three years. I hope that it

will exceed the \$100 million raised for tenants in the IC² Austin Technology Incubator over the past four years.

More time is needed to see if the investments by JSC in the licensed technology provide an adequate ROI. Industry goals for some decades have been 15-20 percent per annum. We will have to be very much more proactive for such a rate. On the other hand, a real return of 2-3 percent in the next three years is attainable—O.K. with a lot of luck.

The technology transfer through dual missions is an exciting challenge. Those involved with JSC-TTC have had as their goal the short term commercialization of JSC technology currently on the shelf. It will be both awesome and rewarding to change our emphasis, even for a short time, to develop several major technology transfer missions that have major payback opportunities. We look forward to the Goldin challenge. We are proud to be members of JSC-technology transfer virtual network for commercialization.

0.11

**Session C1: INNOVATIVE MICROWAVE AND
OPTICAL APPLICATIONS**

Session Chair: G. D. Arndt

MICROWAVE SENSOR FOR ICE DETECTION

by

G. D. Arndt and A. Chu
NASA/Johnson Space Center

L. G. Stolarczyk and G. L. Stolarczyk
Raton Technology Research, Inc.

ABSTRACT

A microwave technique has been developed for detecting ice build-up on the wing surfaces of commercial airliners and highway bridges. A microstrip patch antenna serves as the sensor, with changes in the resonant frequency and impedance being dependent upon the overlying layers of ice, water, and glycol mixtures. The antenna sensor is conformably mounted on the wing. The depth and dielectric constants of the layers are measured by comparing the complex resonant admittance with a calibrated standard. An initial breadboard unit has been built and tested. Additional development is now underway.

1.0 INTRODUCTION

Ice build-up on the Orbiter low temperature fuel tanks, airfoil surfaces on commercial airplanes, and highway structures can create hazardous transportation conditions. Ice build-up on the Orbiter's low temperature fuel tanks is a safety concern in NASA's Space Shuttle Program. Commercial airline disasters in Washington, D.C., Denver, Colorado, Newfoundland, New York, and Europe have been caused by wing icing prior to the take-off role. Adherence of ice to the wing can reduce lift by as much as 34 percent and on some aircraft ice may dislodge and damage the jet engine. Formation of ice on airfoil surfaces claimed 496 lives between 1977 and March 1992 in fatal commercial aircraft accidents. All of these tragedies could have been prevented had there been a reliable device to determine whether deicing was needed. The Federal Aviation Administration (FAA) has projected that unless safety improvements are made, there could be eight additional air carrier icing accidents within the next 10 years claiming 143 lives. In addition, the cost to improve safety operations during winter time could cost approximately \$181 million. Type I anti-icing fluid is spread on the aircraft prior to take-off. In one airport alone, 700,000 gallons of highly toxic ethylene glycol are used during the winter season. The airport waste water treatment load is equivalent to a city of 500,000 people.

Icing on highways is another national concern. Nearly four million miles of U. S. Roadways and 500,000 bridges are subject to icing conditions at some time during the year. Almost one third of all U. S. traffic accidents occur on wet, ice, or snow/slush covered roadway surfaces.

This microwave ice detection sensor project will improve transportation safety by developing thin conformally mounted sensor technology. The sensor and associated microcomputer controlled electronics can determine the electrical parameters of the contaminant layer overlying the sensor.

Theoretical and experimental studies of a resonant circular microstrip patch antenna sensor have shown that its measurable electrical properties are highly sensitive to the contaminant layer properties. The resonant frequency and terminal impedance of the patch antenna depend upon the layer depth, resistivity, and dielectric constant of the overlying layer. Laboratory tests have shown that the sensor and associated electronics can discriminate between water, ice, snow/slush, and antifreeze layers. Measurements can determine the thickness of the layer, as well as the ethylene-glycol mixture.

2.0 PRELIMINARY LABORATORY TESTS WITH PATCH SENSOR

Theoretical and experimental investigation of the resonant microstrip antenna sensor found that the percentage change in resonant frequency and conductance due to overlying ice and water layer depth could be measured with a practical instrument. The thin microstrip antenna sensor is illustrated in Figure 1.

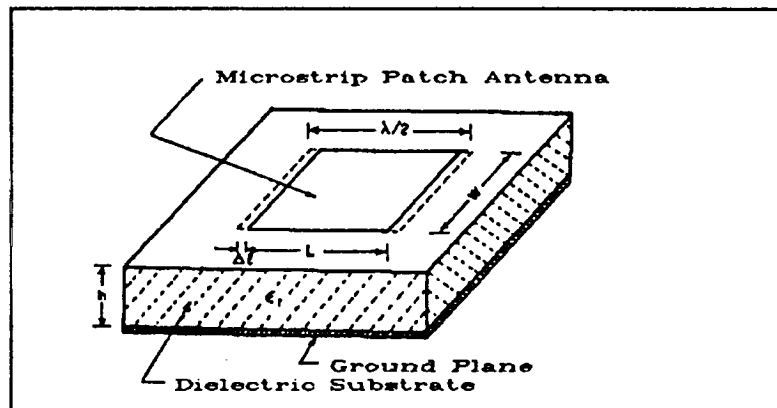
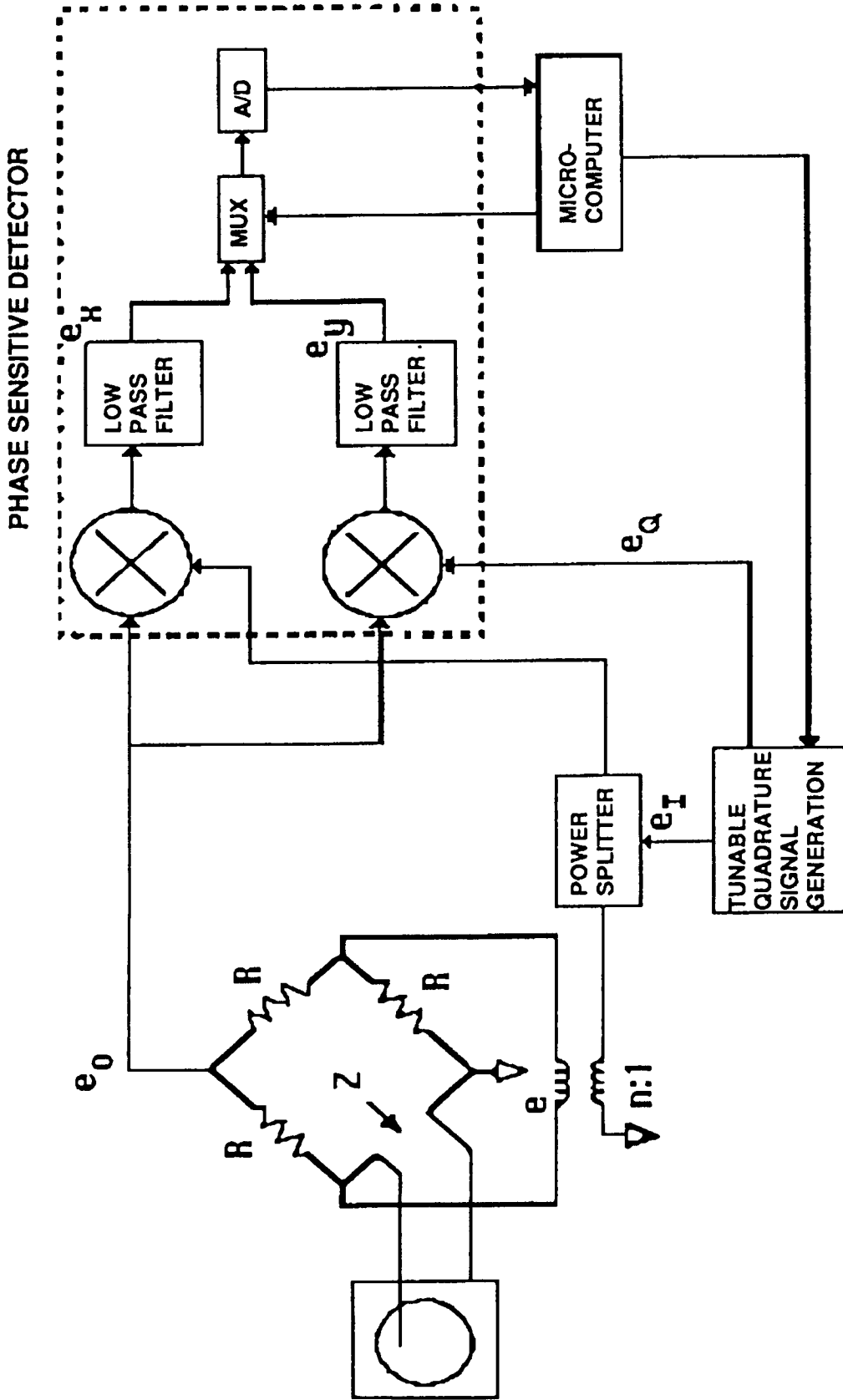


Figure 1. Resonant antenna structure for measuring ice thickness layers.

The block diagram of the resonant antenna sensor and bridge measuring instrument is shown in Figure 2. The instrument is designed to measure the resonant frequency (f) and input admittance of the microstrip antenna.

Figure 2.



BLOCK DIAGRAM OF THE BRIDGE AND PHASE SENSITIVE DETECTOR AND QUADRATURE SIGNAL GENERATION NETWORKS.

The input admittance of the antenna is given by:

$$Y = G + iB$$

where G = input conductance of the antenna in Siemens,
and B = input susceptance.

At the resonant frequency (f_r), the input susceptance (B) is equal to zero.

The antenna admittance measurement will be made with a Maxwell bridge circuit. The bridged output signal (e_o) will be applied to the phase detector (PD) network and measured by the combination of A/D-microprocessor. The microprocessor program includes an algorithm for computing the admittance from the measured values of the PD signals, e_x and e_y . The microprocessor controls the frequency of the Quadrature Signal Generation (QSG). The resonant frequency (f_r) is determined by incrementing the QSG frequency until the measured input susceptance becomes zero.

Computer codes were developed and used to determine the resonant frequency change of a microstrip antenna due to ice buildup. The theoretical results are illustrated in Figure 3.

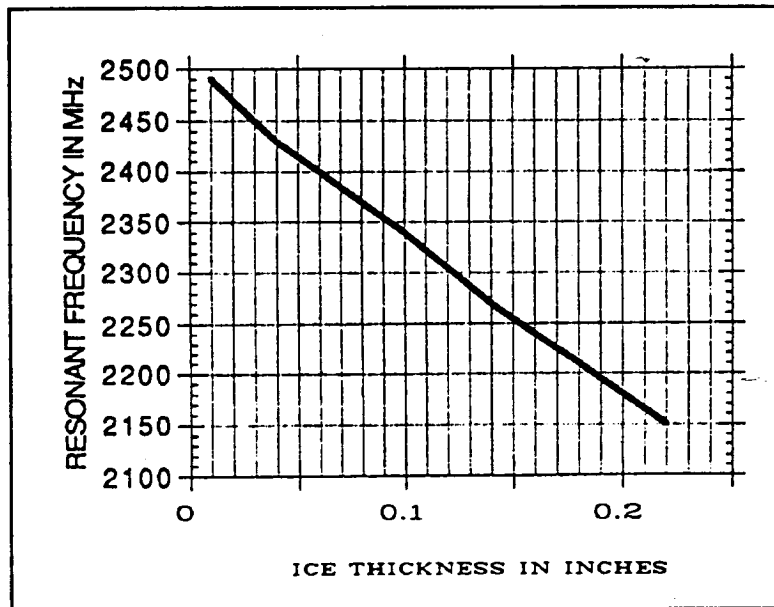


Figure 3. Theoretical resonant frequency versus ice thickness.

The resonant frequency changed by 140 MHz (5.6 percent) for 0.1 inch change in ice layer depth. A 0.25 inch layer of ice caused the resonant frequency of the antenna to decrease by 10 percent. To investigate the ice and ice-water layering

behavior in detail, a series of experimental tests were conducted in a temperature controlled chamber. In these tests, 0.1 inch depth increments of water were added to a tray in which the microstrip antenna sensor formed the bottom of the tray. The resonant frequency and conductance were independently measured after each incremental change in water depth. The measurements were repeated after one hour when 0.1 inch water layer turned to ice. The test data is illustrated in Table A.

TABLE A

CIRCULAR MICROSTRIP ANTENNA
 RESONANT FREQUENCY AND RESONANT CONDUCTOR
 VERSUS ICE AND 0.1 WATER-ICE DEPTH

ICE DEPTH INCH	ICE		0.1 INCH WATER AND ICE	
	f_o (MHz)	G (mS)	f_o (MHz)	G (mS)
0.0	821.21	13.4	797.09	21.5
0.1	812.98	15.4	784.14	28
0.2	815.28	17.0	786.21	30.7
0.3	812.07	17.7	782.0	39.4
0.4	805.02	22.3	772.78	55.4

The measured data is illustrated in Figure 4.

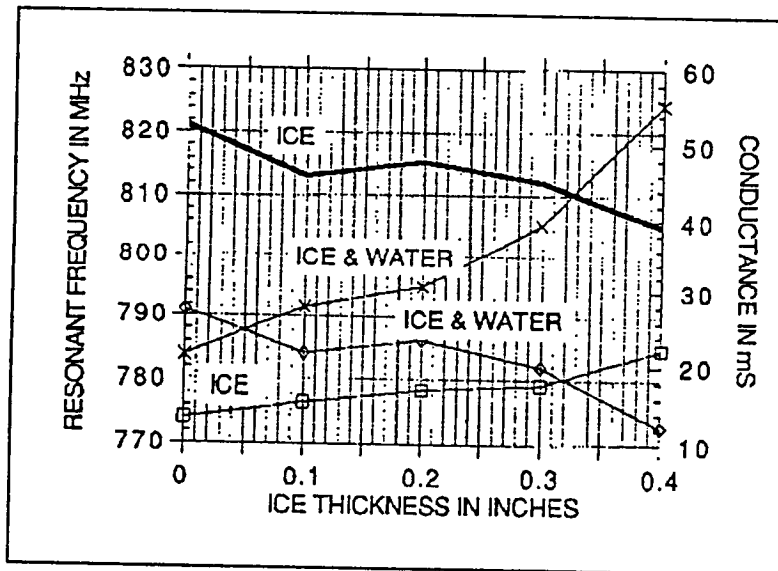


Figure 4. Measured resonant frequency and conductance versus ice thickness at 10°F.

The dark solid and square symbol curves indicate the change in resonant frequency and resonant conductance versus ice layer depth. The diamond and cross symbol curves represent the change in resonant frequency and conductance when 0.1 inch depth of water covers the antenna. The dark solid and diamond resonant frequency curves show that the frequency is significantly changed when a water layer is present.

SUMMARY

Theoretical and experimental research work has found that the contaminant layer properties can be measured with the sensor. The thin conformal sensor can be applied to the critical areas of the wing without altering its aerodynamic properties. Fusion of wing surface temperature measurements data with sensor data will enable estimation of time to freeze. Fusion of piezoelectric data will enable cross check of ethylene-glycol water mixer and freezing point. Adherence of the layer to the wing surface can also be detected with this type of sensor. The sensor fusion goal is to maximize reliability of ice detection for transportation safety.

S2-51

10-40

**Determination of the Residence Time
of Food Particles During Aseptic
Sterilization**

J. R. Carl
Lockheed Engineering and Sciences Company
Houston, Texas 77058

G. D. Arndt
NASA/Johnson Space Center
Houston, Texas 77058

T. X. Nguyen
NASA/Johnson Space Center
Houston, Texas 77058

Abstract

The paper describes a non-invasive method to measure the time an individual particle takes to move through a length of stainless steel pipe. The food product is in two phase flow (liquids and solids) and passes through a pipe with pressures of approximately 60 psig and temperatures of 270-285° F.

The proposed problem solution is based on the detection of transitory amplitude and/or phase changes in a microwave transmission path caused by the passage of the particles of interest. The particles are enhanced in some way, as will be discussed later, such that they will provide transitory changes that are distinctive enough not to be mistaken for normal variations in the received signal (caused by the non-homogeneous nature of the medium). Two detectors (transmission paths across the pipe) will be required and placed at a known separation.

A minimum transit time calculation is made from which the maximum velocity can be determined. This provides the minimum residence time. Also average velocity and statistical variations can be computed so that the amount of "over-cooking" can be determined.

Background

The food industry seeks a non-invasive method (i.e., one that does not impede the product flow) to measure the time it takes for a food particle to move from one end of the tube to another. The mean velocity of all particles cannot be used since there is a velocity distribution. Ten to twenty of food processor association member companies are interested since they are looking for a method to determine the shortest residence time over a statistically representative number of particles. This information will be used to determine if the food has been properly sterilized. Safety of the food product is of primary importance to this industry. The system could be used in the United States and internationally to provide better tasting and safer processed foods for consumers.

Description of Technique

Our technique can best be described by a short discussion of (a) the sensors (detectors), (b) approaches to tagging food particles, (c) an approach to the extraction of information from the microwave signal, (d) velocity calculations, and (e) the need for high data rates.

Microwave transducers consisting of a transmitting aperture and a receiving aperture are used as sensors. The apertures are loaded to match the fluid medium on the pipe end and matched to a 50 ohm transmission line on the other. These devices are referred to as detectors. Two detectors are required, one upstream and one downstream. Detectors are designed to operate in the heat environment associated with sterilization.

Particles of interest are tagged (during test runs) by insertion of a material that is electrically dissimilar to the particles and the fluid in the pipe. However, the tags are small enough to be inserted inside individual food particles and similar in density to the particles so as not to alter the dynamics of the particle flow. Tags are designed to function in the heat environment associated with sterilization. Tags capable of generating harmonics or some sort of modulation on the carrier may be used to identify the passage of particles if necessary.

Variations in amplitude and/or phase of the signal propagating across the pipe provides the expected method of identifying particle passage. In some cases, the particles themselves may provide this signal variation. However, if particles are too similar to the surrounding fluid, a tag must be used to provide a strong property variation in the medium. The detection of a harmonic of the carrier or some modulation added by the tag is also a possibility for identifying particle passage. See Figure 1a for an example of amplitude detection.

Maximum velocity is determined by detecting the minimum transit time between the two detectors of known spacing. Also average velocity and statistical variations about the average may be easily calculated. See Figure 1b.

Fast data rates are needed to provide good time resolution on each detection. Data rates of 10,000 to 100,000 acquisitions per second can be easily accomplished using relatively inexpensive analog to digital converters (data acquisition boards) associated with a personal computer. This would provide a time resolution of 100 to 10 microseconds. For a food particle velocity of 1 meter per second, a 100 microsecond time resolution would yield a distance resolution of 0.1 millimeter.

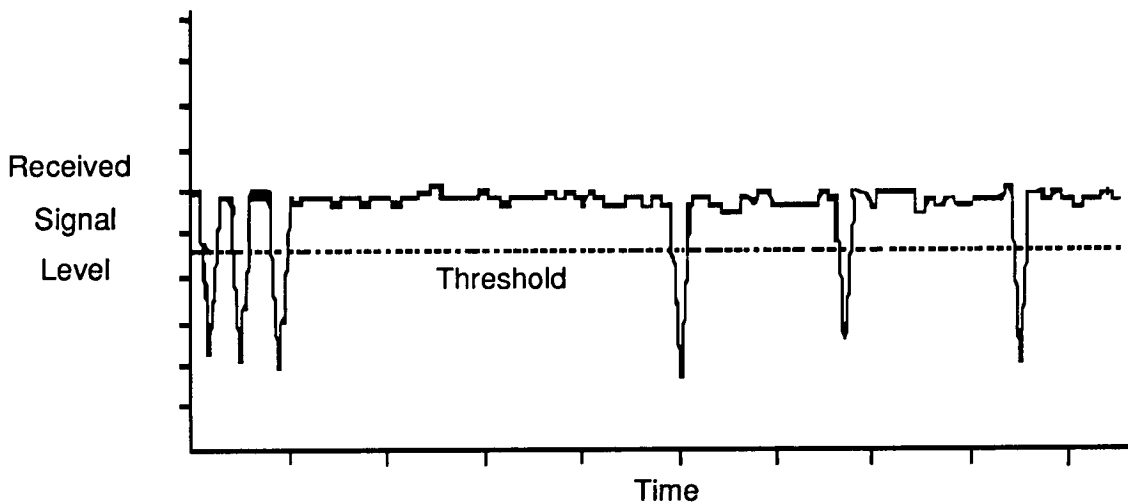


Figure 1a. Amplitude Detection

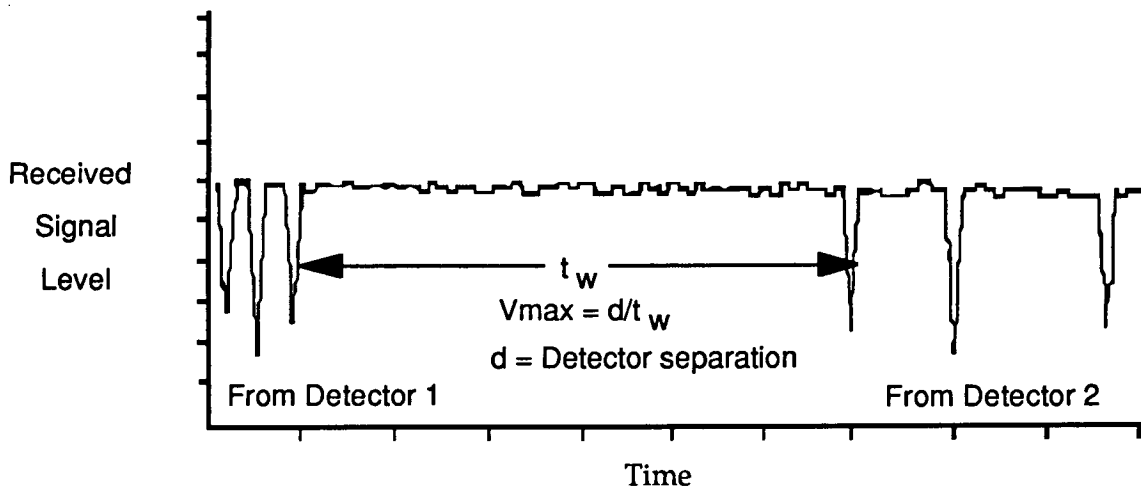


Figure 1b. Maximum Velocity Calculation

An insertion port and an extraction port will be required for test runs. Several test runs can be performed, if needed, to provide good statistical information.

Tags can be removed from food particles (at the extraction port) after each test run and reused or new tags can be supplied for additional runs. Tags are expected to be inexpensive and the latter seems more likely.

System Description

Choice of Frequency

Transmission loss through the food medium is expected to increase rapidly with frequency. The frequency of operation should be low enough such that the transmission loss will not exceed about 100 db for a transmission path of 2 inches. Greater transmission losses would likely make it difficult to avoid leakage paths from transmitter to receiver that would cover up the desired signal. On the other hand, the frequency must be high enough that it is above the cutoff frequency of the pipe when the food medium is present. (When food is not present in the pipe, the chosen frequency is likely to be well below cutoff such that no energy will propagate.) These considerations lead to a frequency of operation between 1 and 4 GHz. Transmission loss and cutoff frequency are expected to increase with temperature. The relationships will need to be established for the various types of food media of interest.

Discussion of the Transmission Medium

It is expected that the medium (food) has a high water content and a high ionic content due to seasoning. The dipolar properties of water molecules will cause the medium to have a high relative dielectric constant probably in the range of 50 to 60 for the frequencies of interest. The conductivity of the medium is likely to be in the range of 1 to 10 mhos per meter for the frequency range and temperature range of interest. The frequency range of interest is approximately 1 to 4 GHz. The expected transmission loss at these frequencies will be discussed later. Some variations in the received signal are expected even when no meat particles pass through the detectors. This is caused by the non-homogeneous nature of the medium.

Marking Particles

Perhaps the most simple way to mark a particle of food is by the insertion of a pellet of some light metal such as beryllium, magnesium or aluminum. The light weight metal should not significantly alter the specific gravity of the food particle but should significantly increase the transmission loss. If this does not provide sufficient marking, the metallic marker can be made into the form of a resonant half wave dipole. A resonant

dipole in the food medium is much smaller than in air. This should provide a distinctive blip in the received signal as it passes between the windows. If a still more distinctive blip is needed, a non-linear device such as a diode or transistor can be used as the marker. The device is arranged to retransmit not only the incident frequency but also harmonics thereof. The receiver is tuned to a harmonic so that it will only receive a signal only when a particle is in the transmission path. Diodes for this purpose should be available for under one dollar each.

Equipment Requirements

The major items of equipment are as shown in Figure 2 and listed below:

- A microwave source (1 to 4 GHz)
- Two microwave receivers/information extractors
- Two special sections of stainless steel pipe equipped with opposing windows as shown in Figure 3a and 3b.
- Two Transducers (antennas) for each detector equipped with appropriate tunable matching devices, and coaxial connectors. Cooling may be provided for the transducer if required.
- A two channel fast A/D converter to digitize the output of the receiver.
- A 486-DX or equivalent computer equipped with 120MB of hard disk storage and monitor. A portable computer is desired for the demonstration hardware.
- Software for control, data gathering, data processing, and display. The data processing may require custom software but of a simple nature.

The microwave source is perhaps nothing more than a crystal controlled oscillator, an inexpensive item. The receivers are likely to be single frequency units. These are also expected to be relatively inexpensive. After the initial design and development, the windows transducers and special pipe section should not be expensive. The personal computer, A/D converter board, and software package are also readily available and not expensive. If custom software is required, this will require, perhaps, a few weeks of engineering time.

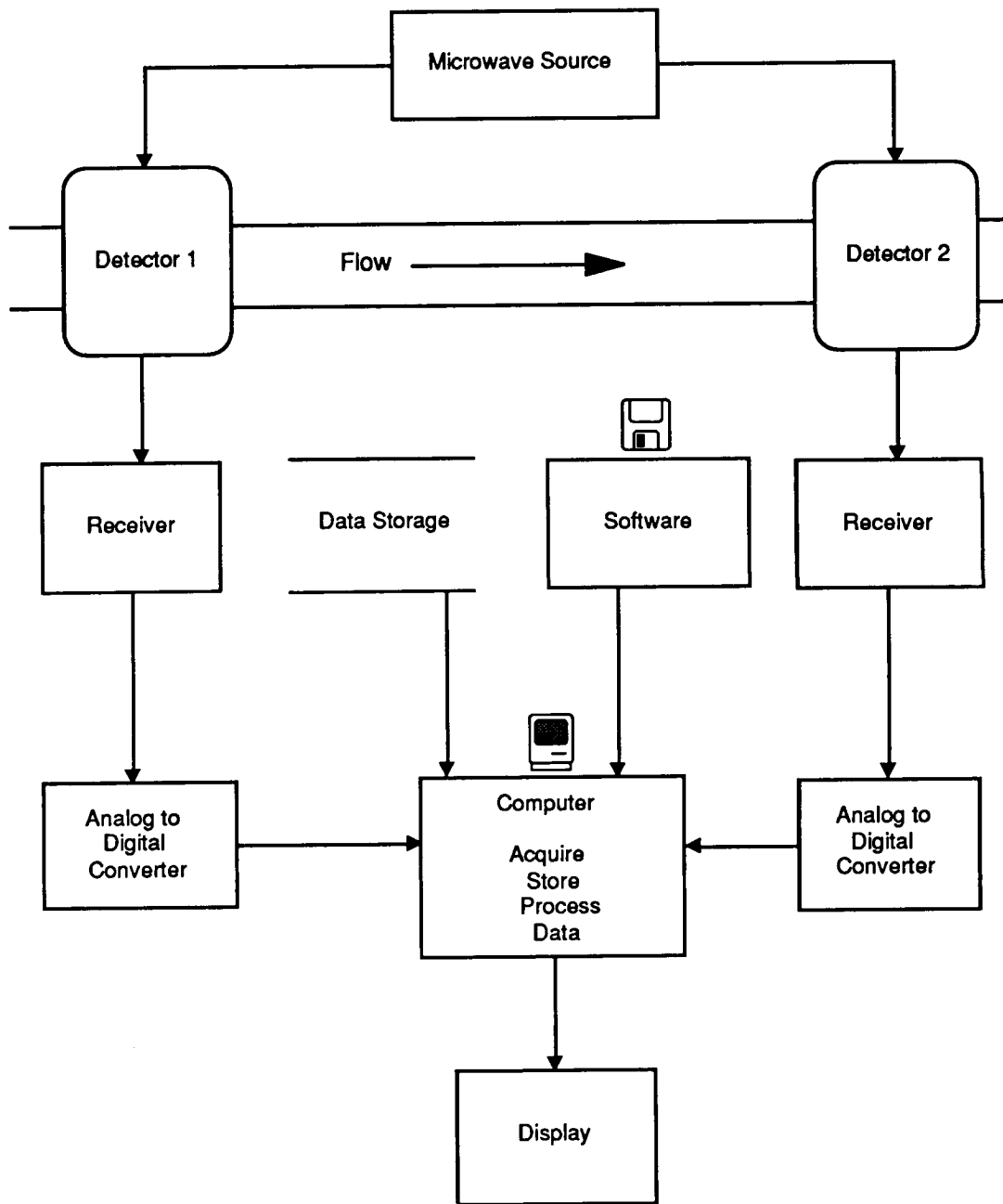


Figure 2. Major Equipment Items

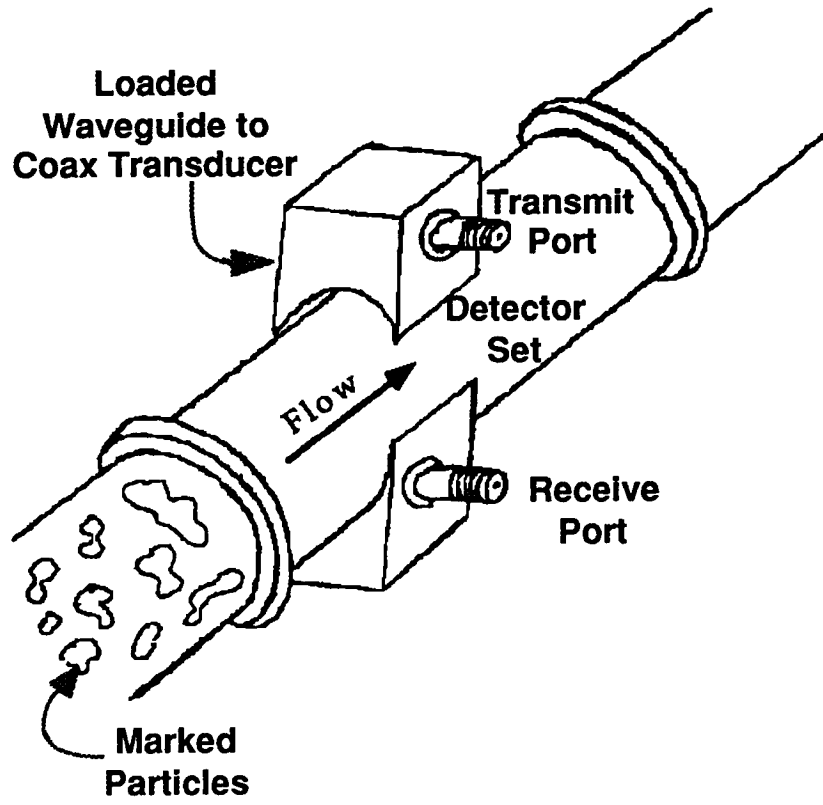


Figure 3a. Microwave Transparent Pipe Section

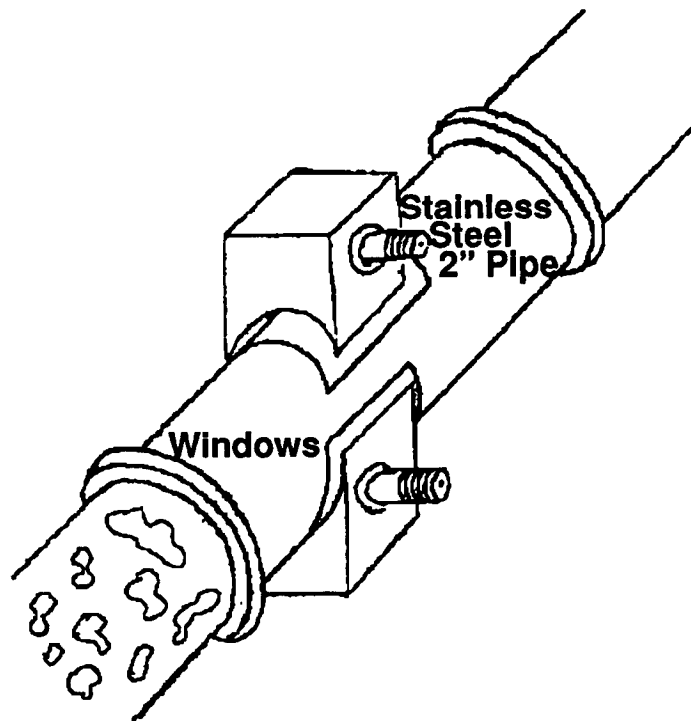


Figure 3b. Microwave Transparent Windows

Planned Test Program

The plan is to develop a breadboard unit consisting of a single transmitter and receiver, data processing hardware and software (personal computer and commercially available data handling software), and candidate antennas. The test parameters of initial interest are the absorption characteristics of the food medium as a function of input frequency (in the ranges of 1 to 4 GHz), the identification signatures of the various individual tags, and the impedance matching of the transmitting and receiving apertures (antennas) to the food medium. This breadboard unit, using mainly standard microwave laboratory equipment, should be ready for testing within eight weeks. The objective is to prove this microwave technique will work in the food environments of interest. Once the detection process is proven, it is straight forward to incorporate a second detection in the hardware and signal correlation software for making individual particle velocity measurements. It is our intention to develop and test the initial laboratory system. Further development of a prototype system for testing in a food plant is dependent upon the interests of the National Food Manufacturer's Association.

Equipment Development Schedule and Cost

The initial laboratory system should be developed within four months. The cost goal for a completed system (hardware and software) for use in a food processing plant is less than \$20,000.

Conclusions

The technique described herein should provide a feasible and inexpensive technique to measure the shortest residence time of food particles during an aseptic sterilization process. Also, average residence time and statistical variations from the average are easily computed to provide "over-cooking" requirements.

Proof-of-concept testing will begin in early February of 1994. Depending on the results of these tests and the availability of funds from the Food Processors Association, the development of demonstration hardware will proceed shortly thereafter.

**ELECTROMAGNETIC PROBE TECHNIQUE
FOR FLUID FLOW MEASUREMENTS**

G. D. Arndt
NASA/Johnson Space Center
Houston, Texas 77058

J. R. Carl
Lockheed Engineering and Sciences Company
Houston, Texas 77058

T. X. Nguyen
NASA/Johnson Space Center
Houston, Texas 77058

ABSTRACT

The probes described herein, in various configurations, permit the measurement of the volume fraction of two or more fluids flowing through a pipe. Each probe measures the instantaneous relative dielectric constant of the fluid in immediate proximity. As long as separation of the relative dielectric constants of each fluid is possible, several or even many fluids can be measured in the same flow stream. By using multiple probes, the velocity of each fluid can generally be determined as well as the distribution of each constituent in the pipe. The values are determined by statistical computation. There are many potential applications for probes of this type in industry and government. Possible NASA applications include measurements of helium/hydrazine flow during rocket tests at White Sands, liquid/gas flow in hydrogen or oxygen lines in Orbiter engines, and liquid/gaseous Freon flow in zero gravity tests with the KS135 aircraft at JSC. Much interest has been shown recently by the oil industry. In this industry, a good method is needed to measure the fractions of oil, water, and natural gas flowing in a pipeline and the velocity of each. This particular problem involves an extension of what has been developed to date and our plans and program to solve this problem will be discussed herein.

INTRODUCTION

The development of a microwave technique for measuring two-phase flow was originally started due to a desire to monitor the flow of monomethyl hydrazine and helium through an inlet pipe during tests at the White Sands Facility of a reaction control system (RCS) thruster jet. The relative amounts of helium and hydrazine flowing into the thruster jet could not be measured on an instantaneous basis. It was realized that since the dielectric constants of helium (approximately 1) and hydrazine (19.2) are sufficiently different, the load impedance seen by a capacitance probe should be sufficiently different to be easily separable.

The microwave technique that is described in this paper measures the magnitude and phase angle of the reflection parameter, S_{11} , associated with reflected energy from a flush-mounted probe. The system is in the process of being modified to include multiple probes within the pipe. This system has other potential space applications in measuring the flow of liquid and gaseous oxygen or hydrogen under zero-gravity conditions within the Space Station. The technique also has ground-based applications in measuring gas-water-oil flow from undersea oil wells as well as

other possible uses in measuring volume fractions and the velocity of multiple liquids having different dielectric constants.

APPLICATIONS

Single Non-Intrusive Probe

There are many potential applications for a single, non-intrusive probe. For example, a single probe mounted at the top of a pipe can perform well as a bubble detector or void detector. A single probe mounted at the bottom of a pipe could be used to continuously monitor the purity of the fluid. A single probe mounted at an appropriate position on a mixing tank could monitor a change from fluid A to fluid B as a function of time. A single probe mounted strategically could be used to identify laminar or turbulent flow. A single probe may be all that is needed to monitor some point of interest in a pipeline. Combined with apriori information, flow regimes may be indentifiable using a single probe. Of course, a single probe could be used to identify a full or empty tank, or an intermediate threshold level.

Multiple Non-Intrusive Probes

Multiple non-intrusive probes could do any of the things mentioned previously. Identifying flow regimes, and calculating volume fractions could probably be accomplished better with multiple probes located at different positions on the pipe and performing additional processing. Velocity computations would require at least two probes at a known downstream spacing.

Multiple Intrusive Probes

In order to monitor directly what is happening in the interior of a pipe or reservoir, intrusive probes must be used (if using the type probe discussed in this paper). If multiple interior locations are to be monitored, then multiple probes are required. An example of this type of requirement comes from the oil industry. They have a need to measure the volume fraction of oil, water and natural gas flowing through a pipe and the velocity of each. It may be necessary, in this case, to gather data at the interior of the pipe.

If probes are internal, they must be made as minimally intrusive as possible. Also, they must be rigid enough to withstand the flow and tough enough to withstand corrosion and abrasion for long periods of time.

System Description

The major components of a single probe system are shown in Figure 1. This system has been built and used in a test program as described later.

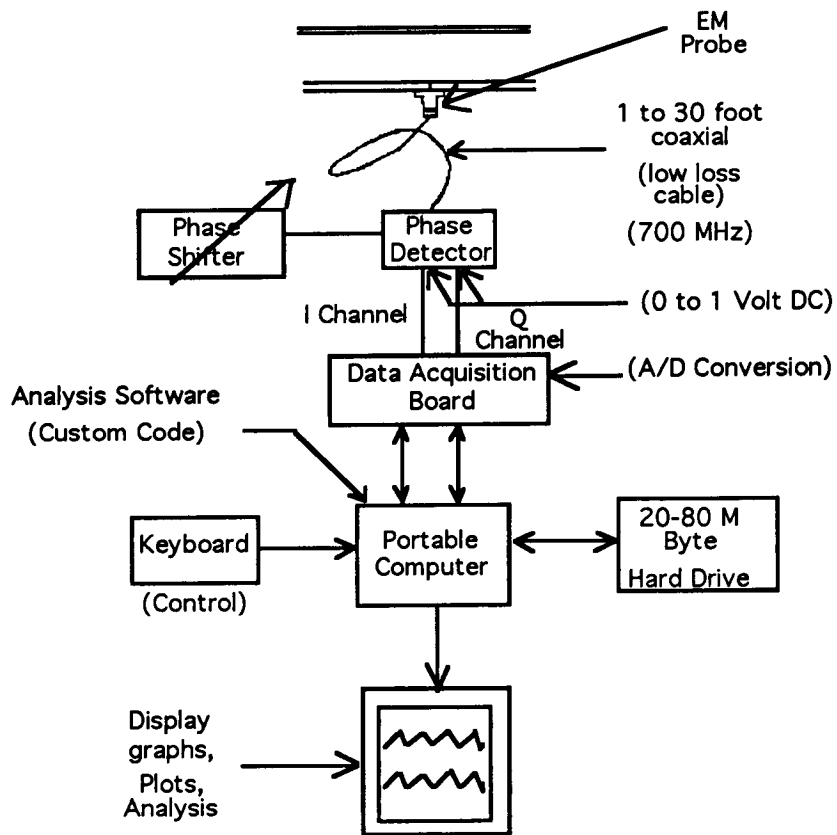


Figure 1. EM Probe System Block Diagram

Probe

The original probe is nothing more than a bulkhead SMA coaxial connector with the center conductor cutoff flush or nearly flush with the inside of the pipe wall. The teflon insulator around the center conductor is cut off flush with the inside of the pipe. Figure 2a shows four such probes mounted around the circumference of a pipe so that it can be determined what is flowing at the top, bottom, and both sides of a pipe. Figure 2b shows two probes at some known spacing to provide the velocity of the fluids flowing at the top of the pipe.

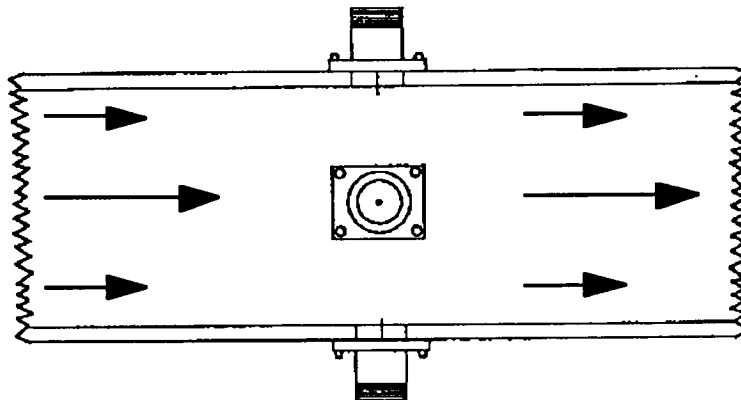


Figure 2a. Multiple Circumferential Probes

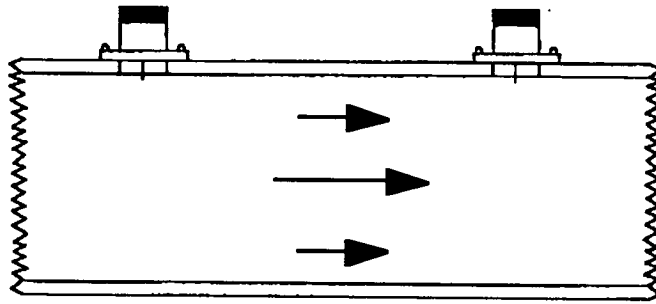


Figure 2b. Multiple Downstream Probes

Phase Detector

A block diagram of the phase detector is provided in Figure 3. This device measures the phase angle on the reflected signal from the probe at approximately 1 GHz. The signal is converted to 100 MHz, amplified, and quadrature phase detected. The two outputs of the phase detector are adjusted to be in the range of 0 to 5 volts. The phase detector was built in-house.

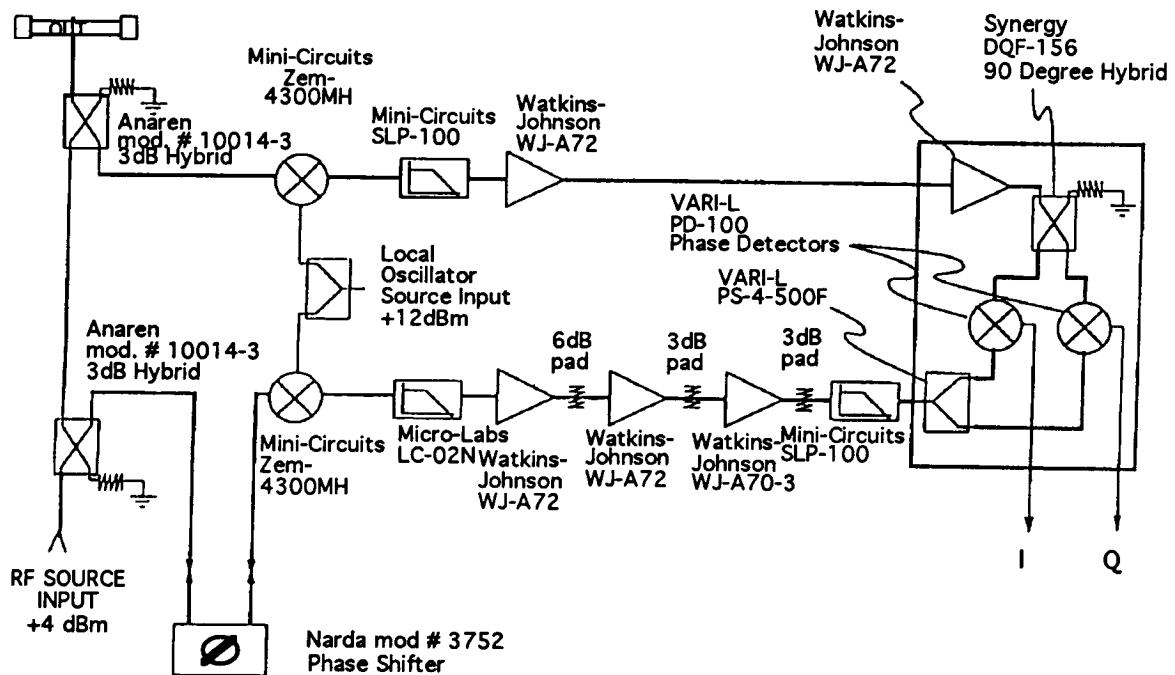


Figure 3. Phase Detector Block Diagram

Data Acquisition Board

There are many commercially available data acquisition boards in the form of expansion boards that are easily inserted into most personal computers (except some of the compact portables and/or notebook computers). Many channels can easily be obtained at reasonable cost and A/D conversion rates are generally available at up to 1 MHz. Conversion rates of 100 KHz are usually adequate.

Portable Computer

A 286 computer with a 20MB hard drive was used for the development of the single probe and used for the testing done to date. However, a 486-DX is being used for the development of the multiple probes with a 120MB hard drive.

Software

The software used for the development and testing of the single probe was written in the "ASYST" language. A commercially available software package called "Viewdac" was purchased for the development of multiple-probe configurations. This package should make it easier to acquire, store, process and display multiple data streams.

THEORY OF OPERATION

The small capacitance of the probe is used as the sensor. This capacitance is a function of the relative dielectric constant of the medium into which the probe is terminated. The short probe "sees" fluid that is no more than a short distance away. If it is required to see further into the medium, it would be necessary to increase the length of the probe so as to increase the volume around the center conductor that forms the probe capacitance.

The phase, or change of phase, associated with the reflected signal at the probe is the quantity measured. The complex "S" parameter associated with reflected energy "S₁₁" is given by:

$$S_{11} = \left(\frac{Z_0 - Z_L}{Z_0 + Z_L} \right) \quad (1)$$

Where: Z_0 is the characteristic impedance of the transmission line from phase detector to probe.

Z_L is the probe input impedance = $R + jX$

Typically, Z_0 is equal to 50 ohms. If there is a negligible energy coupled to the media, the probe resistance is very small. The input impedance is essentially a capacitive reactance in which case equation (1) can be written:

$$S_{11} = \left(\frac{50 - jX}{50 + jX} \right) \quad (2)$$

$$\text{Where: } X = \frac{-1}{2\pi F C \epsilon_r}$$

C = probe capacitance

ϵ_r = relative dielectric constant

F = frequency of operation

The phase " ϕ " on S_{11} , from equation (2), can be extracted as:

$$\phi = \tan^{-1} \left(\frac{100X}{2500 - X^2} \right) \quad (3)$$

for certain cases of interest, where X is large with respect to Zo and where ϕ is small, equation (3) reduces approximately to:

$$\phi \approx -100 / X \text{ radians} \quad (4)$$

$$\phi \approx -5730 / X \text{ degrees}$$

For a certain 1mm probe, the probe capacitance has been measured to have a capacitance of approximately 0.04 pf. Using this value and using a frequency of 700 MHz equation (4) reduces to the convenient form:

$$\Delta\phi \approx \Delta\epsilon_r \quad (5)$$

For this model of the probe, i.e., a capacitive reactance termination for the transmission line, predictions can be made for probe capacitance given frequency, probe length, and the effective relative dielectric constant of the media. Also the probe's sensitivity can be readily formulated.

FLAPPER VALVE TEST RESULTS

The results shown in Figure 4 were extracted from the test results of a "Flapper Valve Experiment" performed at NASA, JSC, Houston Texas in November of 1992. The flush probe was mounted at the top of a 1^{1/2} inch pipe and monitored the flow of distilled water and dry nitrogen flowing at various specific rates through the pipe. The volume fraction of water and nitrogen were varied. The top graph shows that the probe works well as a bubble detector in this configuration. The lower two plots show two different conditions of slug flow. The precise volume fractions and flow rates are not immediately evident from these plots but by processing the data, introducing apriori knowledge, and by influencing the calculation with calibration data, perhaps reasonably accurate volume fractions and flow rates could be determined using only a single probe.

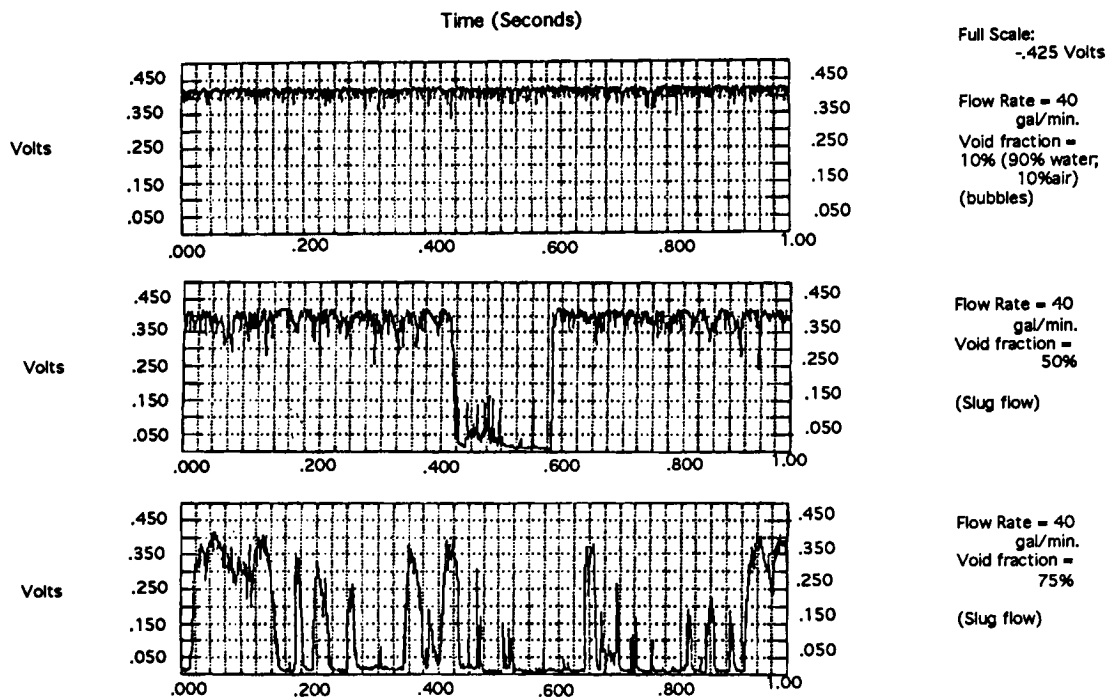


Figure 4. Flapper Valve Experiment Using Distilled Water and Nitrogen

Multiple Probe Design

Certain potential applications of the AC fluid flow probe requires the use of multiple probes placed in the flow stream. For example, consider the problem of measuring the volume fraction of oil, water, natural gas, and oil/water emulsions (flowing in a pipe line). Various flow regimes are possible which may influence the choice of probe locations in the flow stream. By placing probes at various strategic locations within a cross-section of the pipe, the volume fractions of each constituent can be calculated statistically by a technique described herein. Also, by using an identical probe configuration downstream from the first, the velocity of each constituent can, in most cases, be measured.

Probe structures that are placed in the flow stream must be designed to have minimum impedance to flow and minimum effect on the flow while having good strength and durability characteristics. A third generation probe design is now being used as shown in Figure 5.

These probes should be large enough to provide a reasonable capacitance at 100 to 300 MHz (for sensitivity), yet small enough to be relatively invisible to the flow. Also the probes should be self-cleaning and should not interfere with each other electrically or physically.

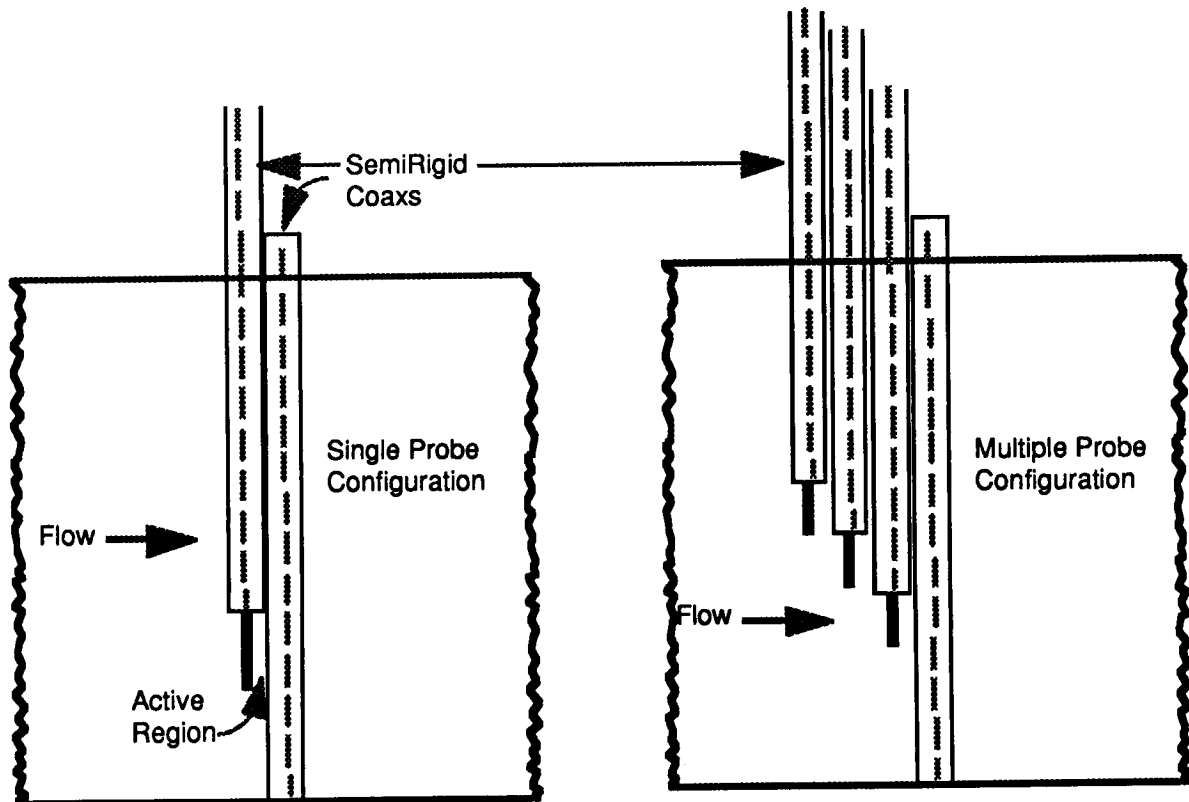


Figure 5. Prototype Stack of 3 Probes

Oil/Water Tests

A flow loop has been constructed in building 14 and is presently being used for oil/water mixture tests. The results thus far, are encouraging in that it has been shown that for slug flow, or flow with some degree of non-homogeneity, flow velocity can be measured. Also, the passage of bubbles larger than 200 microns of oil, air and water can be seen in the data and counted. For homogeneous flow, it has been shown that the voltages from the phase detector provides a measure of the fraction of water in an oil continuous emulsion. Also, the conditions for a switch to a water continuous emulsion have been shown. For salt water or fresh water

continuous emulsion, phase detector voltage also provides a measure of the fraction of oil present provided that the conductivity of the water is below a critical value.

Conclusions

The probes described herein have demonstrated their ability to measure fluid velocities, slug passage, slug lengths, bubble passage and bubble size. Also, volume fractions can often be measured. Although the probes show promise of working effectively for the expected flow regimes in an oil well line, they do not appear to work well for small scale homogeneous fluid mixtures. Work is continuing to determine the conditions under which the probe is useful.

References

- A. Cartellier and J.L. Achord, *Local Phase Detection Probes in Fluid/Fluid Two-Phase Flows*, American Institute of Physics, February 1991.
- D. Brown, J.J. den Boer and G. Washington, *A Multi-Capacitor Multiphase Flowmeter for Slugging Flow*, Shell Research Paper 2.2, North Sea Flow Measurement Workshop, October 1992.

54-52

**DUAL USE OF IMAGE BASED TRACKING TECHNIQUES:
LASER EYE SURGERY AND LOW VISION PROSTHESIS**

Richard D. Juday, Ph.D.
R. Shane Barton
Hybrid Vision Project
Tracking and Communications Division
NASA Johnson Space Center
Houston TX 77058

ABSTRACT

With a concentration on Fourier optics pattern recognition, we have developed several methods of tracking objects in dynamic imagery to automate certain space applications such as orbital rendezvous and spacecraft capture, or planetary landing. We are developing two of these techniques for Earth applications in real-time medical image processing. The first is warping of a video image, developed to evoke shift invariance to scale and rotation in correlation pattern recognition. The technology is being applied to compensation for certain field defects in low vision humans. The second is using the optical joint Fourier transform to track the translation of unmodeled scenes. Developed as an image fixation tool to assist in calculating shape from motion, it is being applied to tracking motions of the eyeball quickly enough to keep a laser photocoagulation spot fixed on the retina, thus avoiding collateral damage.

INTRODUCTION

This is a difficult paper to write, with respect to striking an appropriate balance between being complete on the one hand, and too thinly spreading out information of use to the target audience on the other. I will try to leave much of the technical development to the reference material, presuming a technically competent reader who is not an expert in these disciplines.

NASA's Johnson Space Center (JSC) has been actively developing its image based tracking technology for two medical applications: noncontact retinal position stabilization for laser photocoagulation, and image warping as a prosthesis for certain field defect forms of human low vision. In the retina surgical application, JSC is working with Pinnacle Imaging, a small company dedicated to the development of surgical and robotic methods and equipment for the surgical market. In the image warping application, we are working with the University of Houston's College of Optometry (UHCO) and the University of Pennsylvania. As an Agency, NASA has formal connections with an industrial and medical group that has built a heads-up display system that can directly incorporate our results into prosthetic hardware.

Ophthalmic laser procedures

In the eye surgery application, our central problem is maintaining the aim point of a laser during retinal photocoagulation. Similar image processing hardware may allow tracking assemblyline parts in jumbled orientation, as on a conveyer belt, with no conditions on their placement or orientation. NASA is developing an applicable technology: image correlation based tracking. The eyeball (and the retina with it) undergo continual small-angle jitter, along with the larger jerky saccades that are typified by eye motion while reading. Both are detrimental to the photocoagulation procedure. Incorrect locations on the retina may be struck by the laser; in addition to causing collateral damage to the retina, this limits how close eye surgeons are comfortable working to the fovea (the highest resolution part of the retina -- the part we use when looking directly "at" an object).

JSC has patented¹ a technique that provides fixation; that is, regardless of exactly *what* is in the scene, the input sensor is given the information to allow it to stare at a single location in the scene even as the spacecraft moves or jerks. This is important to simplify analysis of the dynamic scene for information about the three dimensional nature of the unknown surface. "Optic flow" is the technical term for the intra-image motion that results from a shifting perspective of a three-dimensional surface, and computation of optic flow is far eased if fixation is maintained on some point in the image. Our method is based on what is technically described as the

optically derived joint Fourier transform correlation. Full details are presented in the patent and in other literature^{2,3,4,5}.

Field Defect Low Vision Prosthesis

NASA has an Agency presence in image processing technology for low vision sufferers. JSC's Tracking and Communications Division and an associated organization, the University of Houston's College of Optometry (UHCO), originally proposed image warping for field defect amelioration. We have developed hardware that either exists^{6,7} or is in advanced design^{8,9} that implements our specially designed mathematical image transformations^{10,11,12,13}, and we have also begun clinical studies whose initial results^{14,15} are favorable to the notion that video rate image warping may yield increased visual function for such disorders as central field blindness (maculopathy) or peripheral field loss (retinitis pigmentosa). If known portions of the retina are dysfunctional, an image may be warped so as to minimize the part of the world view that lands there, while retaining a balance between distorting the world view too much and losing part of that world view. We intend that the end result of this program will be visual prostheses (high tech "glasses"¹⁶) that are more capable, smaller, lighter weight, and less expensive compared with other methods. We intend that the current program will divine mathematical image warpings that maximally increase visual function as tested by reading speed, facial recognition, and so on, but implementing the warpings in inexpensive, lightweight hardware remains an unsolved problem. Among other practical items, we need to determine the range of visual function loss that can benefit from our technique, and then find a minimum set of transforms that span that range.

Fourier Optics

Fourier optics is an important discipline in which the diffractive properties of light become important in the way information is carried and processed. Journals such as Applied Optics, Journal of the Optical Society of America-A, and Optical Engineering carry current articles concentrating heavily in this discipline. Societies such as SPIE - The International Optical Engineering Society publish many Proceedings per year in the discipline. I will limit the discussion here to saying that we bring information to a Fourier optics correlator, usually in video image form, and that an element known as a spatial light modulator (SLM) has its light transmitting properties altered by that video image. Then as coherent light passes through the SLM, the information within the video signal is encoded onto the light beam. Diffraction can then be arranged so as to take the Fourier transform of the encoded image, and subsequent transmission (or reflection) of the light through other elements causes the information to be processed and then presented in simpler form (i.e. bright spots).

Optical Correlation Based Pattern Recognition

(Technically incorrectly, this section will mix some concepts from conjugating Fourier optics [the so-called VanderLugt, or 4-f, correlator] and joint transform correlation [JTC]. I am going to go ahead and do it, even if the expert would squawk; this is more of a tutorial heuristic device than a technical compendium. Just don't jump me for what looks like an error. There is not room to be rigorous. I am also using one-dimensional notation for a two-dimensional signal, pointing out that to do so has become common in Fourier filtersmithing.)

Suppose we have a function $f(x)$ and we wish to filter information from it with a linear system whose impulse response is $h(x)$. The convolution is $y(x) = f(x)*h(x)$:

$$y_{conv}(x) = \int_{-\infty}^{+\infty} f(x-\tau)h^*(\tau) d\tau \quad (1)$$

but we will not technically distinguish between a convolution, $f*h$, and a correlation, $f \star h$

$$y_{corr}(x) = f \star h = \int_{-\infty}^{+\infty} f(x+\tau)h^*(\tau)d\tau \quad (2)$$

since a time reversal of $h(\bullet)$ changes between the two, and $h(\bullet)$ has been arbitrary. As is well known in linear system theory, convolution can be expressed as the inverse Fourier transform of the product of transforms of the functions. That is, if $f(\bullet)$ and $F(\bullet)$ are a transform pair, as are y and Y , and h and H , then

$$\begin{aligned} y(x) &= \mathcal{F}^{-1}\{\mathcal{F}\{f(x)\} \times \mathcal{F}\{h(x)\}\} \\ &= \mathcal{F}^{-1}\{F \times H\} \end{aligned} \quad (3)$$

Also as is well known, there is a shift in $y(\bullet)$ corresponding to a shift in the input signal $f(\bullet)$, a property known as shift invariance. Correlation uses this property to locate a signal that has been identified by the presence of values of $y(\bullet)$ rising above an identification threshold. Optics offers the ability to do the transforms easily by the diffraction of light (once the signal is impressed on it) and to do the multiplication easily (by interacting the light with an SLM on which the function H has been created). There have been three difficult parts. One has been to manufacture SLMs on which desired functions for encoding the signal and the filter. The second is how to make appropriate use of the SLMs we do have; signal theory prescribes ideal values for the filter values, H , that are ordinarily not realizable. Recently¹⁷ we have achieved a breakthrough in optical filter theory that make some of the job easier, and we continue developing modulators and their characterization^{18,19} for other aspects of the problem. The third is how to accommodate the rotation of the retinal image that occurs as the eyeball rotates within the socket, as in response to signals from the vestibular function. Pattern recognition by image correlation is usually quite sensitive to rotation, so we continue our effort to achieve tailored amounts of rotation invariance, and the results will be fed into our activities under Memorandum of Understanding with our medical partner, Pinnacle Imaging. This effort is ongoing under university grant (see Formal Colleagues, below) following concepts laid down several years ago²⁰.

Somewhat simplistically, a Fourier optics image processor converts the video image of the object it has been trained to find into a bright spot of light, and then infers information from that spot. The difficult parts of the information extraction are done off-line and at leisure; the easy part is what happens on-line and in real time. As distinct from most digital image processing, this system determines the correlation between highly synthesized reference objects and the viewed object, and then extracts the information from those correlations. Synthesizing a reference object is a congruent process to obtaining the inverse transform of the optimal filter that can be realized within the limitations of the filtering spatial light modulator. The correlations run at a very high frame rate (tens to thousands per second), allowing comparisons of the input video object with all members of a library of views of that object. Careful crafting of the reference image set has the effect that the library element producing the highest correlation corresponds to the actual position and attitude of the viewed object.

In further distinction from most digital processing, optical pattern recognition operates on the whole input image, rather than individual features, and thus is robust against partial obscuration. The difficult technical challenges are principally in the physical devices -- spatial light modulators -- that encode the incoming video image and the correlation image.

VIDEO RATE IMAGE WARPING

Shift Invariance in Optical Pattern Recognition

As is well known in signal theory, the convolution (equivalently the correlation) of two signals is a linear and shift invariant operation. Also well known is that scale and rotation changes in a Cartesian geometry do not produce shift invariance. If we know $[x_0, y_0]$, the center about which scale and rotation changes occur, though, we can convert scale and rotation into translational shifts by performing a coordinate transformation. The log polar transformation is what we seek.

$$\begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} = \begin{bmatrix} r \cos\theta \\ r \sin\theta \end{bmatrix} \quad (4)$$

shows the conversion between the Cartesian $[x,y]$ and the polar $[r,\theta]$ as coordinate systems. Go one step further and define u as $\log r$. Then as the original image coordinates undergo a change expressed as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = k \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \quad (5)$$

we will find that the coordinate pair $[u,\theta]$ undergoes the change

$$\begin{bmatrix} u' \\ \theta' \end{bmatrix} = \begin{bmatrix} u + \log k \\ \theta + \alpha \end{bmatrix}, \quad (6)$$

so that the rotation through α and scaling by k become translations by α and $\log k$. For video rate pattern recognition that is to be shift invariant to scale change and rotation about $[x_0,y_0]$, and the difficult part is to create the log polar transformation at full video rate.

The Programmable Remapper

With the Programmable Remapper we have solved that problem, at least on a laboratory scale. Its generality allows us to create virtually unlimited geometric transformations at video rate. Conversion to practicality as a prosthesis remains an unsolved, but significant and potentially productive, interest. Read the References for technical details, but suffice it for the present purposes to say that the Remapper can wreak an arbitrary static geometric transformation (at video rate and incurring only a three frame delay) on an incoming video image. We conceived the Remapper as a tool for creating a log polar video image, but we rapidly understood its potential for manifesting other warpings. Human low vision field defects was the first such extracurricular application.

Field Defect Application for Low Vision Humans

As reported in a number of the papers in the References, we have investigated the applicability of image warping for various field defects. For examples of images, see in particular Ref[12]. One of the two principal ones is maculopathy, in which a lesion causes loss of function at the high resolution central portion of the normal field of view. The other principal application is RP (retinitis pigmentosa), also known as tunnel vision, where peripheral vision is progressively lost. Our intent is to move the active portions of a video image so that it falls onto portions of the retina that are still functional, thus having a more nearly complete mapping of the world into the visual cortex (albeit at the expense of local distortion in the normal world-to-cortex representation).

The fovea is the portion of the eye where the resolution is the highest; it is the part of the eye one naturally places into conjunction with the part of the world scene where visual detail is sought. It is the part of your field of view you use when looking "at" something. Loss of foveal function is particularly expensive to detailed functions like reading and facial recognition. Our method is to "rubber-sheet" a video image. Imagine that a television screen is made of rubber, that you poke a hole in the middle, and then stretch that hole out until it is just larger than the blind spot in the center of the field of view. There are many loose parameters in this description, but the core idea is there. In our initial tests, we used normally-sighted volunteers, and we simulated the central blind spot by forcing them to look at an obscuration on a television screen with text scrolling by. An eye tracker followed the gaze of the eye, controlling mirrors that directed the gaze onto the obscuration. The words scrolled by at various font sizes, angular rates, etc. There was significant (though highly variable from subject to subject) increase in reading rate

with image warping. Our next studies will use actual low vision volunteers, it will not force foveation, and we have some advanced mathematical transformations to try. These studies are beginning in Fall 1993.

To deal with the converse problem, we wish to squeeze the normal world field of view into a smaller solid angle. This is easily accomplished, of course, with a minifying telescope, and such is one weapon in the ordinary armamentarium of the RP patient. The problem is that the angular subtense of all elements in the field of view is reduced. Even though you see more of the things one needs to interact with in moving about (door frames, walls, stairs, etc.), the reduced size of objects may give difficulty in reading signs or otherwise resolving the world. Our approach is to give spatially varying magnification, so that localization of gross objects is possible, but full central magnification is maintained. Perhaps in Spring 1994 we will begin field tests with the Remapper in locations like grocery stores, to see if visual functionality is improved.

JOINT TRANSFORM TRACKING

When NASA was recently contemplating autonomous lunar and planetary landings, this project proposed two techniques for lander vision, and we set about demonstrating certain of the technical elements. The first technique was navigation as guided by image correlation, in which images of landmark features would provide reference to absolute geometric coordinates. The second was image fixation by joint Fourier transform correlation, in which unmodeled image structure correlates with itself and shows sequential error in pointing. It is the JTC, as applied to tracking retinal motions, where we have been vigorously spinning off to medical practice.

Dynamics of Eyeball Tracking

Tracking the human eyeball is very challenging. Consider the problem from the other side for a moment; think how quickly you can jerk your eyes over a really large angular distance and settle your viewpoint in on a small object that has caught your attention. Now consider keeping a laser beam fixated on an individual structure within the retina during that motion, to a precision at the same resolution of the smallest object you can resolve. The harmonic content of the eyeball motion extends to as high as 170 Hz, with speeds up to several hundred degrees per second. Those are really astounding figures. Nyquist sampling theory indicates immediately that to track the motion, about 400 independent measurements per second are necessary. If the eyeball motion is able to break lock (e.g. if the eye jerks during a blink), then it is important to know that track lock has broken and in which direction to move to reacquire, but the precision requirements on knowing the translation are reduced since the surgical laser will be inhibited immediately upon loss of track. The optical engineers in the project (the author cheerfully admits to this persuasion) think that the JTC is the preferable method because of the ease with which two qualities are achieved: spatially variable resolution, and a wide field of regard. If tracking is done by digital methods by time-domain correlation, the size of the field of regard is directly impacted by the serial nature of the digital correlation. In contrast, optical processing is inherently parallel, and the size of the field of regard is limited only by the number of pixels in the spatial light modulators. The project currently is headed for a showdown between digital and optical implementation, as the digital method (if practical) would be slightly less expensive to implement.

THE JOHNSON SPACE CENTER PROGRAM

The Hybrid Vision project at JSC has several civil servants (Richard Juday, Tim Fisher, Shane Barton, and Jennifer Yi), NRC post-doctoral fellow Colin Soutar, and in-house contractor personnel employed by Lockheed Engineering and Sciences Corporation (Stanley Monroe, Carlton Faller). Additionally we have either active or recently concluded contracts, fellowships, and grants at a number of universities and corporations (University of Missouri, Carnegie Mellon, Tennessee Technological University, University of Colorado at Boulder, University of Houston, Physical Optics Corporation, Boulder Nonlinear Systems, Physics Innovations) and working relationships with other Government laboratories (Army Missile Command, Air Force Rome Laboratory, ARPA, Air Force Wright Laboratory, Sandia), and others. The project bibliography for the past seven or eight years has over fifty entries including four issued patents (others in work). We are active in image warping, optimal filter theory, optical correlator architectures and applications, spatial light modulator development, with specific thrusts in robotic vision and human low vision. The author will provide copies of the bibliography upon request.

ACKNOWLEDGEMENTS

Financial, moral, and technical support from our military colleagues, mentioned just previously, is particularly appreciated.

REFERENCES

1. U. S. Patent 5,029,220; Richard D. Juday; **Optical joint correlator for real-time image tracking and retinal surgery**; Jul. 2, 1991.
2. Jerome Knopp and Richard D. Juday, "Optical joint transformation correlation on the DMD", Proc. SPIE 1053, 208-215 (1989).
3. Eddy C. Tam, Francis T. S. Yu, Don A. Gregory, and Richard D. Juday, "Autonomous real-time objects tracking with an adaptive joint transform correlator", Optical Engineering 29, 314-320 (April 1990).
4. Eddy C. Tam, Francis T. S. Yu, Aris Tanone, Don A. Gregory, and Richard D. Juday, "Data association multiple target tracking using a phase-mostly liquid crystal television", Optical Engineering 29, 1114-1121 (September 1990).
5. K. L. Schehrer, M. G. Roe, and R. A. Dobson, "Rapid tracking of a human retina using a nonlinear joint transform correlator", SPIE Proceedings vol. 1959, April 1993 (Orlando).
6. U.S. Patent 5,067,019; Richard D. Juday and Jeffrey B. Sampsel; **Programmable remapper for image processing**; Nov. 19, 1991.
7. Timothy E. Fisher and Richard D. Juday, "A programmable video image remapper", Proc. SPIE 938, 122-128 (1988).
8. U.S. Patent 5,208,872; Timothy E. Fisher; **Programmable Remapper with single flow architecture**; May 4, 1993.
9. Timothy E. Fisher and Richard D. Juday, "An improved architecture for video rate image transformations", Proc. SPIE 1098, 224-231 (1989).
10. Richard D. Juday and David S. Loshin, "Some examples of image warping for low vision prosthesis", Proc. SPIE 938 163-168 (1988).
11. Richard D. Juday and David S. Loshin, "Quasiconformal remapping for compensation of human visual field defects: advances in image remapping for human field defects", Proc. SPIE 1053, 124-130 (1989).
12. David S. Loshin and Richard D. Juday, "The programmable remapper: Clinical applications for patients with field defects", Optometry and Vision Science 66, 389-395 (1989).
13. Richard D. Juday, Alan T. Smith, and David S. Loshin, "Human low vision image warping: channel matching considerations", Proc. SPIE 1705 (1992).
14. David S. Loshin, Janice Wensveen, Richard D. Juday, and R. Shane Barton, "Design of reading tests for low vision image warping", Proc. SPIE 1961, Orlando, April 1993.
15. Janice Marie Wensveen, "Reading Rate with Simulated Central Scotoma", Master's thesis, University of Houston College of Optometry, August 1993.

16. Anonymous, "High-Tech Help for Low Vision", NASA Tech Briefs vol. 17 no. 2, pp. 20-22 (April 1993).
17. Richard D. Juday, "Optimal realizable filters and the minimum Euclidean distance principle", Applied Optics 32, 5100-5111 (10 September 1993).
18. Colin Soutar, Stanley E. Monroe, Jr., and Jerome Knopp, "Complex characterization of the Epson liquid crystal television", Proc. SPIE 1959, Orlando, April 1993.
19. Colin Soutar, Stanley E. Monroe, Jr., and Jerome Knopp, "Measurement of the complex transmittance of the Epson liquid crystal television", Optical Engineering (accepted).
20. Richard D. Juday and Brian Bourgeois, "Convolution-controlled rotation and scale invariance in optical correlation", Proc. SPIE 938, 198-205 (1988).

omit

Session C2: COMMUNICATIONS AND DATA SYSTEMS

Session Chair: Andrew Benjamin



02115 TO
P42

ESTL HIGH RATE OPTICAL COMMUNICATIONS SYSTEM (EHROCS)

TEAM MEMBERS

D. JALUFKA
W. WALLINGFORD
W. PHILLIPS
L. CAMPBELL



AGENDA

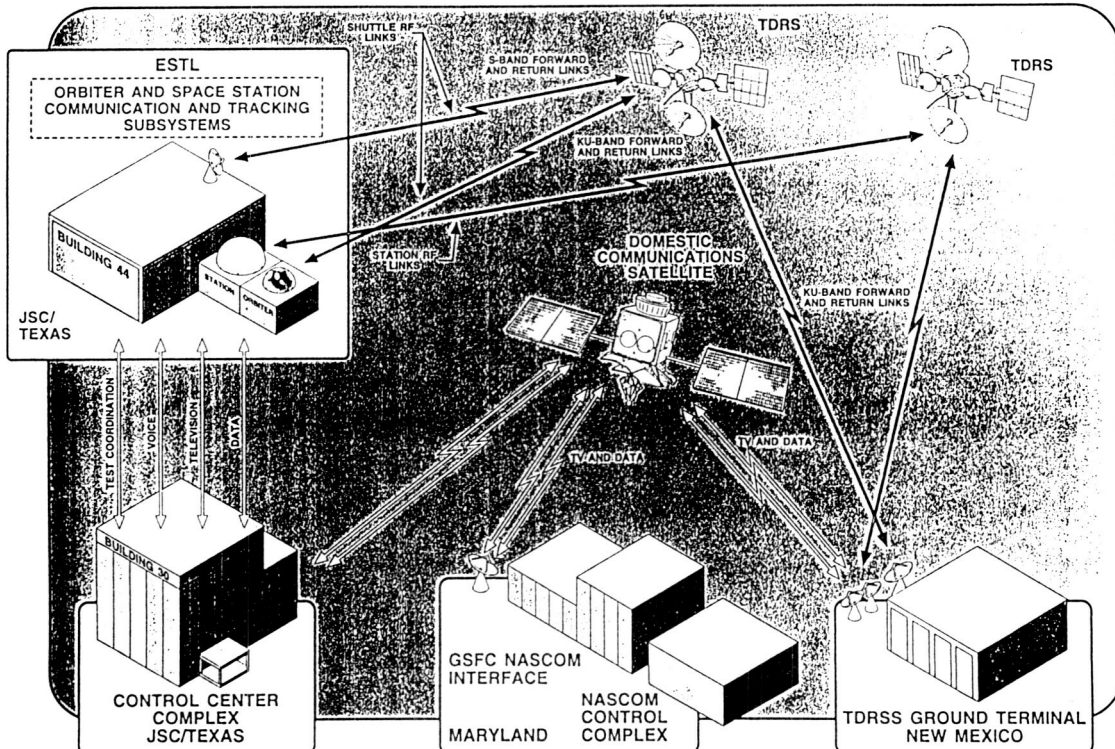
- INTRODUCTION
- SYSTEM OVERVIEW
- TRANSCIEVER UNIT
 - FSK MODULATION
 - OPTICAL MODULES
- CONTROL/REMOTE UNITS
 - 488 BUS EXTENSION USING DTMF
- INTERFACE UNIT
 - 6 SIGNAL TYPES IN
 - 5 SIGNAL TYPES OUT



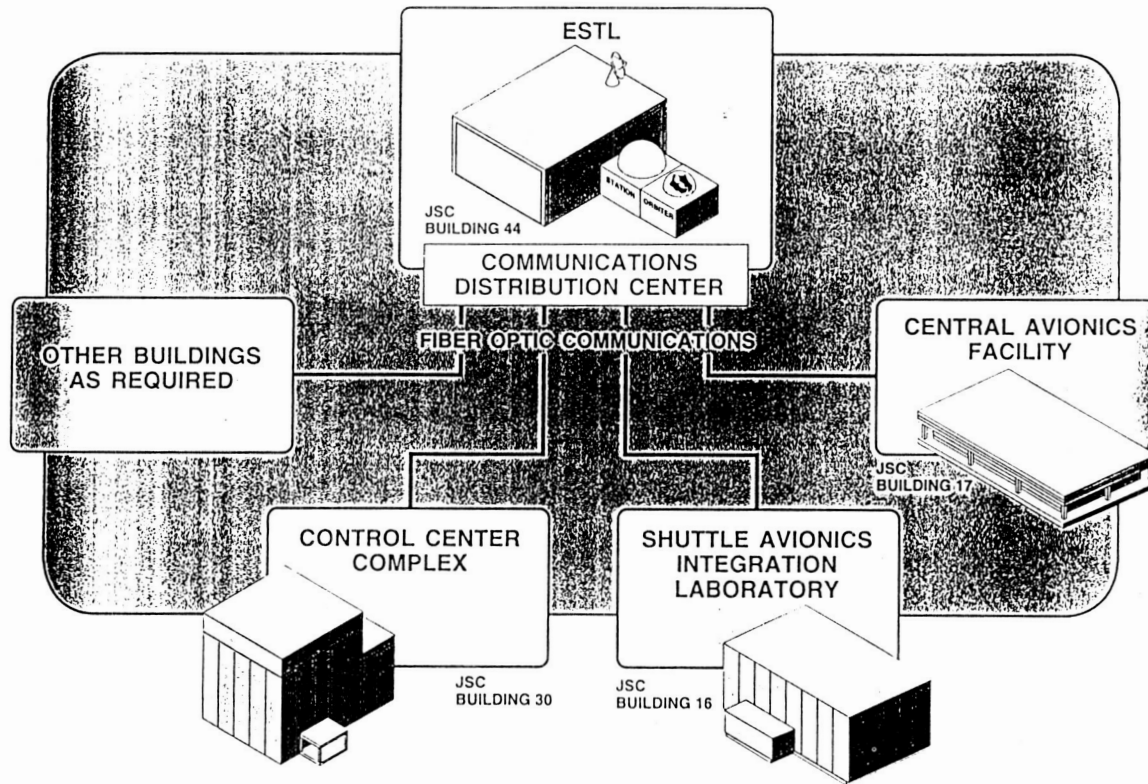
**ELECTRONIC SYSTEMS TEST LABORATORY
(ESTL)**


- COMMUNICATIONS SYSTEMS TEST FACILITY AT JSC
- END-TO-END TESTS (USUALLY INVOLVES ONE OR MORE RF PATHS)
- ANALOG AND DIGITAL SIGNALS
 - VOICE
 - TELEVISION
 - DIGITAL SIGNALS WITH/WITHOUT CORRESPONDING CLOCKS

**ESTL/TDRSS INTERFACE
TEST CONFIGURATION**



ESTL FIBER OPTIC INTERFACES FOR INTEGRATED TESTING



 **Lockheed**
Engineering & Sciences Company

ESTL HIGH RATE OPTICAL COMMUNICATIONS SYSTEM

DATE:

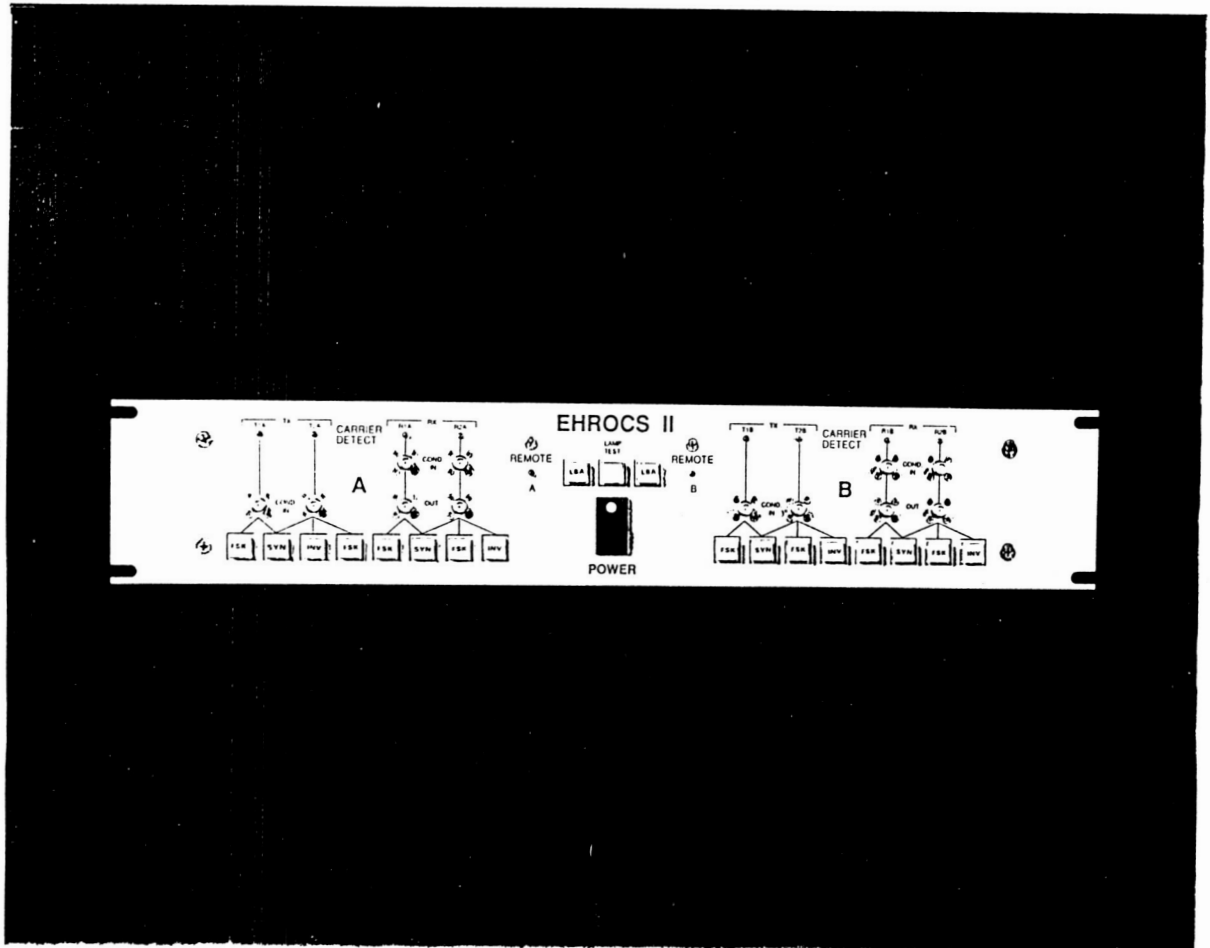
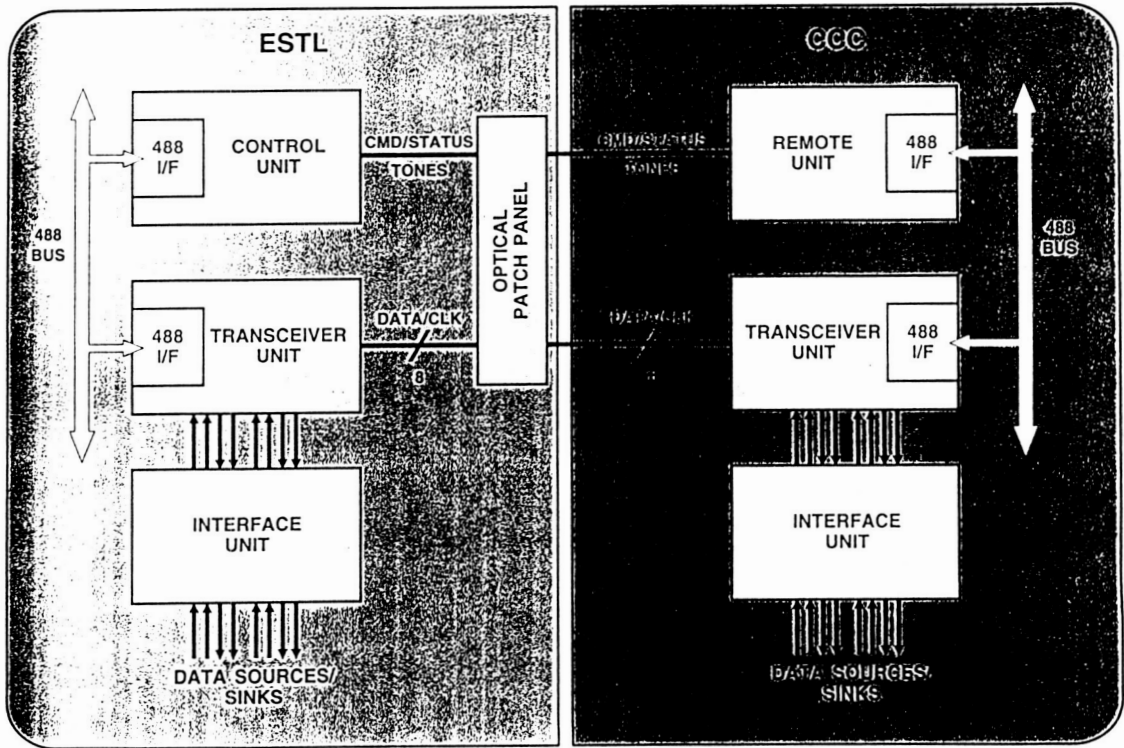
2/1/94



SECOND-GENERATION ESTL HIGH RATE OPTICAL COMMUNICATIONS SYSTEM (EHROCS II) OVERVIEW

- SINGLE-MODE (1300nm)
- FOR DIGITAL SIGNALS ONLY
- DATA AND CLOCK SIGNALS TRANSMITTED POINT-TO-POINT
- PARTY-LINE OPERATION FOR CONTROL AND STATUS
- BROADBAND
 - NORMAL (BASEBAND) RESPONSE IS DATA - DEPENDENT
< 10 KBPS TO > 350 MBPS
CANNOT BE USED FOR BURST DATA
 - FSK RESPONSE IS NOT DATA - DEPENDENT
DC TO 100 MBPS
CAN BE USED FOR BURST DATA
- SYNCHRONOUS OR ASYNCHRONOUS OPERATION

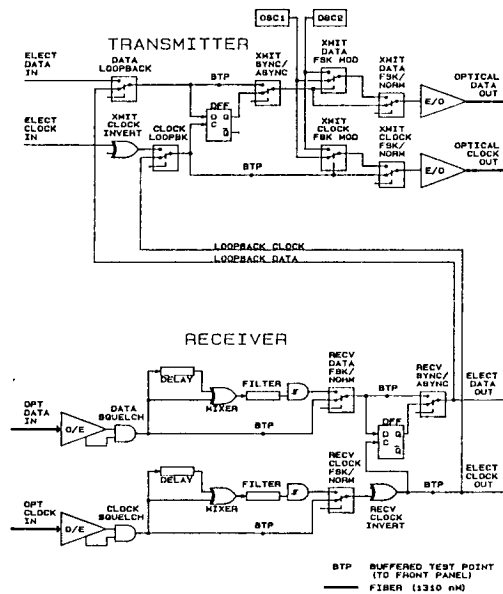
ESTL/CCC FIBER OPTIC INTERFACE CONFIGURATION





EHROCS II TRANSCEIVER UNIT

- 4 TRANSMITTER AND 4 RECEIVER CHANNELS
- CONTROLLABLE SYSTEM PARAMETERS
 - LOOPBACK
 - SYNCHRONOUS OR ASYNCHRONOUS OPERATION
 - CLOCK INVERSION
 - FSK OR NORMAL OPERATION
- MONITORED PARAMETERS
 - ALL CONTROLLABLE PARAMETERS
 - OPTICAL SIGNAL PRESENT
- LOCAL CONTROL OF TRANSCEIVER UNIT OPTIONAL



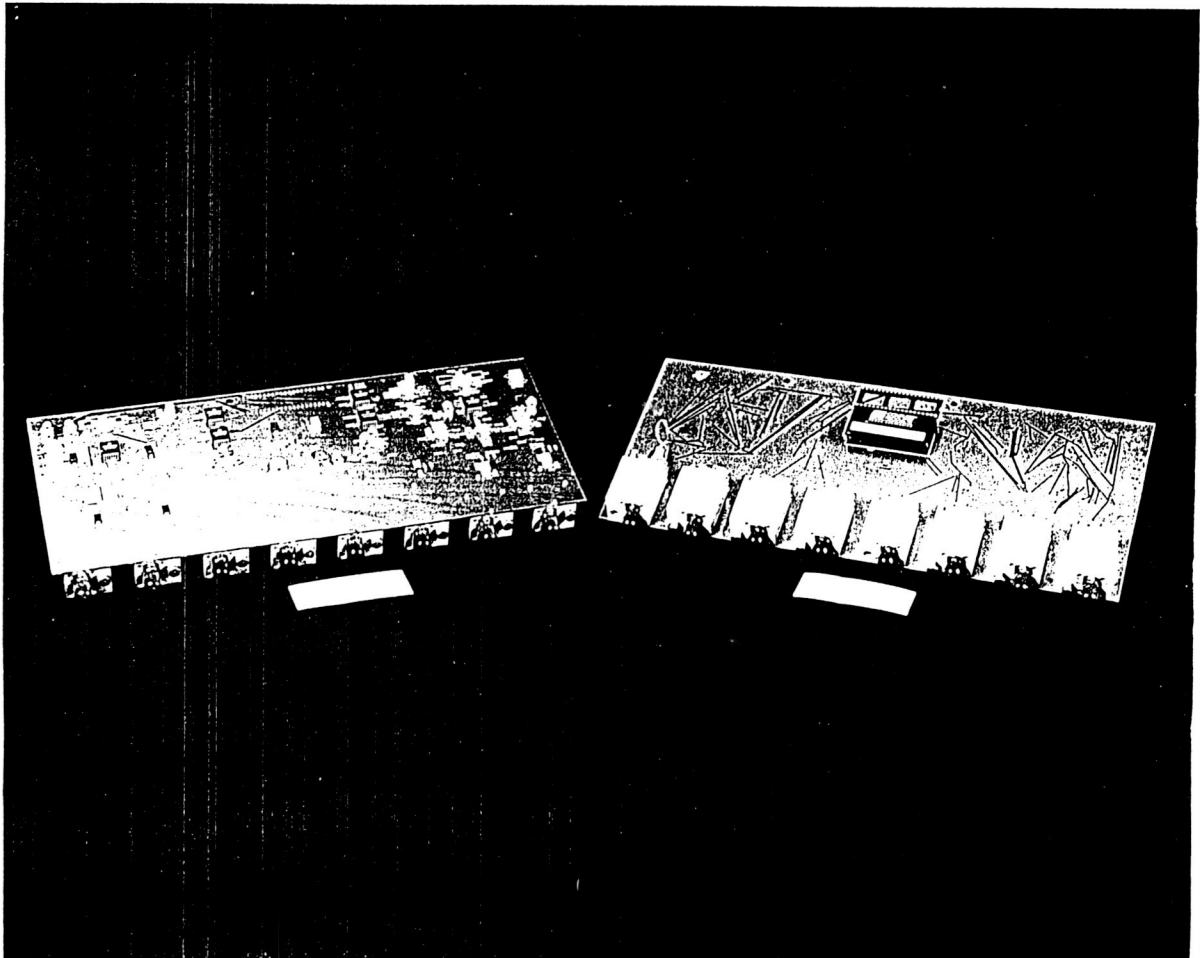
EHROCS II TRANSCEIVER UNIT
FUNCTIONS

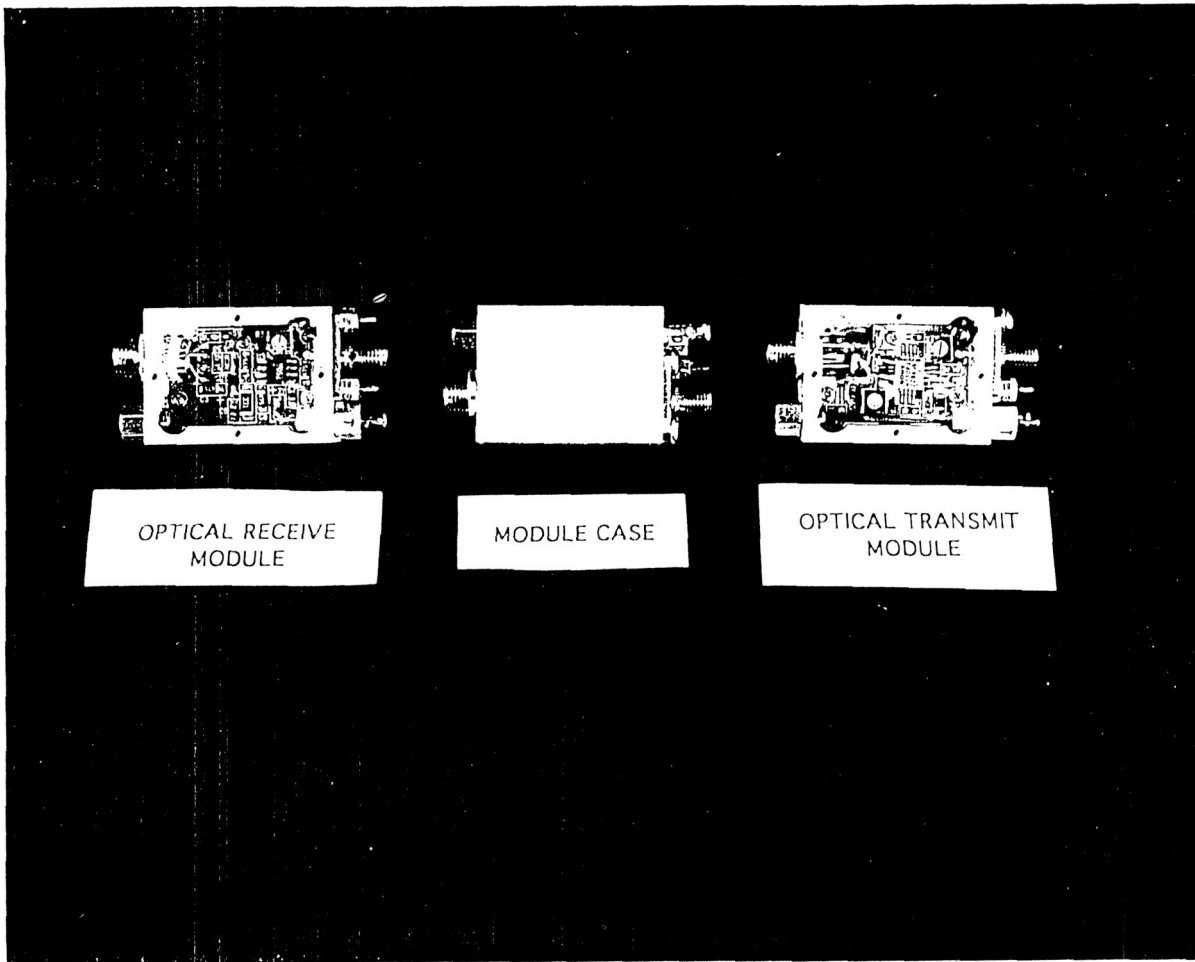
ENOCBDU



FREQUENCY - SHIFT KEYED MODULATION

- SWITCHED - OSCILLATOR MODULATOR
 - NO FREQUENCY DRIFT
 - WIDE DEVIATION
- DELAY-LINE DISCRIMINATOR
 - 19-INCH COAXIAL CABLE FOR DELAY
 - EXCLUSIVE OR USED FOR PHASE DETECTOR
 - NO OUTPUT DRIFT
- FILTER AND SCHMITT TRIGGER
 - SIMPLE ADJUSTMENT
 - LOGIC LEVEL OUTPUT





OPTICAL RECEIVE
MODULE

MODULE CASE

OPTICAL TRANSMIT
MODULE

 **Lockheed**
Engineering & Sciences Company

ESTL HIGH RATE OPTICAL COMMUNICATIONS SYSTEM

DATE:

2/1/94



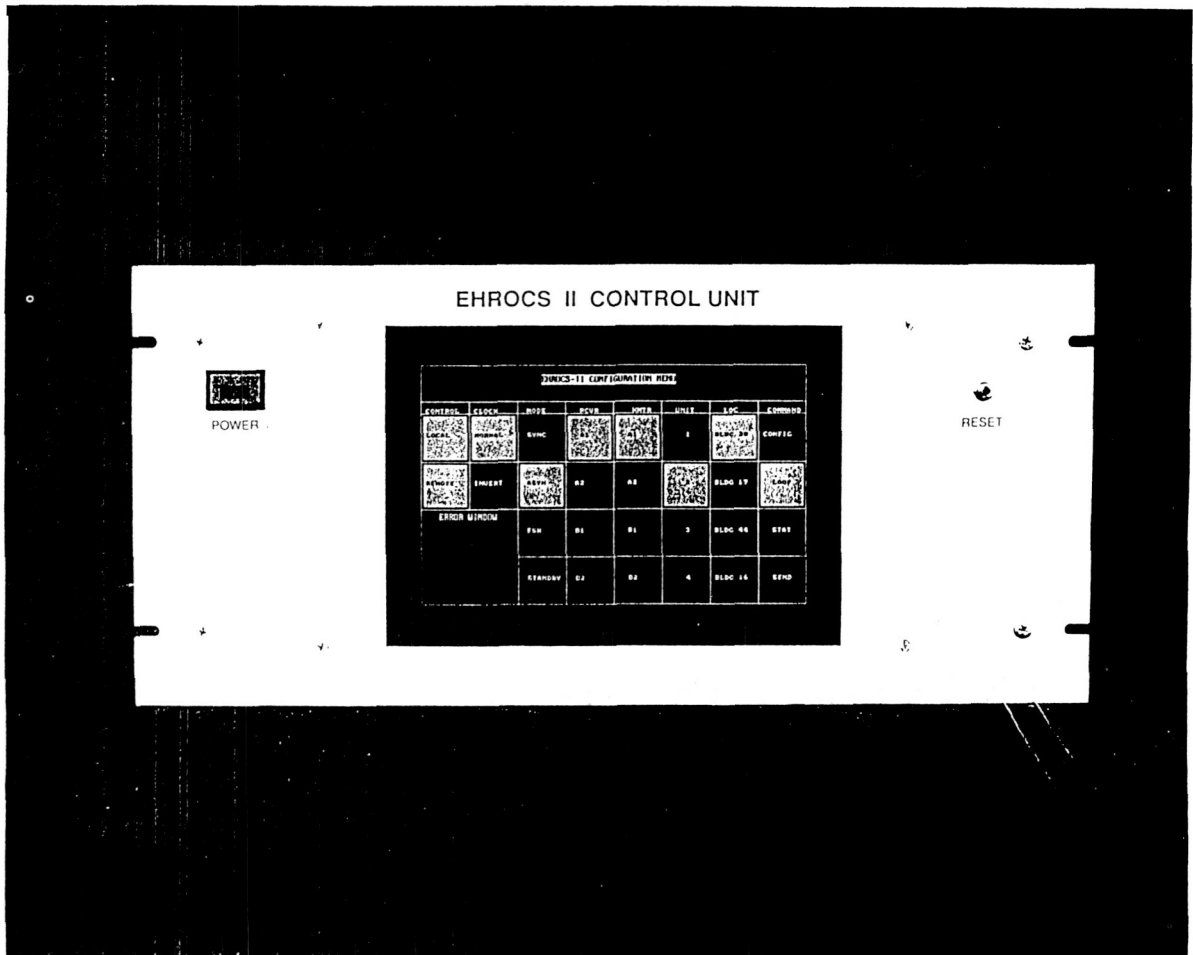
OPTICAL MODULES

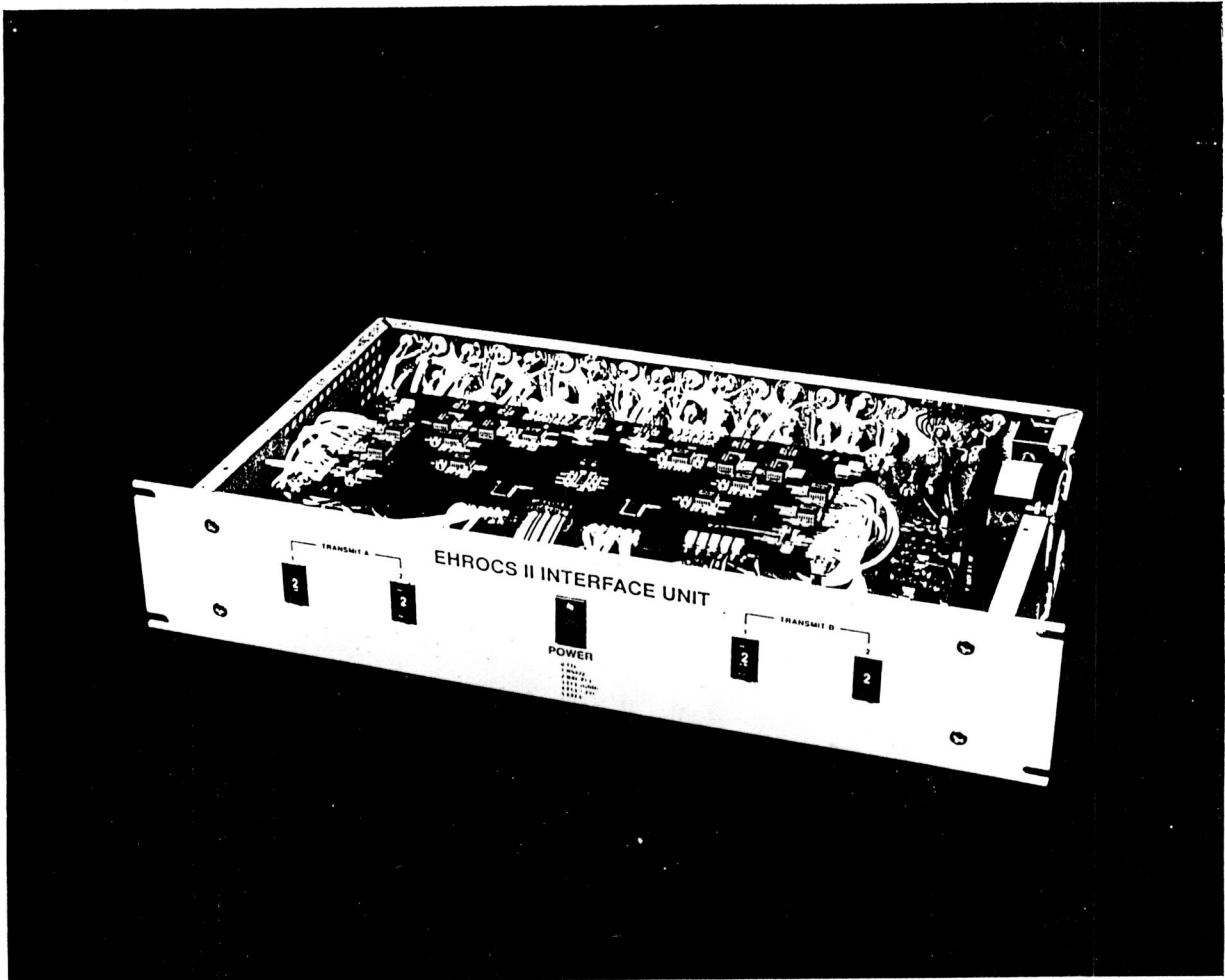
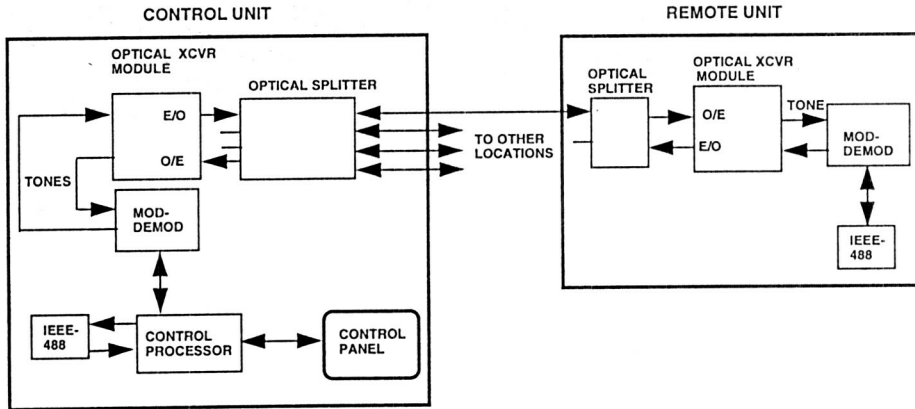
- TRANSMITTER: LASER DIODE
- RECEIVER: PIN PHOTODIODE, PREAMPLIFIER PACKAGE
- DYNAMIC RANGE: > 12 dB (RANGE TO 25 km)
- MODULE ENCLOSURE
 - MOUNTED IN ALUMINIUM BLOCK
 - MECHANICAL AND ELECTRICAL ISOLATION
 - MECHANICAL AND ELECTRICAL STABILITY
 - REDUCED NOISE



EHROCS II CONTROL AND REMOTE UNITS

- CONTROL AND STATUS CHANNEL
 - TRANSMITTER: EDGE-EMITTING DIODE
 - RECEIVER: InGaAs PHOTODIODE
 - DYNAMIC RANGE: >15 dB (RANGE > 25 km)
- IEEE-488 BUS EXTENSION/BRIDGE (IN EXCESS OF 25km)
 - DUAL TONE MULTI-FREQUENCY (DTMF) TRANSMISSION
 - TRANSMISSION IS MICROPROCESSOR CONTROLLED
- CONTROL AND STATUS AS NEEDED
 - CONTROL AND STATUS OF ALL TRANSCEIVER FUNCTIONS
 - TOUCH SCREEN USER INTERFACE





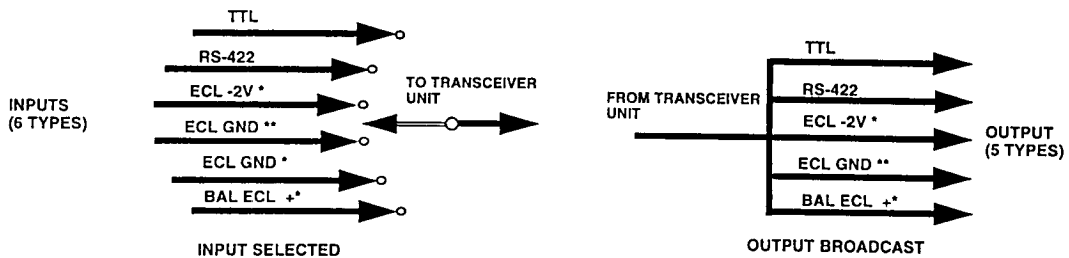


**EHROCS II
INTERFACE UNIT**

- PROVIDES ALL INTERFACES (ELECTRICAL) TO USERS
- ONE I/F UNIT USED WITH EACH TRANSCEIVER UNIT
- FOUR TRANSMIT AND FOUR RECEIVE CHANNELS PER UNIT
- SIX INPUT SIGNAL TYPES
- FIVE OUTPUT SIGNAL TYPES



EHROCS II INTERFACE UNIT



* MINUS VOLTAGE SIGNAL
** PLUS VOLTAGE SIGNAL
+ 78-OHM LINE-TO-LINE TERMINATION



EHROCS II SUMMARY

- FSK OPERATION FOR LOW FREQUENCY AND BURST DATA
- MODULAR OPTICAL TRANSMITTERS AND RECEIVERS
- REMOTE CONTROL AND STATUS MONITORING

FOR MORE INFORMATION OR TO SEE
EQUIPMENT IN THE ESTL:

CONTACT LINDA BROMLEY
(713) 483-0129

28-32
11.4

A Mobile Communications Space Link Between the Space Shuttle Orbiter and the Advanced Communications Technology Satellite

Patrick Fink, Dr. G.D. Arndt, Dr. P. Bondyopadhyay,* Roland Shaw**
Johnson Space Center/EE

* Lockheed Engineering and Sciences Company

** Shason Microwave Corp.

Abstract

A communications experiment is described as a link between the Space Shuttle Orbiter (SSO) and the Advanced Communications Technology Satellite (ACTS). Breadboarding for this experiment has led to two items with potential for commercial application: a 1-Watt Ka-band amplifier and a Ka-band, circularly polarized microstrip antenna. Results of the hybrid Ka-band amplifier show gain at 30 dB and a saturated output power of 28.5 dBm. A second version comprised of MMIC amplifiers is discussed. Test results of the microstrip antenna subarray show a gain of approximately 13 dB and excellent circular polarization.

Introduction

A communications link between the Space Shuttle Orbiter (SSO) and the Advanced Communications Technology Satellite (ACTS), with potential commercial application for developed hardware, is described. We have proposed to provide a low data rate return link from the SSO through ACTS to the Johnson Space Center (JSC). The work has evolved from a NASA/HQ request for communications experiments that utilize the ACTS.

The objectives of the Orbiter/ACTS flight experiment (OAFE) are: (1) to demonstrate the utility of low-earth orbit to geostationary orbit (LEO-GEO) satellite relays at Ka-band; (2) to demonstrate Ka-band monolithic microwave integrated circuit (MMIC) and hybrid solid-state technologies in a space environment; and (3) to demonstrate a moderate gain antenna using computer steering in the dynamic environment of the SSO.

The uplink from the SSO to the ACTS consists of a phased array in the payload bay transmitting low bit-rate data at 29.5 GHz. The Orbiter payload bay is pointed to the ACTS within 10°, and the array points, open-loop, to ±8° in steps of 1.3°. The downlink consists of 20 GHz transmission from the ACTS to a 1.2 m dish antenna at JSC. Various hardware has been prototyped and will be integrated for a field simulation. The Ka-band amplifier and antenna, described below in greater detail, have potential for commercial application.

Ka-band Power Amplifier

For voice rates, the antenna aperture will consist of 16 subarrays in a 4x4 configuration. Each subarray will be flush-mounted to the end of a 1-Watt power amplifier to provide an estimated EIRP of 32 dBW. In addition to amplification, the amplifier module also incorporates a phase-shifting function by utilizing a 4-bit MMIC phase shifter as the first stage. The first prototype, which has been completed and tested, is a hybrid GaAs FET amplifier. A second prototype is being constructed of MMIC

amplifiers that have recently become available, and will be discussed subsequent to the hybrid version.

Performance.

The hybrid amplifier consists of 7 cascaded stages followed by a coupled output stage. Matching circuits for each stage were implemented on a 0.010" quartz substrate. The results are displayed in Table 1. The output power fell short of the desired 1-Watt due to the limited gain of the output stages and the 2 to 2.5 dB combined loss of the branchline coupler and connector.

Table 1. Amplifier Goals and Results.

	Goal	Result
Gain	30 dB	30.5 dB
Bandwidth	500 MHz	760 MHz
1 dB Compression	28 dBm	25.5 dBm
Saturated Power	30 dBm	28.5 dBm

The gain at small signal levels and at 1.5 dB compression is shown in Fig. 1. The ripple at small signal levels is a result of tuning at compressed levels for maximum saturated output power. The input return loss is shown in Fig. 1b along with the 2 dB compression gain and power (-17.6 dB), and the output return loss is displayed in Fig. 1c.

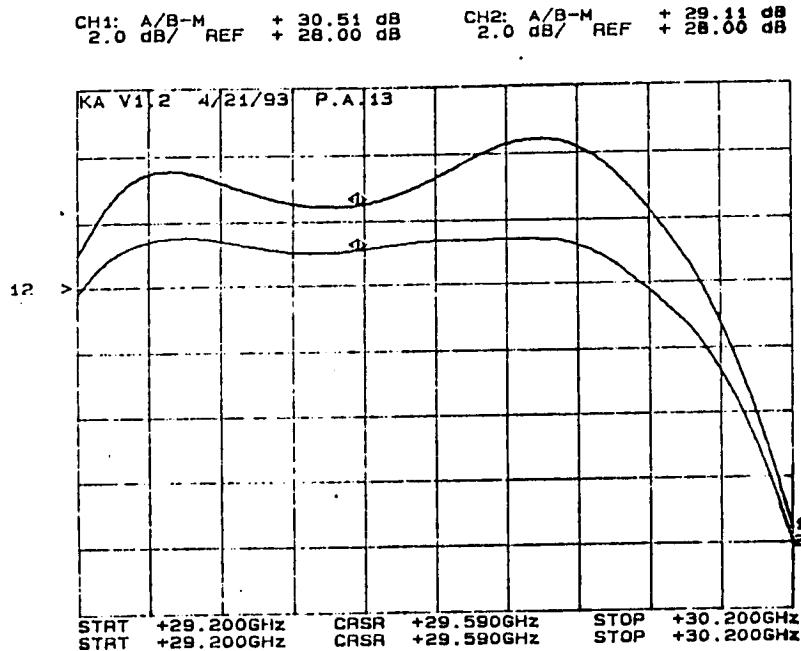


Fig. 1. Small Signal Level and 1.5 dB Compression Gains

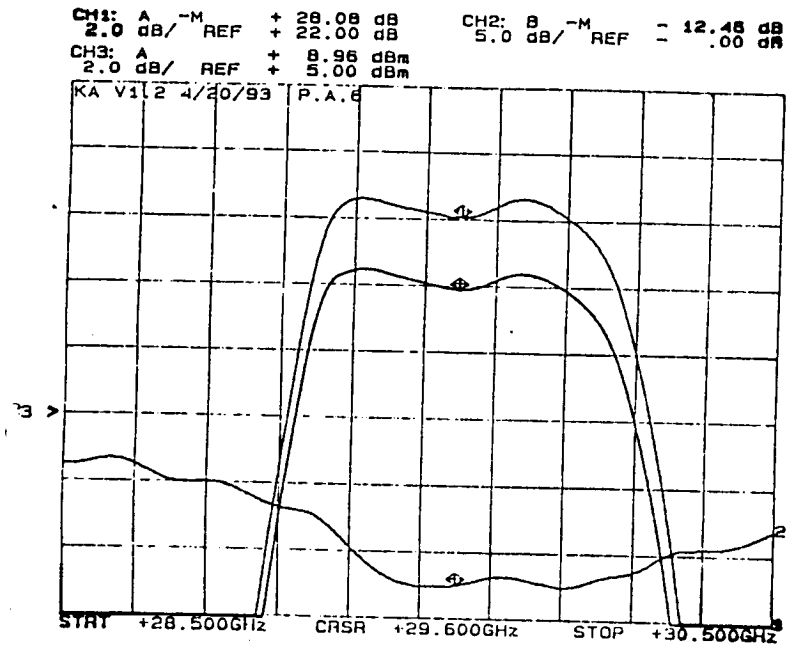


Fig. 1b. Ch1. 2dB Compression Gain.
 Ch2. Input Return Loss.
 Ch3. Power (-17.6 dB).

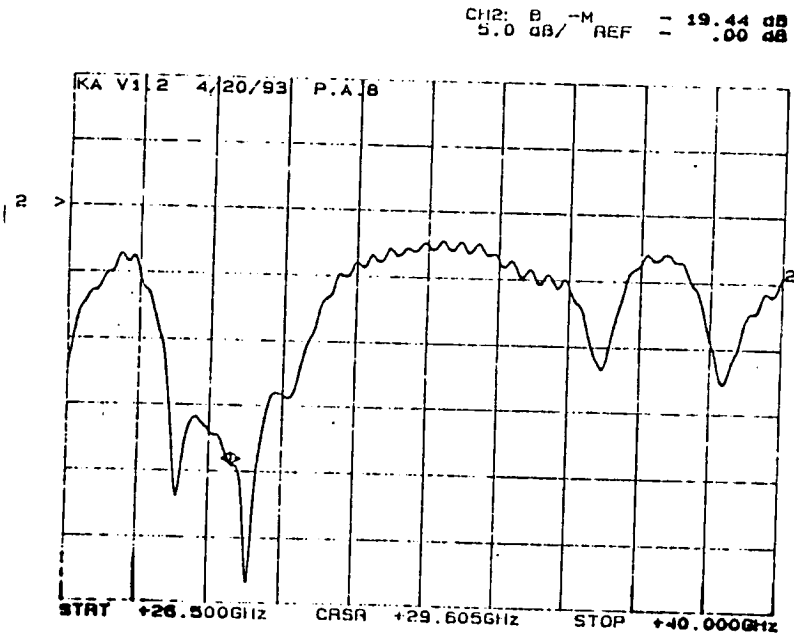


Fig. 1c. Output Return Loss.

Physical Characteristics.

This first prototype was built in a housing that is approximately 8"x1"x1" (Fig. 2). The flanges on the ends of the amplifier serve dual purposes. The deletion of the input flange on the flight unit will allow for a front-loading array structure. For terrestrial application, the dual flanges secure a heat sink around the amplifier. The output flange also provides a surface on which the antenna carrier plate may be flush mounted. A rectangular slot on this end accepts either an SMA connector or a female-female interconnect for flush-mounting.

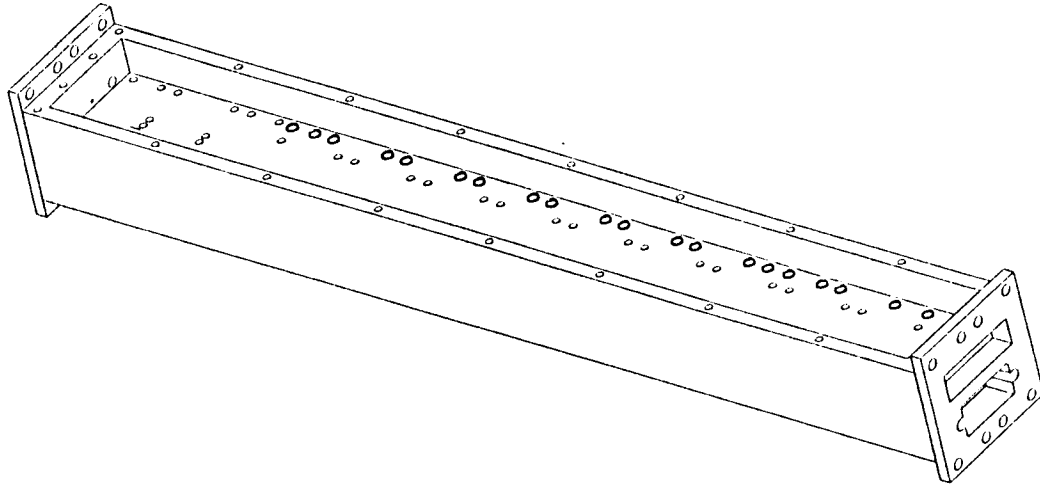


Fig. 2. Ka-band Amplifier Housing.

The required gain was achieved with fewer stages than allowed by the housing, so the next revision has, in addition to enhanced thermal properties, a shorter length. A mating heat sink with a built-in fan was also constructed for terrestrial application. The heat sink supports regulators while the bottom chamber of the amplifier houses the DC bias circuitry.

The power added efficiency of the amplifier is 1.9% at 1 dB compression and peaks at 4%. The need for a higher output power and better efficiency motivated the development of a MMIC version of the amplifier.

MMIC Amplifier.

A MMIC version of the above amplifier is being constructed in the same housing design as the hybrid. The MMIC amplifier will consist of one driver MMIC with 18 dB of gain, followed by a single power MMIC and then a coupled stage of power MMIC's. The power MMIC's have a 1 dB compression point of about 27 dBm and a gain of 11 dB. In addition to a higher output power, better efficiency, and wider bandwidth (approximately 29-34 GHz), the module size for this amplifier can be reduced significantly.

Ka-band Antenna

A circularly polarized microstrip antenna subarray was also developed to mount on the end of the 1-Watt amplifier. As mentioned above, the antenna will consist of 16

subarrays for a total of 256 elements. Each subarray consists of a 2x2 array of a 4-element cluster designed as a single-feed, circularly polarized microstrip compound radiating element. A 4-bit phase shifter in each of the 16 subarrays will provide electronic scanning in a conical space of approximately 10°.

This design of the compound radiating element employs dual-feed, square microstrip patch elements in conjunction with sequential feeding and rotation technique to enhance the axial ratio bandwidth. Impedance and axial ratio bandwidths of over 1 GHz at the center frequency of 29.5 GHz are very easily achieved. The 16-element subarray and the associated radiation pattern are shown in Fig. 3 and Fig. 4, respectively.

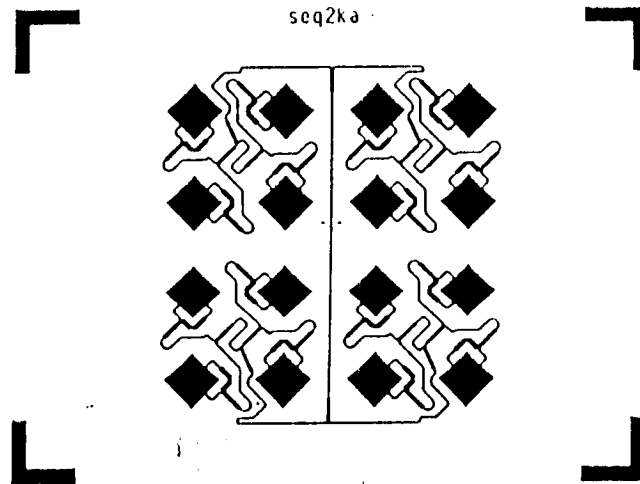


Fig. 3. Sixteen Element Subarray.

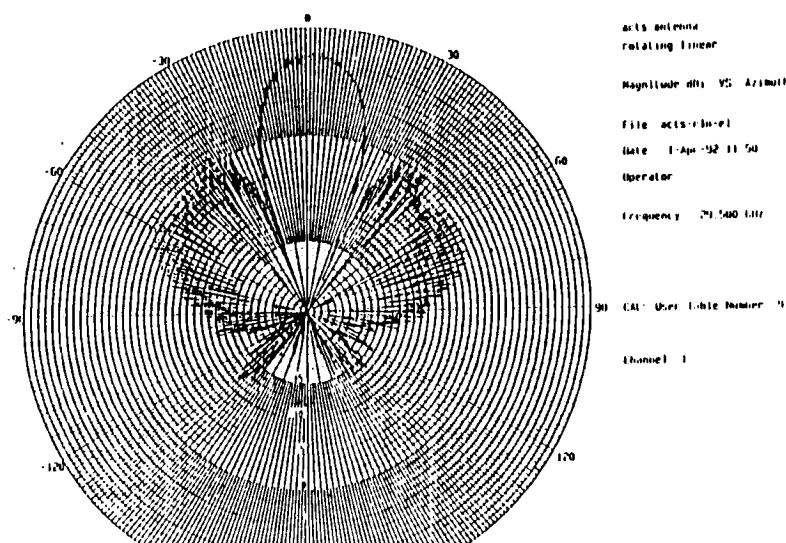


Fig. 4. Subarray Radiation Pattern.

A second version of the subarray is being developed. This version is expected to have a lower feedline loss but a higher axial ratio than its predecessor. As seen in Fig. 5, the feed network is not as intricate as the earlier version.

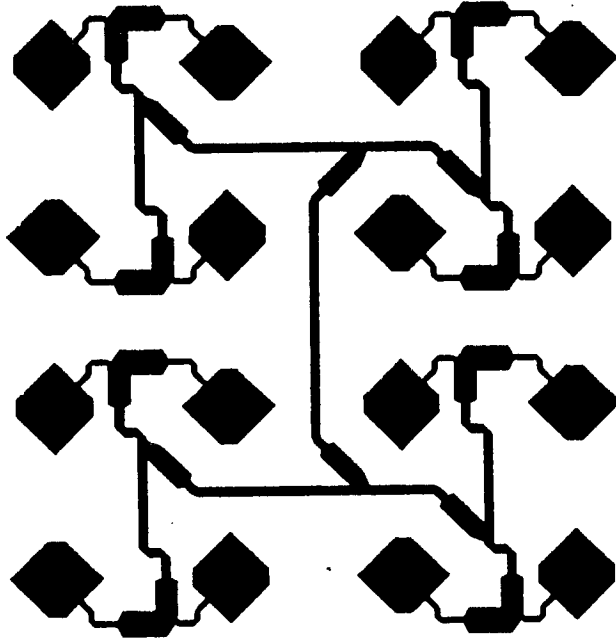


Fig. 5. Ka-Band Subarray: Version 2 (Lower Feedline Loss).

Commercial Application

The transceiver being developed for the OAFE utilizes state-of-the-art Ka-band technology. This work is greatly beneficial in opening the millimeter region of the spectrum for commercial use. For example, the absence in the commercial arena of a Ka-band solid state power amplifier necessitated in-house development by JSC. The advent of components like these is leading to the miniaturization of communications technology and at the same time to the broadening of the usable spectrum. This expansion is evident in the recent FCC authorization of the 27-30 GHz band for point-to-point video broadcast applications.

Conceivable commercial applications of the microstrip antenna developed include scenarios demanding moderate gain (< 33 dB), low profile, and arbitrary attitude. This description is exactly that depicted by airline communications, where low profile antennas are required, and control of the antenna polarization vector is not possible. The employment of this antenna in a briefcase-size portable terminal is also being investigated.

TDMA ALGORITHM for the SPACE to SPACE COMMUNICATIONS SYSTEM

Keith Adam Gerhards
Lockheed Engineering and Science Company
Houston, Texas 77058

ABSTRACT

Currently, a five user Time Division Multiple Access (TDMA) communications system, based on a fixed frame architecture and no dedicated master station, is being developed in the Electromagnetic Systems Branch of the Tracking and Communications Division. This Space to Space Communications Subsystem will be utilized on the Space Shuttle, Extravehicular Mobility Unit (EMU), and the Space Station.

Each communications system is composed of an RF transceiver, a MODEM, and a Baseband Signal Processor (BSP). The RF transceiver provides the upconversion and downconversion from the UHF band to an Intermediate Frequency (IF) of 21.4 Mhz. The MODEM demodulates the 21.4 Mhz IF to baseband data that has a bit rate of 512 Kbps. The baseband data is then processed by the frame synchronizer algorithm and sent to the DSP. The DSP decommutates the data to provide continuous 64 Kbps mu law digital voice, 4 Kbps analog data, and 8 Kbps digital data. The DSP also stores the continuous data coming from the user for future burst transmission.

This paper discusses in detail the design, implementation, and test results of the Modulator-Demodulator (MODEM) portion of the proposed radios. This paper will also elaborate on laboratory test results including bit error rate performance, frame lock, false lock, cochannel interference and receiver sensitivity. Also bit timing, bit threshold, and acquisition time are discussed.

INTRODUCTION

Communication between the Extra Vehicular Activity (EVA) astronauts and the Orbiter is currently established using a VHF amplitude modulated (AM) system. The biomedical, pressure, and suit data are transmitted via a subcarrier that is frequency modulated. The system will accommodate 2 EVAs simultaneously and more EVAs in simplex mode. The VHF frequency band used by the present system is a restricted band, and as such, there is a requirement to move to a new band. In addition to moving to a new frequency band, other recommendations were made to improve on the current EVA radio.

The first and foremost improvement is the upgrade to full duplex communication between four EVAs and the Orbiter without any master control station. This improvement allows the EVAs to communicate without Orbiter control and also provides added users to the network. The next improvement is in the electronic packing of the EVA radio itself.

In order to service the current EVA radio electronics, the radio must be first dissected of its multiple interconnects and modules before any of the electronics can be exposed. The positioning of some of the modules in the current radio requires a special tool to be inserted to remove the cabling. To improve on the old design, a complete modular approach was conceived. The RF module would have its own interconnect support and the remaining circuitry would be in a card format interconnect through a common bus.

As a final improvement to the current communications system, a digital data stream was added to the network to provide communications between the Orbiter, EVAs, and the Space Station. The primary users for the data channel are the Orbiter and the Space Station but the EVAs can also access the information.

As mentioned earlier, each radio is made up of an RF transceiver, a MODEM, and a Baseband Signal Processor (BSP). The main focus of this paper will be on the MODEM section of the radio.

MODULATOR DEMODULATOR DESIGN

There are three types of data to be received and transmitted by each user in the system: digital voice, status data, and digital data. All these data forms are encapsulated into an eight millisecond frame format that is divided among five users. In addition to the three types of data, there are preamble and unique word bits that are transmitted before each packet of data. The preamble is a 64 bit binary data stream of alternating "1"s and "0"s that is used by the DC offset estimator and bit synchronizer to extract statistical and timing data from each burst reception. The unique word is a 32 bit pattern that enables the receiver to identify the beginning of the three types of data and is also used to update local frame timing.

The frame description is shown in Appendix A and displays each user's data format in the eight millisecond frame and also the burst transmission calculation. The burst transmission calculation shows that each user must transmit at a 512 Kbps rate for 1.4 milliseconds every eight milliseconds in order for other users to receive valid data in a continuous fashion. The valid data is composed of 608 bits, 512 of which are mu law encoded digital audio, 64 are digital command data, and 32 are status data. Each user transmits the packet of framed information to other users using a single RF channel. The transmission occurs by modulating the packet data to an IF frequency and then upconverting the IF signal to a UHF transmission band.

Choosing a specific type of modulation technique for this system involved tradeoffs with power consumption, acquisition time, bandwidth, and complexity. The acquisition time for the modulation scheme is especially critical since each user is transmitting their information in 1.4 millisecond bursts. The acquisition time of any modulation scheme chosen must provide a received signal to the demodulator before the beginning of valid data is received. The most effective scheme of modulation is a noncoherent type that does not require the carrier of the received signal to be extracted. For a noncoherent modulation scheme, a Continuous Phase Frequency Shift Keying (CPFSK) digital modulation method has been proven to be the most efficient. The three criteria for an effective CPFSK are to filter the received signal by 1.4 times the data rate, to filter the demodulated data at the data rate, and to use a modulation index of 0.7 times the data rate. It has been shown that with these constraints in the system, a bit error rate of $10 \text{ E-}5$ can be achieved for a bit power per noise density of 12.7.

The MODEM handles all of the modulation, demodulation, and frame timing necessary to communicate between the users. The MODEM is composed of two analog sections and five digital sections. The first analog section is used to process the demodulated data from the RF transceiver and the other analog section is used to send modulated IF to the RF transceiver. The digital sections are used for DC offset estimation, bit synchronization, frame synchronization, automatic user identification, and baseband data transfer. These seven key MODEM components will be discussed in detail in this paper.

Analog Processing of Baseband Data

The analog processing section of the MODEM interfaces with the transceiver analog data that is being demodulated. The transceiver downconverts the received UHF signal to an IF frequency that is then sent to the demodulator. The sensitivity of the transceiver is specified to be -90 dBm. The demodulator is an FM discriminator that uses a limiter and a

quadrature detector to provide the demodulation of the received CPFSK signal. The sensitivity of the demodulator is specified at -60 dBm. The 30 dB of gain is provided by the transceiver Low Noise Amplifier (LNA). The discriminator's analog output contains a DC component that varies with temperature and frequency offset. The temperature variation of the radio actually changes the quadrature capacitor and inductor values which then change the slope of the discriminator. The frequency offset is caused by the differences in the local oscillators of the transmitting and receiving radios with the maximum offset expected to be 10 KHz. Both of these factors add a 0.5 volt uncertainty to the analog data. The DC offset component is tracked in the digital controller section of the MODEM. A seven pole Low Pass Filter (LPF) with a cut off at 512 KHz is used to first filter the analog data from the transceiver to eliminate any aliasing signals to a flash Analog to Digital Converter (ADC).

The sample rate of the flash ADC is set at 4.096 Mhz which is 16 times faster than the received analog data rate. The LPF before the flash ADC ensures that no analog signal above 512Khz is being sampled. Therefore, the digital spectrum at the output of the flash ADC should be the main signal with a bandwidth of 1.024 Mhz repeated every 2.048 Mhz. Since the LPF is not ideal, some noise beyond the LPF cutoff is getting sampled by the flash ADC. A digital Infinite Impulse Response (IIR) filter is used to eliminate that noise. The transfer function of the IIR filter is given by the following equations.

$$H(Z) = \frac{0.25}{(1 - 0.5 Z^{-1})^2} \quad (1)$$

$$H(e^{-j\omega}) = \frac{0.25}{(1 - 0.5 e^{-j\omega})^2} \quad (2)$$

$$|H(e^{-j\omega})|^2 = \frac{(0.25)^2}{|(1 - e^{-j\omega} + 0.25 e^{-j2\omega})(1 - e^{-j\omega} + 0.25 e^{-j2\omega})|} \quad (3)$$

$$|H(e^{-j\omega})|^2 = \frac{0.25}{\sqrt{(1 - \cos(\omega) + 0.25 \cos^2(2\omega)) + (-\sin(\omega) + 0.25 \sin^2(2\omega))}} \quad (4)$$

The Z^{-1} term denotes a digital delay and has an analog equivalent of $e^{-j\omega}$ for a normalized frequency response. Once transformed into the frequency domain, the transfer function can be solved and the magnitude versus frequency response shown in Figure 1. The plot shows that the digital IIR filter reduces the resolution of the received data stream by 3 dB, or 1 bit of the flash ADC. The normalized frequency in the plot refers to the sample rate of the ADC which is 4.096 MHz. This digitally filtered data is then used by the DC offset estimator for further processing.

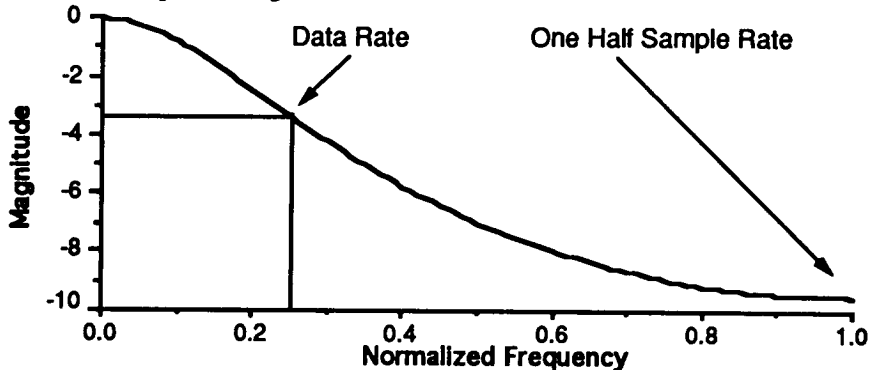


Figure 1. Magnitude Response for IIR Filter

DC Offset Estimation

There are several methods to determine the DC value of the analog signal that is being received from the transceiver. One method is to continuously average the received signal from the IIR filter. This method would require the incoming data to have a 50% duty cycle for the majority of the received data in order to get a somewhat accurate DC estimation. If long dropouts occurred in the data then, the estimated DC value would follow the dropout and produce inaccurate data.

The DC estimation process used for this system is a congruent process whereby the received data format is connected to the DC offset estimation process and the user tracking process. The preamble portion of the packet of received user data is a clock pattern at a frequency of one half the data rate. A complete period of this clock is the perfect data type for DC offset estimation since the incoming data during this period is a sine wave with a DC offset. If the sine wave is sampled at 16 times its period then the DC offset value can be extracted from those 16 samples. The problem is to estimate the DC value during the preamble time and turn off the DC offset estimation after that.

During acquisition, the DC estimation process is continually averaging data from the IIR filter. The estimation of the DC value during preamble time consumes 32 out of the 64 preamble bits, leaving 32 bits for the bit synchronizer to extract a clock. After the preamble, the unique word is received and processed in the DC estimation. The unique word is configured such that it has a maximum transition density so that it minimally changes the DC estimation value. This maximum density in the unique word allows the unique word correlator to still recognize a correlation during the acquisition period. The correlation is used by the frame controller to determine the lock status. Once the frame controller has determined a tracking status, the unique word is used to mark the time at which the user was received. That value is subtracted from the number of bits in the unique word and the preamble plus jitter time to obtain a comparison for the next received frame. The comparison value is used to start and stop the DC offset estimation process. The estimation is only calculated during the first user preamble in the frame. The other user's preamble is not used for DC estimation. Therefore, the DC value that is obtained during the first user preamble is used throughout the entire frame.

The averaging is implemented by first accumulating 16 samples in a 12 bit storage register and then dividing that value by 64 to obtain the new threshold value. The threshold register value is then used to subtract the value of the next 16 filter samples coming from the IIR filter. These next 16 samples are then accumulated, divided, and summed with the current threshold value. Subtracting the threshold value from the current IIR filter value results in an eight bit bipolar value with the most significant bit used for the sign of the value. The sign of the value is then used to make the determination of a zero or a one. This determination is used by the bit synchronizer to extract the bit clock.

Bit Synchronization

The sign value of the DC offset register is used in determining the threshold of the data. The threshold value triggers the clock synchronizer which aligns its rising edge with the rising edge of the data. The bit synchronizer algorithm uses an all digital approach to obtain a synchronized clock. The rising edge of the clock must consistently follow the rising edge of the incoming data. Since the analog data is sampled at 16 times the data rate, a phase differentiation of 45° can be extracted. A timing diagram showing the sample, data, and synchronized clock is shown in Figure 2.

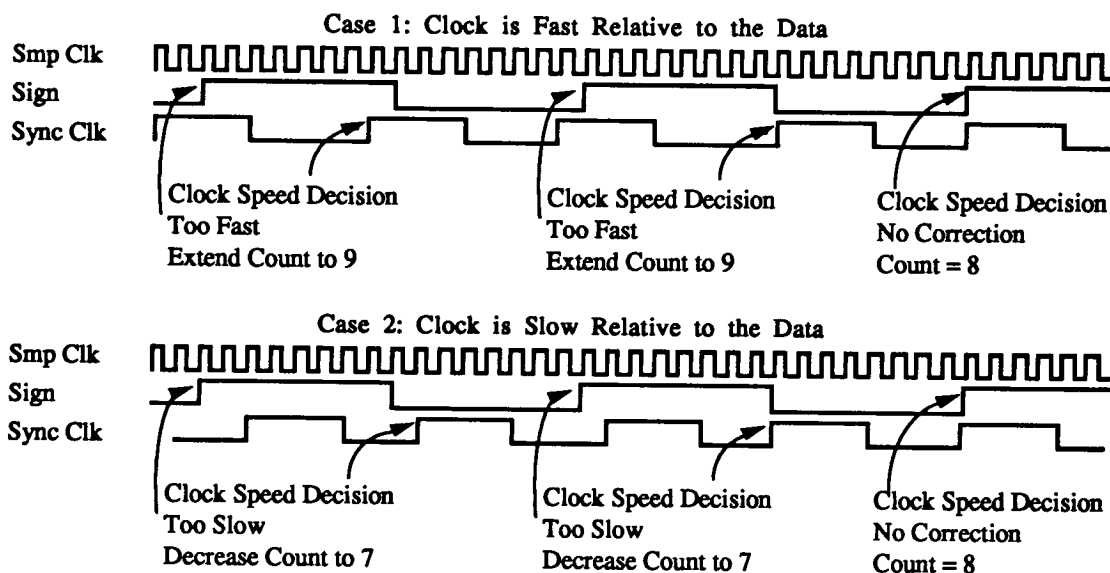


Figure 2. Bit Synchronizer Timing Diagram

Figure 2 shows two cases where the phase of the bit clock is changed. The first example is when the clock is fast relative to the data. In this instance, the rising edge of the data is received after the rising edge of the clock, and therefore the clock must be slowed down to align the two edges. The clock speed is determined by a four bit binary counter that is clocked by the 16 times clock. This counter usually restarts at count seven which provides a 512 KHz clock from the third most significant bit of the binary counter. To slow the 512 KHz clock, the four bit binary counter's restart count is changed from seven to eight. By increasing the restart count, the 512 KHz clock has now changed frequency to 455.11 KHz for one period. This frequency change also changes the phase relationship between the clock and data by + 45°.

The second example demonstrates what occurs when the clock is slow relative to the data. In this case, the rising edge of the data is received before the rising edge of the clock and therefore the clock must be sped up to align the two edges. The clock speed is accelerated by adjusting the restart count of the four bit binary counter from seven to six. The 512 KHz clock is then changed to a 682.67 KHz clock for one period, which also changes the phase relation by - 45°.

The bit synchronizer is first triggered by the transition detection section. When the sign value from the DC offset estimator changes from zero to one, the transition detector sends a pulse to the bit synchronizer. The bit synchronizer has three counters it uses to determine the clock frequency change. Two counters, PREG and NREG, keep track of how many transitions occurred for both of the cases previously described. When the clock is faster than the data, the PREG is incremented and when the clock is slower than the data the NREG is incremented. When either the PREG or the NREG is incremented, a third counter, the GREG, is incremented. When the GREG reaches a value of two, the PREG and NREG values are compared. If the PREG value is greater, then the clock is sped up; if the NREG value is greater, then the clock is slowed down. Or, if PREG and NREG are equal then all registers are reset, and no change is made to the clock. The effect of the GREG counter is to filter out the transition pulses which provides a less jittery clock. The synchronized clock and data are then sent to the frame controller.

Frame Synchronization

One of the most critical elements of the TDMA algorithm is the strategy used to distribute equal portions of an eight millisecond transmission time frame between each of the 5 radio users. Further, the algorithm must determine precisely when during that frame time each user should be transmitting information.

To accomplish this time allocation, each radio determines where in the frame to transmit based on the vacancy of each time slot. If all the slots are vacant of users, then the radio begins transmitting assuming itself to be the first user that has powered on. If no users have been received for 255 frame periods, then the radio stops transmission to listen for other users during its own transmit time. This collision avoidance scheme provides each user with the ability to initiate a communications link if a previous link has failed, or if no link has been established.

The sequence of events for a radio to lock on to other users begins by first "listening" for 255 frames in an attempt to lock on to any valid users. If two users are already in the system, and they are locked on to each other, then the third radio will lock on to the second user and begin transmitting in the next open slot. The third user locks on to the second user by adjusting a 12 bit frame counter which continuously counts the number of bits in the 8 millisecond frame - nominally 4096 bits. The frame counter clock period is adjusted in the same manner as the bit synchronizer clock to achieve a longer or shorter frame count which allows the radio to adjust its transmit time. In the example of a third user locking onto two previously established users, the frame counter is adjusted until the time at which transmission by the third radio occurs is positioned before the first user transmits. To user one, user three looks like he is transmitting in the 4th slot and to user two he is transmitting in the 5th slot. A frame diagram showing the three different frame times and their relation to the others is shown in Figure 3.

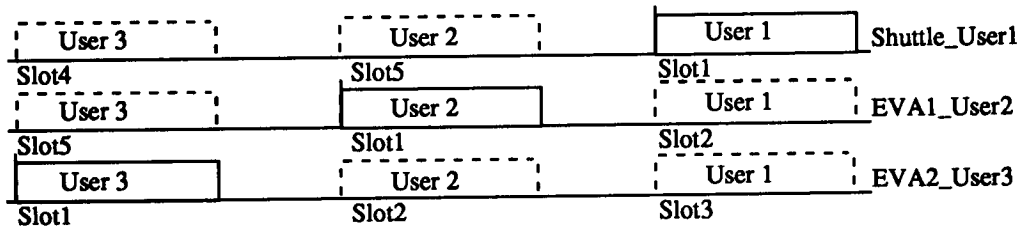


Figure 3. Three User Transmission Relationship

The dashed line boxes in Figure 3 represent received users and the solid line boxes represent transmission. User three acquires the other two users by not allowing any users in his frame time of 0 through 704. If any users are received during this frame time then a frame error of positive 400 is generated. That frame error essentially pushes the received user towards slot five and away from his own transmission time. As an example, consider the case where user three has powered on and received user two at a frame time of 600. User three would add 600 to the 400, divide that sum by two and adjust his frame to slow down by 500 bits. The frame adjustment will change the user two received time from 600 to 100. For this frame time, user three would shift his frame counter by 250 and receive user two on the next round at frame time 3945. At this point user one would be received at frame time 668 and would be pushed back to frame time 134. The next frame error would place user one at frame time 3962 and user two at frame time 3144. Since user two was received before user one, user two will be set to position 915 in the next corrective frames which then places user one in the third slot. User three declares acquisition and enters

tracking mode when the frame error has reached a value of less than nine. The sequence of frames for correction are shown in Appendix B.

Once user three has acquired user two, user three begins transmitting information when the frame counter reaches zero. User three positions user two in slot two and user one in slot three. A table showing different frame counter values for received times for each user in the above system is shown in Table I. A zero in a column represents a transmit time for the user in that row.

Table I. Frame Values for User Receive Times

	User 1	User 2	User3
User 1	0	915	1734
User 2	3373	0	915
User 3	2553	3373	0

For the EVA to Orbiter mode, the different time slots do not represent an identification problem if the EVAs are sequenced during start up procedures. In the airlock, each EVA must turn on one at a time. When the first EVA turns the radio on and locks on to the Orbiter, the Orbiter will receive that EVA in slot five. The first EVA will place the Orbiter in slot two. When the second EVA turns his/her radio on, the algorithm will place EVA one in slot two and the Orbiter in slot three. To the Orbiter, EVA two will be seen in slot four. The rest of the EVAs will power on in a similar manner, and will be placed in slots three and two in the Orbiters frame. If this power on sequence is followed, then each EVA will be identified in the order that they were powered on.

Frame Adjustments

The frame adjustments start to occur when two receptions of the unique word are made with no errors. During the first reception, the frame count value is stored in a buffer to be compared with the current count of the frame counter. On the next frame, a window is generated that is 32 bits wide about that buffer count value. If the unique word is received within this window, then a frame calculation is initiated. Therefore, a frame error is not calculated until the unique word has been received and then verified during the following frame. The unique word is also used to mark the beginning of valid data for the baseband signal processor.

Baseband Data Transfer

The packet data that is being sent and received by the users is first processed by the Baseband Signal Processor (BSP). The BSP commutates the audio, status, and digital data to be transmitted by the MODEM in the frame format discussed earlier. The BSP also decommutates the received packet data from the MODEM to provide real time audio, status, and digital signals to the user. The packet data must first be verified as valid by the MODEM before the data is sent to the BSP. The MODEM verifies that the data is valid by entering the tracking state as described earlier. The MODEM then sends data to the BSP every time a unique word is received in the frame.

TEST RESULTS

An engineering model radio for the EVA that represents the TDMA communications system discussed in this paper was developed in the Electromagnetic Systems Branch of the Tracking and Communications Division. Laboratory tests were performed on the radio including bit error rate performance, frame lock, false lock, cochannel interference and receiver sensitivity.

The Bit Error Rate (BER) performance of the radio was measured at 10 E-5 for a received RF signal level of -91.5 dBm at ambient temperature. The test was conducted with the engineering model radio set to continuous receive mode which allows the BER measurement to be measured at a 512 Kbps rate. At temperature extremes, the RF signal level was increased to -90.5 dBm to achieve the required 10 E-5 BER. The full BER curve versus the theoretical curve is shown in Appendix C. The actual BER performance is 2 dB off the theoretical curve which is mainly due to the noise figure of the receiver.

The acquisition time was measured with the minimum required input level of -90 dBm to the RF receiver. The time to acquire with the stated constraints was recorded to take sixteen frame periods or 128 milliseconds . There is a possibility that the acquisition time of any user will take as much as 12.5 seconds if both users are powered on exactly at the same time. This longer acquisition time is due to the users transmitting on top of each other's time slot until their clock uncertainty slides the users apart.

The minimum signal level that was measured to the RF transceiver where a user was unable to acquire was measured at -99.5 dBm . With no RF signal into the receiver, the frame lock status was monitored to record any false lock indications. After operating four hours, no false locks were detected.

The two main tests conducted for cochannel interference on the engineering model included a wide band and a narrow band interferer. The first test introduced a 99 Khz FM modulated interferer at 20 dB above the minimum level of the main signal. The interferer was swept over a 3.5 MHz frequency range on either side of the main signal. The result was a degradation in BER performance once the interferer was inside the maximum rejection band of the IF bandwidth of the receiver, which is 1.5 Mhz . The second cochannel interference test involved a 1 Khz FM modulated interferer at the main signal frequency with a varying power level. This interferer test showed that the main signal to interferer ratio must be greater than 15 dB to maintain a 10 E-5 BER.

The receiver quieting sensitivity was measured using an external frequency synthesizer and a spectrum analyzer to measure the IF amplitude from the RF transceiver. The 20 dB quieting FM receiver sensitivity was measured to be -85 dBm .

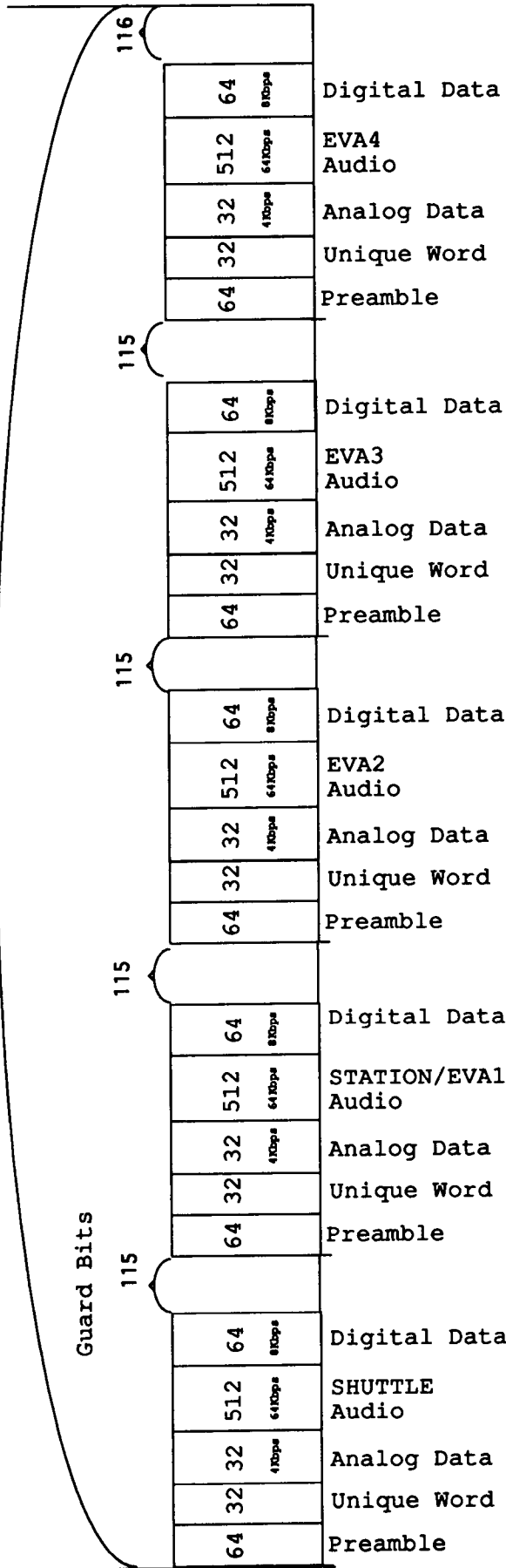
Bibliography

- [1] Wong, Alvin, "Circuits for Wide-band FM Demodulation," RF Design, Englewood, Co. ,December, 1990, pp. 27-34.
- [2] Richley, Edward, "Design of Quadrature Detectors," RF Design, Englewood, Co. ,May, 1991, pp. 68-72.
- [3] Tjhung, Theng and Wittke, Paul, "Carrier Transmission of Binary Data in a Restricted Band," Transactions on Communication Technology, Vol. COM-18, No.4, August, 1970, pp. 295-304.
- [4] Batson, B., Seyl, J., and Smith, B., "Experimental Results for FSK Data Transmission Systems using Discriminator Detection," National Aeronautics and Space Administration, Johnson Space Center.
- [5] Rohde, Ulrich and Bucher, T.T.N., "Gain, Sensitivity, and Noise Figure," COMMUNICATIONS RECEIVERS PRINCIPLES & DESIGN, ed. 1, Vol. 1, McGraw-Hill, New York, New York, 1988, pp. 65-68.

APPENDIX A

Frame Description

Frame Rate = 8ms



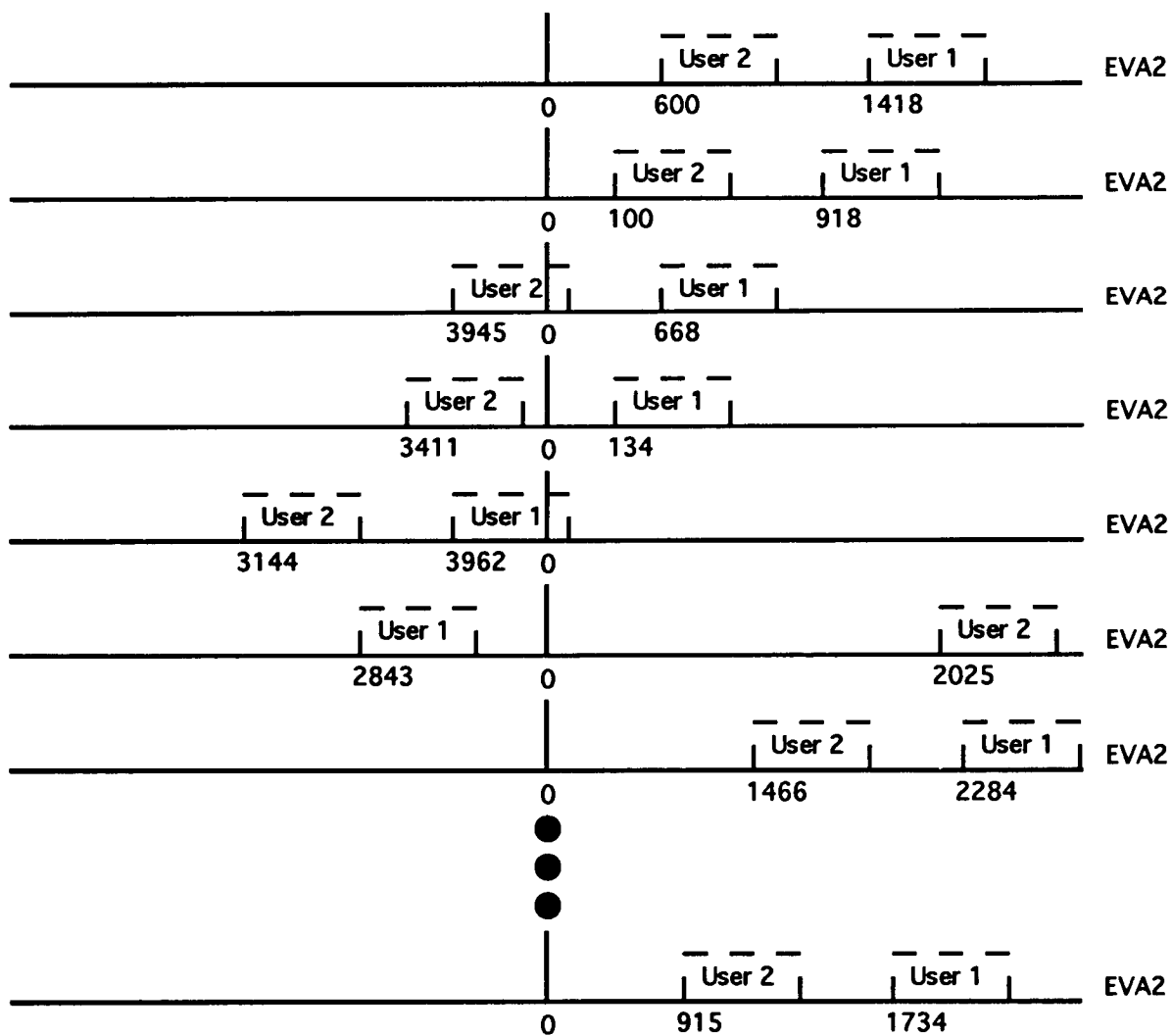
Information	Data Bits
1.) Audio	2,560
2.) Shuttle Data	96
3.) EVA Data	384
4.) Unique Word	160
5.) Guard Space	576
6.) Preamble	<u>320</u>

Total Bits per Frame 4,096

$$\text{Burst Rate} = \frac{4,096}{8\text{ms}} = 512\text{Kbps}$$

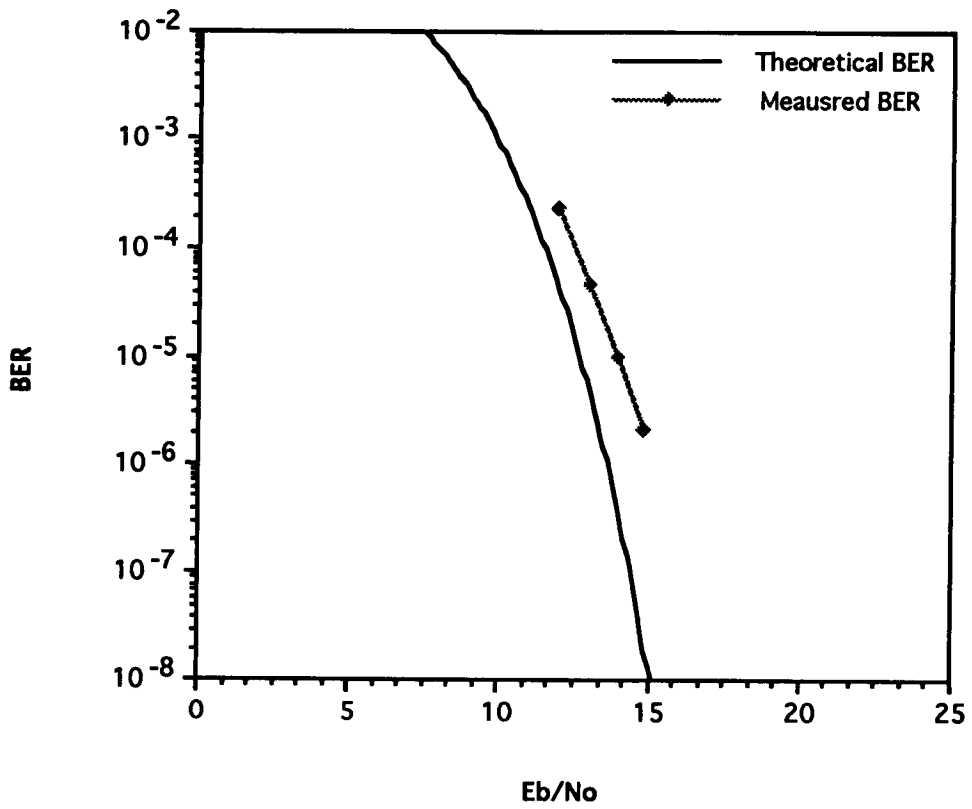
APPENDIX B

Power On Sequence of Frame Correction Events for a Third User



APPENDIX C

Theoretical Versus Measured Data for the Bit Error Rate Measurement



56-32

TRANSITION FROM NASA SPACE COMMUNICATION SYSTEMS TO COMMERCIAL COMMUNICATION PRODUCTS

Farzad Ghazvinian
LinCom Corporation
Vice President
5110 W. Goldleaf Circle
Los Angeles, CA
(213) 293-3001

William C. Lindsey
LinCom Corporation
CEO/COB
5110 W. Goldleaf Circle
Los Angeles, CA
(213) 293-3001

ABSTRACT*

Transitioning from twenty-five years of space communication system architecting, engineering and development to creating and marketing of commercial communication system hardware and software products is no simple task for small, high-tech system engineering companies whose major source of revenue has been the U. S. Government. Yet, many small businesses are faced with this onerous and perplexing task. The purpose of this talk/paper is to present one small business (LinCom) approach to taking advantage of the systems engineering expertise and knowledge captured in physical neural networks and simulation software by supporting numerous National Aeronautics and Space Administration (NASA) and the Department of Defense (DoD) projects, e.g., Space Shuttle, TDRSS, Space Station, DCSC, Milstar, etc. The innovative ingredients needed for a systems house to transition to a wireless communication system products house that supports personal communication services and networks (PCS and PCN) development in a global economy will be discussed. Efficient methods for using past government sponsored space system research and development to transition to VLSI communication chip set products will be presented along with notions of how synergy between government and industry can be maintained to benefit both parties.

1.0 INTRODUCTION

The communications and information industry currently being energized will lead and drive the global economy well into the twenty-first century. LinCom Corporation is pleased to have this opportunity to share with NASA our plans for further evolving our commercial division and our government divisions in the new environment created by the destruction of the Berlin Wall. The major purpose of our commercial division is to exploit our current system architecting, engineering and design skills so as to develop telecommunication chip sets for integration into marketable **Systems on Silicon**.

* This paper was presented at the Dual-Use Space Technology Transfer Conference, February 1-3, 1994, Houston, Texas, NASA/Johnson Space Center

LinCom Corporation is a 20 year old, stable, successful, high-tech telecommunications company which does systems research and development. LinCom is also an experienced telecommunication system architecting, systems engineering software development house offering telecommunication system services using highly skilled communication system engineers augmented by a host of system analysis and proprietary computer simulation tools. In 1989, LinCom established the goal of transitioning from a company which does science and technology development for the DoD, NASA and other government agencies into a company which does commercialization of science and technology based upon its past contracting experience with the government. In this regard, LinCom has (1) surveyed the future of telecommunications, (2) established the potential telecommunications market value, (3) evaluated LinCom's needs to make a successful transition, (4) defined a product line which uses LinCom's ability to perform the sophisticated algorithmic designs needed for the implementation of electronic chip sets and integration of these chip sets into **Systems on Silicon**.

LinCom is organized into three technical groups. These three groups have been reorganized to serve as **technology incubators** whose purpose is to interact, develop and transfer, dual-use communication chips into **Systems on Silicon**. In making the transition from a government contracting organization to one which energizes a commercial division, there are several issues and deficiencies which need to be addressed. These include: (1) Telecommunication and Material Sciences Trends, (2) Telecommunication System Trends, (3) Chip Technology Trends, (4) LinCom's R and D Philosophy, and (5) LinCom's software and hardware technology products.

1.1 Telecommunication and Material Sciences

The state-of-the-art regarding the ability to ultimately achieve the physical limits associated with telecommunication and material science development have been established. It is well known that the ultimate limits in both sciences are rapidly being approached; shifts from these limits will require new paradigms. With regard to telecommunication sciences, Shannon's (1948) limit of the ultimate achievable information transportation efficiency is approximately 2 dB away from the limit. With regard to material sciences, technologists are currently able to build microsystems using semiconductor materials and integrate these into useful chip sets for use in systems. The evolution of chip sets to Planck's limit using the nanosystem technologies predicted by Feynman (1958) are likely to be achieved during the time interval 2005-2010.

1.2 Future Global Telecommunication System Infrastructures

Future telecommunication system infrastructures are evolving rapidly. The **Wireless Quantum World** in which humans, computers and machines of the world will be interconnected and networked at the speed of light by using sophisticated Digital Communication Techniques and algorithms implemented in the form of integrated electronic chip sets is well under way. LinCom's goal over the next five years is to develop, transfer and insert certain of these chip set

technologies into the developing global telecommunication market. This market is expected to be on the order of two and one-half to three trillion US dollars by the year 2000. The market value for the telecommunication products predicted to evolve over the next fifteen to twenty years is in excess of twelve thousand trillion dollars per year. An estimate of the amount of the market value that LinCom can capture depends upon a number of factors, **none** of which are related to its technical capabilities. The major limitations on the percent of the chip and systems market that can be captured by LinCom is constrained by: (1) available working capital, (2) lack of chip materials, and (3) lack of a marketing and manufacturing capability; however, LinCom's strength lies in the strong systems design and engineering background that has been captured while working under contract to NASA, DoD and other government agencies during the past 20 years.

1.3 LinCom Research and Development

The bottom line to LinCom's approach to the commercialization of science-to-technology and technology transfer to the marketplace is the recognition that "positive sum" R and D implies joint efforts. This says that LinCom's cumulated intellectual property, systems architecting and engineering skills and simulation tools must be augmented by several additional ingredients: (1) **Capital**, (2) **Marketing**, (3) **Semiconductor Technology Access**, (4) **Fabrication and Manufacturing** capabilities. LinCom's solution to capitalization is to: (1) continue to contribute LinCom funds (R and D plus profits), (2) use Phase II SBIR funds (we have been very successful in winning Phase I and Phase II's) and (3) seek outside investors. Ingredients (2) and (3) will be met by forming partnerships and alliances with government laboratories, universities and industry.

1.4 Technology Transfer Models

There are sixteen technology transfer models which LinCom will consider for use in further evolving the commercial division. In addition, the notions of: (1) Intellectual Property Licensing, (2) Value-Added Intermediaries, (3) Use of NASA Commercialization Centers, (4) Applied Research Grant Programs, (5) State Venture Funds, and SBIRs also serve as vehicles for LinCom to exploit in transferring R+D into the commercialization of the intellectual properties accrued via LinCom's past 20 years of government contracting.

2. Commercialization for LinCom Product

Over 4 years ago LinCom identified the need to support the new and/or enhanced communication system requirements of the existing and planned space programs such as the Space Shuttle communication Upgrade, the Space to Space Communication System (SSCS) for Space Station, and the multiple access communication links between these elements. LinCom identified the potential cost-savings of employing the same set of common building blocks which form the heart and capture a major portion of the knowledge required to design a flexible, multi-function modem to support a variety of programs. The availability of key building blocks which could perform the core baseband

processing for digital modulation and demodulation would provide NASA with the superior performance and efficiency of very large scale integrated circuit technology without the associated development risk, schedule impact, and non-recurring cost.

LinCom started a IR&D project to develop an innovative multi-purpose VLSI modem chip set to encompass all the requirements of these programs resulting in significant cost savings in the development of the vital portions of the communication system. With the proliferation of VLSI development software and the availability of semi-conductor foundries, it was possible for a small business, like LinCom, to enter this highly specialized yet important commercial VLSI chip market. The result of this effort was the successful development and demonstration of a VLSI breadboard modem capable of operating up to 1.5 Mbps.

Despite the success of this IR&D project, LinCom was unable to implement the VLSI modem due to lack of funding from NASA or the commercial sector. However, the development of the VLSI breadboard had two significant benefits. First it provided LinCom with experience and know-how in the digital implementation of receivers which made it possible for LinCom to help NASA/JSC in the modeling, analysis and evaluation of several receivers. For example, modeling and analysis of the ACS and the SSCS receivers for Space Station program and the Extended Range Payload Communication Link (ERPCL) for the Space Shuttle program all resulted in identification and solution of problems associated with the design of these receivers. The second benefit of the IR&D project was to provide LinCom with the experience to be able to enter the commercial communication market.

2.1 Identification of Market for Product Development

The first step in commercialization of a product is to identify the potential market. This is probably the hardest task in the development cycle. A small company like LinCom does not have the resources to develop large system such as Personal Communication Services (PCS), Personal Communication Networks, or others which require several hundred millions of dollars to billions of dollars to develop. Secondly LinCom does not have access to marketing research that is available to large companies.

2.2 Forming Alliances

In a recent meeting of telecommunication Company executives, the speaker asked the audience which one of them have not used alliances as a means of developing products or services. All the executives indicated that they have used alliance in the development of their product or services because of large investment that is involved or the need for a particular technology or capability.

LinCom in the development any commercial product needs funding, marketing information, and marketing force to sell the product.

2.3 Commercial Communication Product Requirements

There are many differences between the requirements of the commercial and the NASA, or DoD communication system products. However the most important ones are cost of the product and development time. In developing a product for commercial market the cost, performance trade-off is the essential driver in the product design. The design which results in minimum cost, while meeting a minimum set of requirements is usually the design selected for the product. The development time which directly translates in time to market, and revenues is also an essential design and implementation consideration. For consumer electronic product with large potential market, usually the development cost is of secondary importance.

3.0 Wireless Communication Market

The invention of the telephone over one hundred years ago launched a revolution in communications that enabled people to communicate efficiently over long distance. In this decade a new revolution is taking shape-a wireless communications revolution-that will at long last free customers from the limitations of the wiring. This market includes paging services, cordless phones, wireless PBX, cellular phones, Personal Communication Services (PCS) and Personal Communication Network (PCN), wireless LANs, and many others. LinCom has joined forces with a semiconductor manufacturer, Zilog, to design, develop, and market a digital cordless phone.

In the United States, FCC regulations allow 900 MHz spread spectrum wireless devices to use as much as 1W of power, allowing compliant cordless phones to demonstrate significantly better range than the conventional 49 MHz analog cordless phones. LinCom has developed the design and the specifications for all the subsystems of the phone which include the RF frontend, Spread Spectrum Processor (SSP) transceiver/controller and the ADPCM voice encoder/decoder. LinCom has also designed and implemented the SSP as an ASIC.

The cordless phone system design uses adaptive frequency-hopping, 32 kbps ADPCM digital voice coding multiplexed with command structure, Frequency Shift Keying (FSK) modulation, and time-division duplexing (TDD).

Frequency hopping is a form of spread spectrum where the modulated signal is simply switched from carrier frequency to carrier frequency as a function of time. The signal is "Spread" in the sense that, on average, a wide bandwidth is occupied; instantaneously, however, the frequency hopped signal is still just a narrowband signal.

The phone system design involves 142, 180 kHz hop channels in the 902-928 MHz band, where the handset and base station synchronously switch across a set of 64 of those channels, dwelling 4 ms at each channel. Each channel supports a data rate of ~93 kbps. The 64 channels are chosen from the available 142 channels adaptively: an estimation algorithm operating at both the base station and handset monitors the quality of each channel over time, replacing

those channels with poor signal-to-noise ratios. Update of the hop channels is coordinated between the handset and base station through use of the commands multiplexed with the voice data.

Besides adaptive frequency hopping, the most important aspect of the phone's hop scheme is the 4 ms dwell time. This dwell time, the time that the system operates at each frequency channel, was chosen to satisfy FCC regulations, to accommodate the time required to switch from channel to channel, and to minimize the delay incurred as a result of time-division duplexing.

In time-division duplexing, the transmit and receive signals are at the same carrier frequency but offset in time. In other words, first the base station transmits at ~93 kbps (and the handset receives), then it receives (and the handset transmits). The advantage is that the design of the frequency hopping RF section is simplified, although the data rate must be at least twice the desired rate and rate-converted at the transmitter and receiver to fully realize duplex communications. Here, the transmit data rate is ~93 kbps, where that value reflects both TDD and the fact that not all of the 4 ms dwell time can be actively used for communication -- guard times are needed to allow for switching time between channels and between transmit and receive modes.

The phone uses FSK modulation to simplify the design and keep overall system costs down. For example, FSK does not require phase coherent demodulation, does not require linearity of the transmit power amplifier, and is fairly robust to interference.

Since FSK is a digital modulation technique, of course, the speech must be digitally encoded. Here, 32 kbps ADPCM has been used to reduce the data rate (bandwidth) and thus allow more hop channels within the 902-928 MHz band while still maintaining high voice quality. The voice data is multiplexed with command bits to support the relay of user commands and transceiver control signals between the handset and base station.

The SSP chip is a 16.384 MHz 84-pin PLCC part that combines a static DSP core with ~14,000 additional gates and 12k words of ROM space. The DSP core is mainly used as a RISC microcontroller. The ~14,000 gates implement the transceiver, both receive and transmit functions. On the receive side, the input IF signal from the phone's RF section is processed successively by a A/D converter, digital downconverter, limiter/discriminator (to perform FSK demod), and symbol synchronizer. Output of the demodulator is further processed to allow estimates of channel performance. On the transmit side, the source data is FSK modulated and output to a 4-bit D/A, where the output is filtered and amplified to become the 10.7 MHz signal to the RF sections. Data from the receiver and to the transmitter are buffered in hardware, allowing demultiplexing/multiplexing of the command and voice data as well as rate buffering to support TDD.

4.0 Conclusions

The design and development of this ASIC has provided a mechanism for LinCom to develop a commercial division. In the future other avenues, such as SBIR/STTR, IR&D, alliances, and NASA commercialization centers will be explored to expand LinCom's role in the commercial communication market through the development of ASIC. It is hoped that the lessons learned in the development of the commercial communication systems will result in improvements and cost savings in the development of Space Communication systems for NASA.

MIT

**Session C3: COMMUNICATIONS SIGNAL
PROCESSING AND ANALYSIS**

Session Chair: Robert Panneton

SCATTERING EFFECTS OF SOLAR PANELS ON SPACE STATION ANTENNA PERFORMANCE

57-32

Robert J. Panneton, John C. Ngo

NASA/Lyndon B. Johnson Space Center, Houston, TX 77058

Shian U. Hwu, Larry A. Johnson, James D. Elmore, Ba P. Lu, James S. Kelley
Lockheed Engineering & Sciences Company, Houston, TX 77058

ABSTRACT

Characterizing the scattering properties of the solar array panels is important in predicting Space Station antenna performance. A series of far-field, near-field, and Radar Cross Section (RCS) scattering measurements were performed at S-Band and Ku-Band microwave frequencies on Space Station solar array panels. Based on investigation of the measured scattering patterns, the solar array panels exhibit similar scattering properties to that of the same size aluminum or copper panel mockup. As a first order approximation, and for worse case interference simulation, the solar array panels may be modeled using perfect reflecting plates. Numerical results obtained using the Geometrical Theory of Diffraction (GTD) modeling technique are presented for Space Station antenna pattern degradation due to solar panel interference. The computational and experimental techniques presented in this paper are applicable for antennas mounted on other platforms such as ship, aircraft, satellite, and space or land vehicle.

I. INTRODUCTION

The solar panels are rotated dynamically to maintain a preferential orientation with respect to the sun. Due to their large electrical and physical size, the solar panels have been identified as major structure elements causing interference for all Space Station antennas. Scattering effects due to solar panels will reduce signal quality or even cause communication outages. A thorough understanding of the solar panel scattering effects is necessary for the design of a reliable communications system, the siting of antennas, communication systems analysis, and communication coverage prediction. It is also very important for successful operation of the Space Station.

The solar array panel is a composite structure formed by closely spaced solar cells. The solar cells are silicon and are welded on the front surface of the solar array panel, as shown in Fig. 1. A grid of copper strips which collect current is on the back side of the solar array panel, as shown in Fig. 2. Traditionally, the solar panel is modeled as a perfect electrical conductor at Very High Frequency (VHF). Very little information is available on the scattering effects of a solar panel at microwave frequencies,

which is the primary frequency region for Space Station communications. A series of far-field, near-field, and Radar Cross Section (RCS) scattering measurements were performed at S-Band and Ku-Band microwave frequencies on Space Station solar array panels using the NASA/Johnson Space Center Anechoic Chamber test facility.

In the microwave frequency region, reflection and diffraction become the dominant mechanisms of interaction between electromagnetic waves and scattering structures. The Geometrical Theory of Diffraction (GTD) has been used successfully to predict microwave propagation characteristics. A study was conducted to validate the GTD modeling technique for the scattering pattern prediction of a conducting (aluminum or copper) panel mockup which simulated the solar array panels. Results obtained with the GTD model have been compared with measurements for a variety of scattering cases and good agreement has been experienced.

As a summary, sample numerical results are presented. Antenna pattern degradations due to solar panel interference are shown for Space Station Ultrahigh Frequency (UHF), S-Band and Ku-Band antennas.

II. GEOMETRICAL THEORY OF DIFFRACTION (GTD)

In this section, the GTD modeling technique used in this study for obtaining the theoretical scattering patterns is briefly described.

At high frequencies the scattering fields depend on the electrical and geometrical properties of the scatterer in the immediate neighborhood of the point of reflection and diffraction. Thus, the scattering structure can be modeled as a collection of basic geometrical components such as flat plates and cylinders. The size of each component must be electrically large, that is, $kl \gg 1$, where $k = 2\pi/\lambda$ is the wave number and l the minimum dimension of the basic component. The total scattering fields are obtained by summing up the individual contributions of each of the basic geometrical components and the contributions of the interactions among the basic geometrical components.

The diffracted fields are related to the incident fields by means of diffraction coefficients, in a similar way the reflected fields can be obtained using reflection coefficients. Since the diffracted field is determined solely by the incident field and the local nature of the scattering surface, it is possible to derive a diffraction function relating the incident field to the diffracted field for a certain scatter geometry, a so called canonical configuration.

The application of GTD to a given radiation problem is first to decompose the scattering structure into simple geometrical shapes where the diffraction coefficients are known. Next, all field components contributing to the radiation intensity in the field point must be traced, and the individual contributions must be determined. As illustrated in Fig. 3, the resultant field is given by summing all the complex contributing components

$$\vec{E}^{total} = \vec{E}^{direct} + \sum_{n=1}^N \vec{E}_n^{reflected} + \sum_{m=1}^M \vec{E}_m^{diffracted}.$$

III. FAR-FIELD SCATTERING

The far-field scattering effects of a 24.1 cm by 38.4 cm solar cell coupon to be used on the Space Station are experimentally tested for two different separation distances at S-Band frequencies. The test results for the solar cell coupon are compared to measured patterns for a reference metal plate and to GTD computed patterns. The solar cell coupon was obtained from the NASA-Lewis Research Center and comprises a total of 12 solar cells (3 cells by 4 cells). A cell dimension is 8 cm X 8 cm. The solar panel is formed in a two-dimensional rectangular lattice of solar cells.

The setup for the 12-cell solar coupon scattering pattern measurement is illustrated in Fig. 4. The measurements were conducted at 2.0, 2.175, and 2.3 GHz with two different separation distances, $d = 1$ foot (0.30 m) or 2 feet (0.61 m), between the solar coupon and the antenna. A complete measurement data package with all tested and simulated results can be found in report [1].

For $d = 2$ feet (0.61 m) separation distance between the testing plate and the antenna, the measured scattering patterns for the solar cell coupon and for the reference metal plate with the same dimensions at 2.175 GHz are shown in Fig. 5. The results for a 1 foot (0.30 m) separation distance between the testing plate and the antenna, are shown in Fig. 6.

It is observed that the scattering properties of the Space Station solar cell structure are such that the surfaces are highly reflective at S-Band frequencies. Thus, based on these far-field scattering results [1], the Space Station solar panel can be modeled as perfect electrical conductors.

VI. NEAR-FIELD SCATTERING

The near-field scattering pattern measurements of various Space Station solar array partial full-scale mockup configurations were performed. To simulate the Space Station environment, the operating frequencies of the Space Station S-Band and Ku-Band antenna systems, which are 2.175 GHz and 15 GHz, respectively, were used in the experimental measurements and computer simulations. For verification purposes, only selected solar array mockup configurations were simulated and compared to measured data. A complete data package with all mockup configurations can be found in reports [2, 3].

The setup of the near-field scattering measurements for solar array mockups in the Anechoic Chamber is illustrated in Fig. 7. In tests, two orthogonal linear polarized measurements were conducted for each mockup configuration. The parallel E-field scattering patterns were obtained by orienting the source antenna illuminating the solar array mockups so that the incident E-field is parallel to the reflecting surface of the tested mockups. The parallel H-field scattering patterns were obtained by orienting the source antenna illuminating the solar array mockups so that the incident H-field is parallel to the reflecting surface of the tested mockups. Note that the amplitude has been normalized in both measured and computed patterns. All scattering patterns shown in this paper are at a relative power level and can only be compared at the same frequency.

A 2-panel section of Space Station solar array material, with a size of 4.32 m by 0.76 m (170" by 30"), was obtained from the NASA-Lewis Research Center. Similar scattering patterns were observed when comparing the measured results for the front and back side scattering tests [2]. The measured parallel H-field scattering patterns for the Space Station solar array panels and for the same size aluminum plate at the S-Band frequency are shown in Figs. 8 and 9, respectively.

The near-field scattering patterns using conducting plate models (aluminum) are very similar to the measured scattering patterns for the solar array panels [2]. Based on these near-field scattering results, the solar array panels may be modeled using perfect reflecting plate models at S-Band and Ku-Band frequencies.

The measured and computed scattering patterns using conducting plate models simulating the Space Station solar array panels are similar [2]. The solar array panels were modeled using a flat conducting plate in the GTD computations and using an aluminum strip in the measurements.

V. RCS SCATTERING

The RCS is a scattering measurement that is a func-

tion of the scatterer's shape, the frequency or wavelength of operation, the polarization of the transmitter and receiver, the angle of incidence of the incident wave with respect to the scatterer, and the materials composing the scatterer. The RCS patterns can be used to describe the scattering properties of a scatterer.

The RCS pattern measurements were performed on two Space Station solar array panels with a size of 4.32 m by 0.76 m (170" by 30"). The monostatic and bistatic RCS was measured as a function of azimuth angle in a plane perpendicular to the test plate. The bistatic angle, β , is approximately zero for the monostatic tests and is 37.5° for the bistatic tests. Two orthogonal linear polarized measurements were conducted at an S-Band frequency of 2.175 GHz and a Ku-Band frequency of 15 GHz. A complete data package can be found in reports [4, 5].

The setup of the RCS measurements for the solar array mockups in the Anechoic Chamber is illustrated in Fig. 10. The horizontal polarized RCS patterns are obtained by orienting the source and receiving antennas so that the electric field is parallel to the ground. The vertical polarized RCS patterns are obtained by orienting the source and receiving antennas so that the electric field is vertical to the ground.

The solar array panels were tested in the near-field distance illumination condition. The far-field distance is defined as

$$R = \frac{2D^2}{\lambda},$$

where the largest dimension is $D = 170$ inches for the test solar array panels and λ is the wavelength. The corresponding far-field distance is $R = 270$ meters at 2.175 GHz and $R = 1864$ meters at 15 GHz. Such a large distance requirement cannot be realized in the indoor test chamber. Thus, only near-field RCS measurements were possible.

Similar RCS patterns were observed by comparing the measured results for the front and back side scattering measurements [4]. A comparison of measured horizontal monostatic RCS patterns for the front (1) and back (2) sides of the solar array panels at the Ku-Band frequency is shown in Fig. 11. Note that the curve for the front side RCS pattern was arbitrarily raised by 30 dB to avoid overlapping with the curve for the back side RCS pattern.

The solar array panels, with a size of 4.32 m by 0.76 m (170" by 30"), form a large reflector. For such a large aperture, the corresponding far-field RCS pattern should exhibit a sharp pencil-beam with very narrow beamwidth. However, in the near-field distance, the patterns show a rather broad boresight. The ripples on the broad boresight are a result of the interference between the surface reflected fields and the edge diffracted fields.

The measured RCS results indicate that the patterns for the copper panel mockup are similar to the patterns for the solar array panels [4]. A comparison of measured horizontal monostatic RCS patterns for the solar array panels (1) and the copper panel (2) at the Ku-Band frequency is shown in Fig. 12. Note that the curve for the solar panels RCS pattern was arbitrarily raised by 30 dB to avoid overlapping.

The solar array panels exhibited slightly less reflectivity than the copper panel. Based on this RCS scattering experiment, the solar array panels may be modeled using perfect reflecting plates.

VI. ANTENNA PATTERN INTERFERENCE

The results obtained from the above scattering measurements indicate that the solar array is highly reflective at S-Band and Ku-Band. A conducting plate is assumed to be a reasonably good electromagnetic model for the solar panel. In the following analyses, as a first order approximation, and for worse case interference simulation, the Space Station solar panels are modeled as perfect electrical conductors.

a. UHF Antenna

The Space Station UHF antennas are designed to provide a UHF communications link between the Space Station, Extravehicular Activity (EVA) astronauts, and the Space Shuttle Orbiter (SSO). The ability of the UHF antennas to communicate with an EVA immediately behind the solar array panels is a concern.

These regions are expected to have significant shadowing effects due to the large electrical size of the solar panels. The EVA directly behind is blocked from both antennas. The only signal present is due to the diffracted fields, which are usually very small compared to the direct and reflected fields.

The Space Station UHF antenna is a 1/2-turn quadrifilar helical antenna with a beamwidth of $\pm 90^\circ$ from boresight at -4 dBic. A frequency of 415 MHz was used in this analysis. In Fig. 13, the computed signal strength in terms of the electric field behind solar panels is presented and compared to the signal strength corresponding to the 0 dB link margin. The results indicate that this is a critical region where the UHF communications link to the EVA may not be maintained[6].

b. S-Band Antenna

The S-Band subsystem, also known as the Assembly/Contingency Subsystem (ACS), is a bi-directional Radio Frequency (RF) link utilizing the Tracking and Data Relay Satellite System (TDRSS) S-Band Single Access (SSA) service to receive audio, software uploads, and commands from the ground station, and to transmit

audio and core element telemetry to the ground station. In addition, the S-Band subsystem is utilized to support ground based tracking services for station orbit determination.

The scattering effects of the solar array panels on the ACS high gain antenna patterns are analyzed. The Space Station ACS high gain antenna is a steerable conical horn with a 16 dBic peak gain and a 30° 3-dB beamwidth. A frequency of 2.175 GHz was used. The predicted antenna patterns were computed using the GTD technique. The results indicate that, if significant interference (more than 1 dB antenna gain loss in the $\pm 4^\circ$ antenna bore-sight region) with ACS antenna patterns is to be avoided, the ACS steerable antenna 3 dB beamwidth must be kept clear of solar panel blockage. This means that the ACS high gain antennas should not be pointed within 15 degrees of obstructions. Fig. 14 shows a summary of the maximum antenna gain degradation in the $\pm 4^\circ$ antenna boresight region due to solar panel interference[7].

c. Ku-Band Antenna

The Ku-Band Space-to-Ground Subsystem (SGS) is a single direction RF link utilizing the TDRSS Ku-Band Single Access (KSA) service to transmit payload data and video to the ground station.

The solar panels are rotated to maintain a preferential orientation with respect to the sun. The SGS steerable reflector antenna must track signals in the upper hemisphere and will encounter scattering interference from the solar panels.

A 4-foot reflector antenna operated at 13.5 GHz was used to simulate the SGS antenna. Figs. 15 and 16 show the computed and the measured difference mode patterns for the 4-ft reflector with the solar panel penetrating 1 foot and 2 feet into the projected antenna aperture cylinder. The same characteristic behavior is shown in both the measured and computed patterns.

A gain loss in the main lobes and a shift in null position, as well as a decrease in the depth of the null, were observed for the antenna patterns with the solar panel penetrating into the projected antenna aperture cylinder.

In order to avoid significant scattering interference, the SGS antenna projected aperture cylinder must be kept clear from solar panel blockage. This means that the SGS antenna projected aperture cylinder should not be pointed onto solar panels, otherwise significant antenna pattern degradation will occur.

It is also noted that when the antenna was blocked by the solar panel, the strong reflection from the solar panel increases the back lobes to a very high level. These back lobes raise Electromagnetic Compatibil-

ity/Electromagnetic Interference (EMC/EMI) concerns to nearby equipment and EVA astronauts[8].

VII. CONCLUSION

The scattering properties of the solar array panels were characterized for the purpose of predicting Space Station antenna performance. A series of far-field, near-field, and RCS scattering measurements were performed at S-Band and Ku-Band microwave frequencies on Space Station solar array panels. As a first order approximation, and for worse case interference simulation, the solar array panels may be modeled using perfect reflecting plates. Numerical results obtained using the GTD modeling technique were presented. Antenna pattern degradations due to solar panel interference were shown for Space Station Ultrahigh Frequency (UHF), S-Band and Ku-Band antennas.

The computational and experimental techniques presented in this paper are applicable for the following applications:

- Antenna near-field and far-field radiation pattern prediction.
- Radiation hazard prediction.
- Antenna coupling and electromagnetic compatibility analysis.
- Radar cross section computation.
- Signal strength prediction for wireless, personal and indoor communications.

The antennas can be mounted on various platforms such as ship, aircraft, satellite, and space or land vehicle.

References

- [1] S. U. Hwu, L. A. Johnson, R. J. Panneton, "Scattering Effects of Space Station Solar Cell Coupon at S-Band Frequencies," NASA/JSC Technical Report JSC-32206 (LESC-30136), Lyndon B. Johnson Space Center, NASA, Houston, TX, February 1992.
- [2] J. D. Elmore and L. A. Johnson, "Scattering Measurements for Space Station Solar Array Partial Full-Scale Mockup," NASA/JSC Technical Report JSC-32304 (LESC-30832), Lyndon B. Johnson Space Center, NASA, Houston, TX, June 1993.
- [3] S. U. Hwu, B. P. Lu, R. J. Panneton, "Near-Field Scattering Measurement and Verification for Space

Station Solar Array Mockup," NASA/JSC Technical Report JSC-32314 (LESC-30981), Lyndon B. Johnson Space Center, NASA, Houston, TX, August 1993.

- [4] L. A. Johnson, "Backscatter Measurements for the Space Station Solar Array Panels," NASA/JSC Technical Report JSC-32322 (LESC-31047), Lyndon B. Johnson Space Center, NASA, Houston, TX, January 1994.
- [5] S. U. Hwu, B. P. Lu, R. J. Panneton, "Radar Cross Section Measurement and Verification for Space Station Solar Array Panels," NASA/JSC Technical Report JSC-32316 (LESC-30983), Lyndon B. Johnson Space Center, NASA, Houston, TX, November 1993.

- [6] S. U. Hwu, B. P. Lu, R. J. Panneton, G. D. Arndt, "Space Station Freedom UHF Antenna Performance Prediction," AIAA Paper No. 93-1136, February 1993.
- [7] S. U. Hwu, B. P. Lu, L. A. Johnson, J. S. Fournet, R. J. Panneton, D. S. Eggers, G. D. Arndt, "Scattering Effects of Space Station Structure on Assembly/Contingency Subsystem (ACS) Antenna Performance," AIAA Paper No. 92-1538, March 1992.
- [8] S. U. Hwu, L. A. Johnson, J. S. Fournet, R. J. Panneton, D. S. Eggers, G. D. Arndt, "Scattering Effects of Solar Panel on Space Station Space-to-Ground (SGS) Antenna Performance," AIAA Paper No. 92-1941, March 1992.

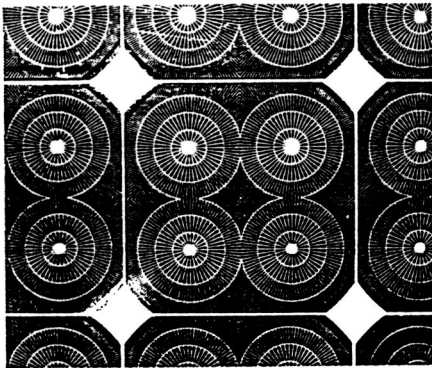


Figure 1: Front surface of the solar array panel.

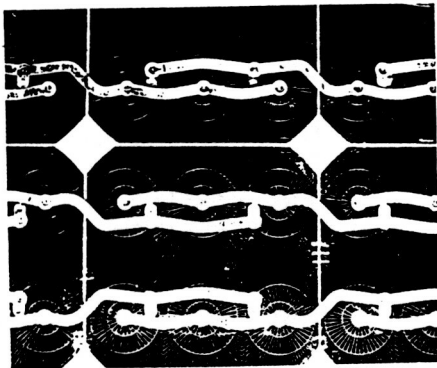


Figure 2: Back surface of the solar array panel.

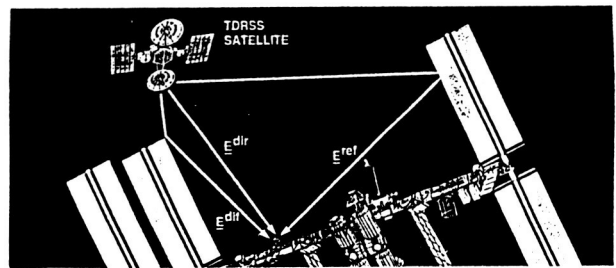


Figure 3: Total field summing all contributions.

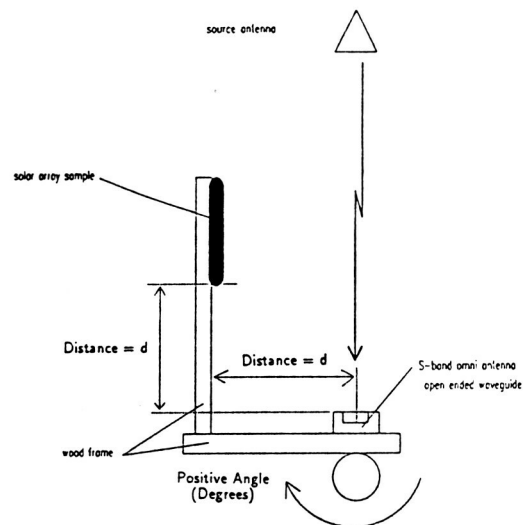


Figure 4: Setup for far-field scattering measurement.

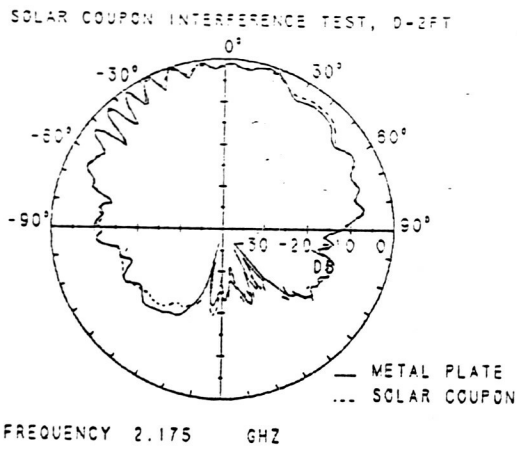


Figure 5: Far-field patterns for 0.61 m separation.

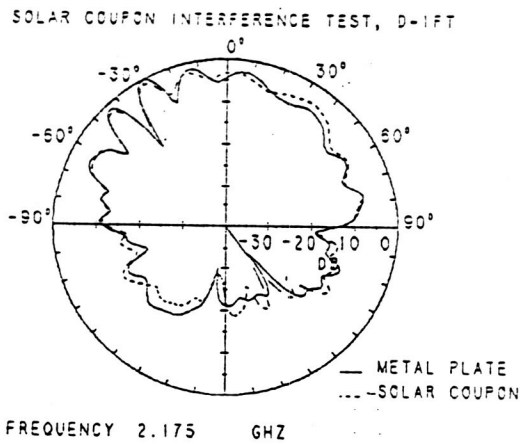


Figure 6: Far-field patterns for 0.30 m separation.

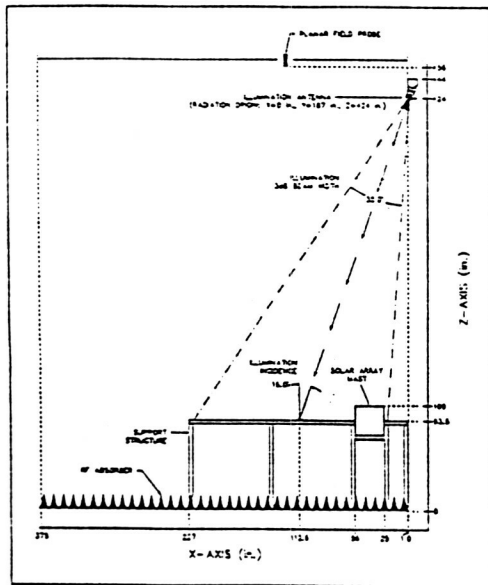


Figure 7: Setup for near-field scattering measurement.

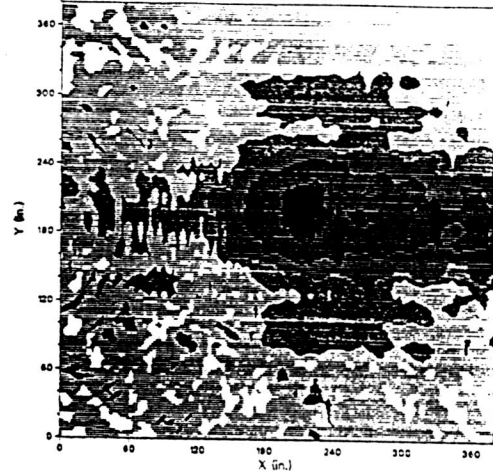


Figure 8: Near-field patterns for solar panels.

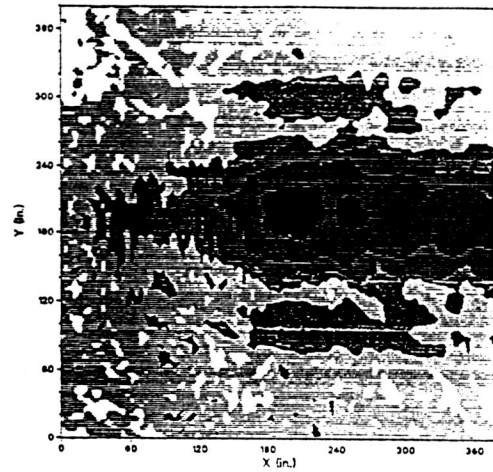


Figure 9: Near-field patterns for aluminum plates.

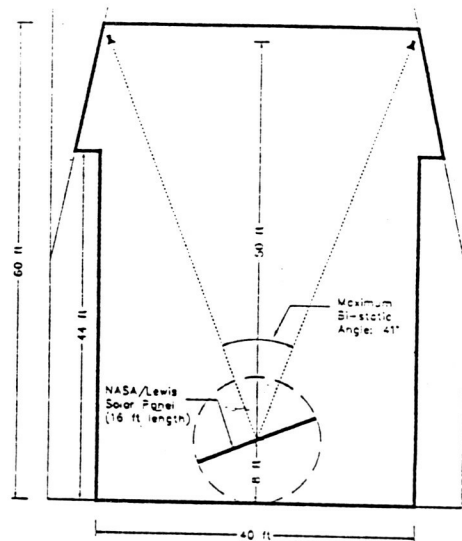


Figure 10: Setup for RCS scattering measurement.

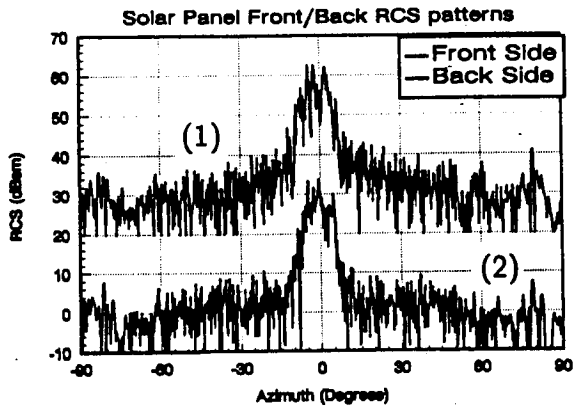


Figure 11: RCS patterns for solar panels.

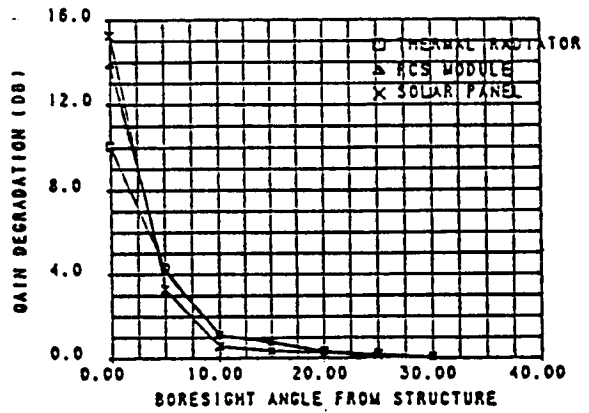


Figure 14: S-Band antenna gain degradation.

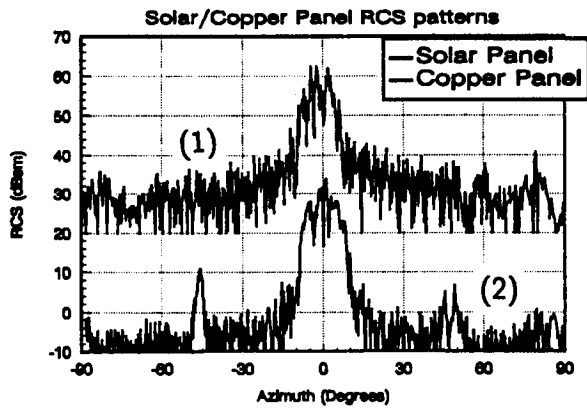


Figure 12: RCS patterns for copper plate.

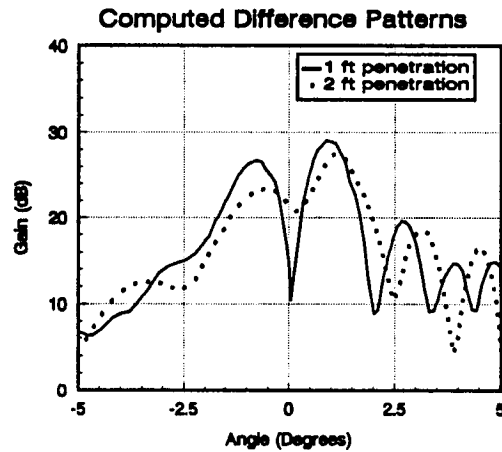


Figure 15: Computed Ku-Band antenna patterns.

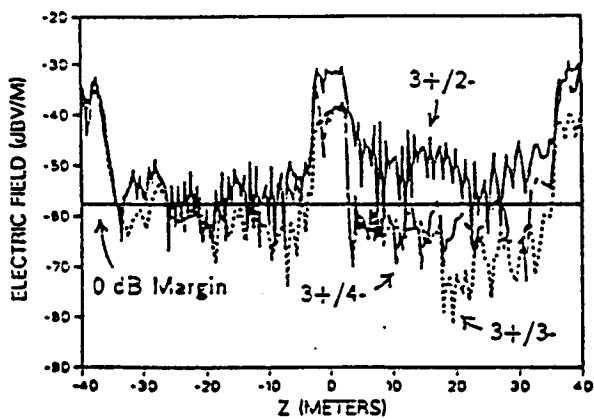


Figure 13: UHF signal strength behind solar panels.

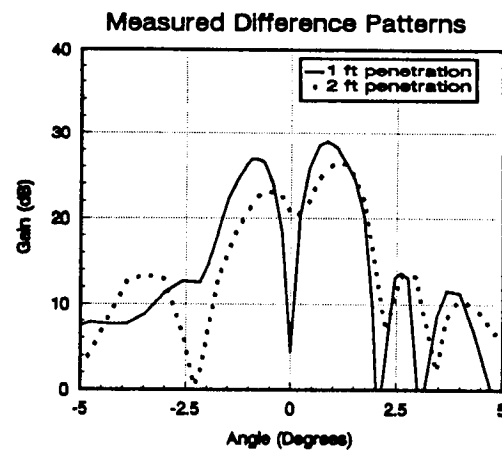


Figure 16: Measured Ku-Band antenna patterns.

10-32 ✓

DYNAMIC ENVIRONMENT COMMUNICATIONS ANALYSIS TESTBED (DECAT) AND ITS APPLICATIONS TO DUAL USE SPACE TECHNOLOGY

Antha A. Adkins, William C. Gadd, Q. D. Kroll, Robert H. Sparr, Douglas J. Steel
Lockheed Engineering and Sciences Company
2400 NASA Road 1
Houston, Texas 77058

ABSTRACT

The Communication System Simulation Laboratory (CSSL) at the National Aeronautics and Space Administration (NASA)/Johnson Space Center (JSC) provides a facility for analyzing and simulating space communication systems. The CSSL has simulation tools that model NASA communication systems and the environment in which they operate. One of these tools is a radio frequency coverage analysis tool called the Dynamic Environment Communications Analysis Testbed (DECAT). This paper focuses on DECAT and its applications to the Space Station and Space Shuttle programs.

1. INTRODUCTION

The CSSL is located at NASA/JSC and is managed by the System Engineering Branch of the Tracking and Communication Division of the Engineering Directorate. DECAT was developed by the Lockheed Engineering and Sciences Company to perform simulations in the CSSL. This paper describes the DECAT simulation tool and its applications to both the Space Station and the Space Shuttle Orbiter (SSO) with emphasis on how DECAT simulations assist in the analysis of antenna placement, Tracking and Data Relay Satellite (TDRS) coverage studies, Global Positioning System (GPS) coverage, and interference and frequency allocation analyses.

2. DECAT SIMULATION TOOL

The DECAT architecture, graphical user interface (GUI), and post processor are discussed in this section. DECAT is described in more detail in reference 1.

2.1 DECAT Architecture

The motion of DECAT vehicles and planets is defined hierarchically, so the user is not required to handle coordinate transformations. An example is the Earth orbiting the Sun, the space vehicle in a Local Vertical-Local Horizontal (LVLH) orbit around the Earth, and the space vehicle in a fixed attitude such as gravity gradient with respect to the LVLH orbit. DECAT is flexible in that any number of vehicles in any configuration may be defined. The Sun and planets may be included as necessary.

The DECAT users can create multi-vehicle dynamic radio frequency (RF) coverage simulations without programming. Figure 1 shows an overview of the DECAT system. Functions and models are chosen and initialized using a GUI during simulation setup. The GUI writes text files describing the simulation setup which are read by DECAT. DECAT writes result files for post processing or sends information across the network to be displayed by a three dimensional (3D) Graphics workstation. DECAT is written in American National Standards Institute (ANSI) C programming language and can be run on either UNIX or VAX/VMS workstations.

New functionality can be added to DECAT as required to model the vehicle's environment or to calculate new parameters. DECAT has two basic kinds of defined functions; functions which describe vehicles and their environment and control functions which perform calculations during the simulation.

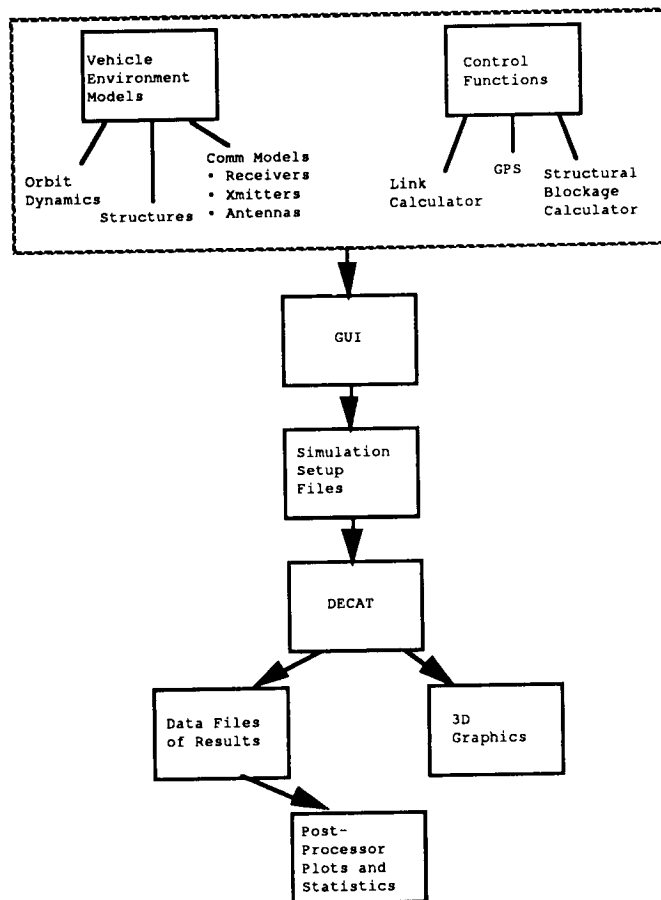


Figure 1 - DECAT System Overview

2.1.1 Vehicle Models

Vehicle models include orbital dynamics, structures, and communication system equipment such as antennas, receivers, and transmitters. Simple orbital dynamics models (LVLH circular or elliptical orbits, fixed attitude, etc.) are available to perform RF coverage analysis with DECAT. These models enable the user to quickly investigate the RF coverage for a new orbital scenario with a minimum of input parameters. Trajectory tapes from other organizations which specialize in trajectories are obtained and input into DECAT when coverage analysis is required for trajectories such as vehicle ascent or entry, or translunar. Trajectory tapes are also acquired for Earth orbits when very accurate coverage results are required or if maneuvers in Earth orbit are to be included in the simulation.

Structures define the vehicle physically. The structures are used by DECAT to determine if the vehicle's structure obscures the communication link. Each segment may have some particular

motion assigned to it. For example, solar panels track the Sun and radiators anti-track the Sun. A structure data base from which the desired structure can be selected is maintained.

The communication systems for each vehicle are defined by receivers, transmitters, and antennas. Any number of transmitters or receivers may be connected to the same antenna. Each transmitter and receiver may be connected to only one antenna. Communication links are defined by connecting a transmitter/antenna pair on one vehicle to a receiver/antenna pair on another vehicle.

DECAT antennas are defined by their position on the vehicle, initial boresight pointing directions, antenna gain function, and boresight motion function. Antenna gain patterns can be deterministic functions or data from files. Boresight models include fixed boresight and ideal tracking. New antenna gain or boresight functions can be added to model specific antennas.

DECAT receivers use a simplified "link margin" approach to model the actual receiver. Parameters which define the receiver include losses, noise floor, bandwidth, and frequency. DECAT transmitters follow a similar approach as the DECAT receiver model. Parameters which define a transmitter include transmitter power, losses, frequency and bandwidth.

2.1.2 Control Functions

Control functions provide a way to add functionality to DECAT. A number of control functions for DECAT have been developed, for example the blockage calculator control function. It uses information in DECAT such as planet position, vehicle position, vehicle structure, and vehicle antenna information to calculate the structure and planet blockage from the vehicle antenna to another antenna. This blockage is calculated using line-of-sight vectors which can be a line, a cone, or a cylinder. The cone has been used to define a protective zone for structural multipath analyses.

2.2 DECAT GUI

The DECAT GUI can only execute on the VAX cluster. The GUI was developed using the Transportable Applications Environment¹ which builds X Window System displays. X Window System displays were developed for vehicle information files such as receivers, transmitters, and orbital dynamics, and for control functions as they become operational. The GUI writes text files which describe the simulation setup as input to the DECAT executive. The X Window System display of the GUI may be distributed over the ethernet to any workstation that is X Window System-compatible. With this capability DECAT simulations may be set up with the GUI at the engineer's desk on an Apple Macintosh computer. Figure 2 shows the hierarchical representation of the vehicles and planets via the DECAT GUI.

¹ Transportable Applications Environment was developed by NASA/Goddard Space Flight Center and is distributed by the Computer Software Management and Information Center (COSMIC).

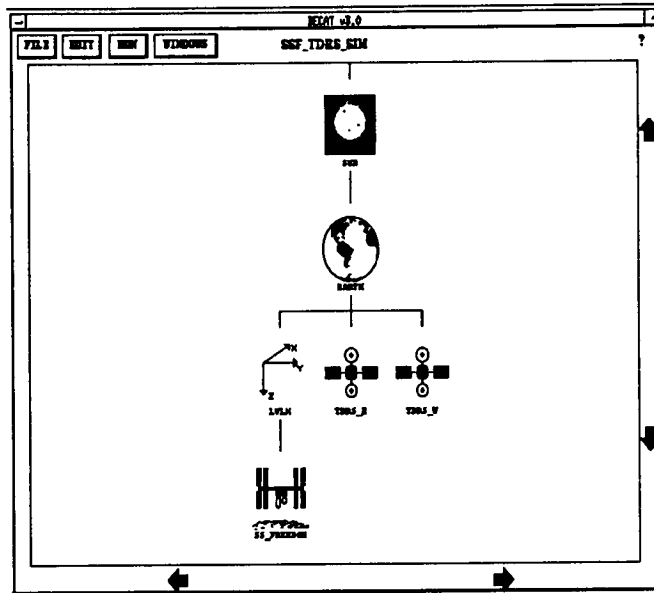


Figure 2 - DECAT GUI

2.3 DECAT Post-Processor

Probes of variables can be defined during DECAT simulation setup so that during the DECAT simulation the values of the variables can be saved at each time step in a text file. The Overall Link Availability Post Processor (OLAPP) processes DECAT output text files to produce statistics on RF coverage, time history plots of RF coverage, and x-y signal plots. It can include structural blockage, planet blockage, RF interference (RFI), acquisition outage, and power flux density violation time. RF coverage statistics include mean coverage time, mode, and the number of communication periods longer than x minutes. For time history plots, outage data from the OLAPP are filtered and scatter plots are produced in the user's favorite plotting package. While the statistics provide the percentage of time that communication coverage is available, the time history plots show the distribution of communications availability which cannot be determined through statistics alone.

The 3D graphics display control function sends vehicle position, boresight vectors or cones, and structure information to a Silicon Graphics Iris workstation with the Tree Display Manager² software for 3D graphical display of the simulation. This display allows engineers to see the antenna's line-of-sight as it is obscured by the vehicle structure.

3. DECAT APPLICATIONS FOR NASA SPACE PROGRAMS

3.1 Space Station Communications Coverage Analysis

The Space Station has two communications subsystems to communicate with the ground, the Ku-band Space-to-Ground Subsystem (SGS) and the S-band Assembly Contingency Subsystem (ACS). Both of these systems are designed to operate with the NASA Tracking and Data Relay Satellite System (TDRSS). The TDRSS consists of several geosynchronous orbiting satellites which relay the signals from Space Station to the White Sands Ground Terminal in New Mexico for processing and transmission to the Space Station Control Center Complex (CCC) located at JSC.

² Tree Display Manager was developed by NASA/JSC and is distributed by COSMIC.

During the 1991 restructuring of Space Station Freedom (SSF), the command data requirement was removed from the SGS. The ACS was selected as the primary operational communications system and the restructured SGS link remained for sending video and attached payload data to the ground. To meet the demand of becoming the primary communications link, the data rates of the ACS were increased and a steerable horn antenna was added to support the required increase in the link margin.

During the Delta Preliminary Design Review (DPDR) of the SSF communications systems, the ACS antenna locations were found to be unacceptable for a number of reasons.³ The primary reason was that the ACS antenna was located in the vicinity of major obstructions such as the Reaction Control System (RCS) modules and the SGS antenna.

The Electromagnetic Systems Branch at JSC analytically determined the degradation that would occur to the ACS antenna pattern whenever the physical boresight of the antenna pointed in the vicinity of SSF structural members. For the ACS horn antenna design, there could be as much as 1-dB of degradation to the antenna gain specification value if the boresight of the antenna was pointed within 15° of the structure. To ensure that reliable communications exist for the ACS, the absolute line of sight to the TDRS was protected by a 15° half-cone angle mask. This criteria became the accepted standard of definition for degraded system performance of the SSF ACS communications system.

A working group consisting of NASA Engineering Directorate, NASA Mission Operations Directorate, and McDonnell Douglas Space Systems Company (MDSSC) personnel was formed to determine the location for the antennas to provide reasonable coverage. However, not only was coverage to be considered, but also the impacts of weight, power, thermal constraints, Extra Vehicular Activity (EVA), operations, and of course, cost. The issue was further complicated since the ACS was originally required to perform in only two different flight orientations during the assembly phase of SSF. Selection of the ACS as the prime communication system for the life of SSF required the consideration of a third, long-term flight orientation.

Communications with SSF are required by Mission-Build 2 (MB-2), or the second assembly flight. This greatly reduced the number of possible locations for the ACS antenna. The use of the first SSF stage was eliminated because a solar panel was placed on this segment leaving only the MB-2 stage. DECAT was set up with two orbiting TDRS models representing TDRS-West and TDRS-East and an orbiting SSF model. Monte Carlo simulations were run to analyze 14 potential antenna locations in the three attitudes for various SSF structure configurations. The selection of cases represented several stages of SSF construction and growth.

DECAT utilized a ray tracing routine to determine obscuration of the optical line of sight to the desired receiver by the SSF structure. This computation used accurate polygon structure models built from official NASA Computer-Aided Design (CAD) data bases. A new line of sight or "cone" blockage control function was required to compute the degraded performance as defined by the Electromagnetic Systems Branch. Computationally efficient algorithms were developed and implemented, which resulted in a relatively quick blockage computation using the same structure data bases. DECAT can compute obscuration in both time step and angular spatial intervals. The combinations of these data outputs can be analyzed to give insight into the cause and effect of obscuration. Using the time history plots, attitude profiles, and obscuration profiles, the structural component that causes the obscuration can be identified easily. Statistics can be generated from

³ RID 2-57, DRB Action Item 264 DRB 920123 (3-2), Jan. 23, 1992.

the time step data, such as the percentage of time the antenna pattern is not degraded by SSF structure when the TDRSS is in view, and the maximum and minimum time duration that these events occur. These are valuable units of measure for evaluating the performance of the antenna at various locations.

The selection of ACS antenna locations was narrowed down to the two locations which provided optimum coverage for the limited amount of mounting space available. One of the locations was found to interfere with the berthing adapter required on the MB-2 flight. Since the DECAT simulations had shown that any other ACS antenna location would be less than optimum for the life of SSF, it was accepted that the antenna would remain in the proposed location and a special antenna mounting adapter would be designed to deploy the ACS antenna out of the way of the berthing adapter for the single flight. The ACS antenna would be moved back into its permanent position after MB-2. Simulations of this interim configuration were also performed using DECAT.

This is one example of the use of the DECAT cone blockage control function and post processing tools. Not only can Monte Carlo analyses be performed, but DECAT can also process specific flight trajectories to verify unique Space Station operational scenarios. DECAT has been used extensively to verify communications coverage and determine deficiencies in SSF operations (ref. 2). Several alternatives, recommendations, and solutions were established and presented to various different SSF operations panels and working groups to solve communications coverage problems.

3.2 Space Shuttle and ORFEUS/SPAS Payload GPS Analysis

DECAT can analyze a space vehicle's usage of the Global Positioning System (GPS) from a communication link point of view by using control functions to calculate which GPS Space Vehicles (SVs) are available based on Earth blockage, structural blockage, vehicle antenna pattern, and signal margin. DECAT also contains control functions to calculate the Geometric Dilution of Precision (GDOP) and Position Dilution of Precision (PDOP) given combinations of available SVs. Additional control functions can be written which model particular receivers. For Space Transportation System (STS)-51, DECAT was used to investigate the usage of GPS for relative positioning of both the SSO and the Orbiting and Retrievable Far and Extreme Ultraviolet Spectrometer (ORFEUS)-Shuttle Pallet Satellite (SPAS) payload. How often the SSO can receive four GPS SVs and how often the SSO and ORFEUS-SPAS select the same four SVs for solution were investigated.

During the STS-51 mission, the SSO deployed and recovered the ORFEUS-SPAS. Both the SSO and the ORFEUS-SPAS carried GPS receivers which could be used to perform differential GPS calculations to measure the distance between the SSO and the ORFEUS-SPAS. The ORFEUS-SPAS transmitted its GPS-derived state vector to the SSO, which differenced this state vector with its own to arrive at their separation in real time.

The pre-mission simulations modeled the SSO, the ORFEUS-SPAS, and the GPS constellation in orbit around the Earth. Several DECAT simulations for different combinations of SSO trajectories and antenna placements were performed. The GPS constellation orbital parameters were downloaded from the GPS Bulletin Board Service (BBS) at Holloman Air Force Base, and each GPS vehicle was modeled with an elliptical orbit and was given a transmitter and an antenna. The ORFEUS-SPAS vehicle was placed in another elliptical orbit and was given a receiver and an omni directional antenna. The SSO vehicle was placed in a trajectory relative to the ORFEUS-SPAS vehicle and was given a receiver and three GPS antennas. The GPS antennas were placed at

the geometric center of three of the SSO windows and oriented with their boresight axes normal to the window plane. To model the effects of the SSO structure, the antennas were given a visibility cone of 65°. The results generated by these simulations gave mission planners and controllers information on how the system should perform during the actual flight.

After the mission, DECAT simulations using the Best Estimated Trajectory and actual SSO antenna positions will be performed and compared to the actual data. The data from the SSO will be available first, so the first post-mission simulations will compare DECAT simulation data to SSO GPS availability data. When data from the ORFEUS-SPAS becomes available, simulations will be performed to compare DECAT simulation data to the ORFEUS-SPAS GPS availability data and DECAT simulation data to the joint SSO/ORFEUS-SPAS GPS data as described in the pre-mission analysis.

3.3 Frequency Interference Measurement Modeling

The Space-to-Space Communications Subsystem (SSCS) is an Ultra-High Frequency (UHF) system designed to provide voice and data communications between the SSO, the Space Station, and Extravehicular Activity (EVA) astronauts. The SSCS will replace the existing SSO to EVA communications system which operates in the 250-300 MHz frequency band, and will move to a proposed 410-420 MHz frequency band (ref. 3). Because the SSCS must operate with the presence of ground stations in the band, the effects of ground-based interference to the SSCS is a concern.

To determine ground-based interference, a Frequency Interference Measurement (FIM) was conducted for the frequency band 410-420 MHz on STS-56. STS-56 was a 57° inclination, 300 km altitude SSO mission that collected frequency interference data over several regions of the earth using a spectrum analyzer (ref. 3).

A ground-based interference model was built for DECAT consisting of statistically placed ground stations distributed around major cities, mostly in the industrialized nations, especially the United States, Europe, and Japan. The 410-420 MHz band allocations consist of mobile and fixed low-power stations and were modeled with two different types of ground stations. A data base was created of 693 cities, with location and population information included. Ground stations were distributed according to population. Location information was obtained from a world atlas, and population estimates and yearly growth rates were found in various almanacs (ref. 4). The number of ground stations assigned to each city was computed by estimating the current number of assignments in the U. S. in the 410-420 MHz band, and assigning a portion of them to each U. S. city based on the city's population.

DECAT was used to simulate the STS-56 FIM and to test the functionality of the ground station interference model. The SSO antenna was fixed in the +z direction (toward the Earth). A sample of the received interference power from the ground stations was taken every 3 minutes to model the sample rate of the spectrum analyzer. The spectrum analyzer bandwidth was 300 kHz so 33 DECAT receivers were used to sweep the 10-MHz band. The noise floor was set to -95 dBm, which is approximately equal to the measured interference seen by the spectrum analyzer over the oceans. Results of the simulation compare favorably with the FIM experiment.

3.4 Usage of the UHF Ground Based Interference Model

To illustrate the functionality of the ground interference model (ref. 5), DECAT was used to compute the signal-to-interference plus noise ratio $S/(I+N)$ for the EVA to Space Station SSCS

link. A DECAT simulation was run which placed the Space Station in a 28.5° inclination, 330-km altitude orbit for a 4-hour period. An EVA astronaut was placed along the Space Station truss with a fixed orientation facing in the orbit direction. Antennas were placed on both the EVA and Space Station and the system parameters input into DECAT. Results show that when the SSF is over land (especially Europe), the EVA-SSF link degrades below the required threshold needed to maintain the specified bit error rate (BER). The link availability, defined as the percentage of time the EVA to Space Station link is not degraded below the required threshold was computed.

This simulation serves as an example of how the ground-based interference model can be used to predict UHF link degradation.

4. CONCLUSIONS

DECAT has provided RF coverage analysis for the Space Station, Space Shuttle, and other advanced space program communication systems. DECAT's flexibility has allowed a wide range of analyses to be performed accurately and efficiently. Plans have been developed to enhance the capabilities of DECAT to meet new project requirements. DECAT will be distributed by NASA's Computer Software Management and Information Center in the future so that other organizations can make use of DECAT's capabilities for their communication system simulations and analyses.

5. ACKNOWLEDGMENT

The authors wish to acknowledge the CSSL manager, Laura Hood, for her assistance in developing DECAT and her major contributions to this paper. They would also like to thank their colleagues Shayla E. Davidson, Dr. Y. C. Loh, Penny Saunders, and Anna Kettler for their valuable comments on this paper.

6. REFERENCES

1. Best, R. E.; Land, K. P.; Gadd, W. C.; and Steel, D. J., "DECAT User's Manual Version 3.0", JSC-25697, NASA/JSC, Houston, Texas, March, 1992.
2. Kroll, Q.D.; Steel, D.J.; Garcia, C.L., "Communications Coverage Analysis for the Integrated Operations Scenarios Mission Build (MB)-1 through MB-7 and Utilization Flight (UF)-1 for the Assembly/Contingency Subsystem (ACS) and the Space-to-Ground Subsystem (SGS)", EE7-93-301, LESC-30851, NASA/JSC, Houston, Texas, June, 1993.
3. Larsen, D. W., "Data Package and Analysis of STS-56 Ultra-High Frequency (UHF) Frequency Interference Measurement for the 410-420 MHz band", EE7-93-708, LESC-30866, NASA/JSC, June, 1993.
4. Crandall, G. A. , "Spectrum Resource Assessment in the 406.1 - 420 MHz Band", NTIA 82-91, National Telecommunications and Information Administration, January 1982.
5. Larsen, D. W., "Space Station Freedom (Manned-Base) Circuit Margin Data Book", EE7-93-602, LESC-30871, NASA/JSC, Houston, Texas, June 1993.

SOFTWARE-IMPLEMENTED FAULT TOLERANCE IN COMMUNICATIONS SYSTEMS

Dr. Rex E. Gantenbein
Associate Professor
University of Wyoming
Department of Computer Science
Laramie, Wyoming 82071-3682

ABSTRACT

Software-implemented fault tolerance (SIFT) is used in many computer-based command, control, and communications (C³) systems to provide the nearly continuous availability that they require. In the communications subsystem of Space Station Alpha, SIFT algorithms are used to detect and recover from failures in the data and command link between the Station and its ground support. The paper presents a review of these algorithms and discusses how such techniques can be applied to similar systems found in applications such as manufacturing control, military communications, and programmable devices such as pacemakers.

With support from the Tracking and Communications Division of NASA's Johnson Space Center, researchers at the University of Wyoming are developing a testbed for evaluating the effectiveness of these algorithms prior to their deployment. This testbed will be capable of simulating a variety of C³ system failures and recording the response of the Space Station SIFT algorithms to these failures. The design of this testbed and the applicability of the approach in other environments is described.

INTRODUCTION

Nearly continuous availability is important to many computer-based command, control, and communications (C³) systems. To maintain their availability in case of a failure in a component, most of these provide redundant backup components that can be configured into the system. The detection and diagnosis of failures, the isolation of the failed component(s) of the system, and the restoration (or recovery) of the system to a functional state are collectively known as *fault tolerance*.

For low-level components (switches, gates, etc.), the detection and diagnosis of failure and the reconfiguration of the system are simple enough to be done with hardware logic. As systems get increasingly complex, however, the possible causes and results of failure increase, as do the potential responses to a detected failure, to the degree that software subsystems must manage the diagnosis and recovery. Such *software-implemented fault tolerance (SIFT)* is essential in maintaining availability in complex systems.

Central to the problem of providing SIFT is the problem of designing the algorithms used in the software to carry out the failure detection, isolation of failed components, and recovery in the system. If these algorithms are incorrect or incomplete, then the system they are intended to "protect" from failure is still vulnerable and unreliable. It is essential that the design of such algorithms be thoroughly evaluated and tested both before and after implementation to ensure that the software operates correctly when installed.

This paper describes the design of SIFT algorithms intended to maintain the availability of the communications system aboard Space Station Alpha. The potential causes of failure for this system are considered, and a computer simulation now being constructed to test the effectiveness of the SIFT algorithms in preventing such failure is discussed. The paper concludes by showing how SIFT can be effectively applied in other environments.

COMMUNICATIONS IN SPACE STATION ALPHA

The Assembly/Contingency Subsystem (ACS) is the primary communications link for the Space Station. It provides two-way audio and core data communications and supports ground-based tracking of the Station. The ACS is composed of two encapsulated *strings* of components. The primary components, or Orbital Replacement Units (ORUs), of each string, as shown in Figure 1, are:

- the ACS baseband signal processor (ACBSP), which provides the interface between the ACS and other onboard computer and audio subsystems,
- the ACS transponder (XPDR), which converts ACS information between analog radio frequency (RF) and digital form, and
- the ACS radio frequency group (ACRFG), which provides for transmission and reception of S-band signals between the Station and the ground via the Tracking and Data Relay Satellite System (TDRSS). This group consists of a high-gain and a low-gain antenna and the amplifiers for these antennae.

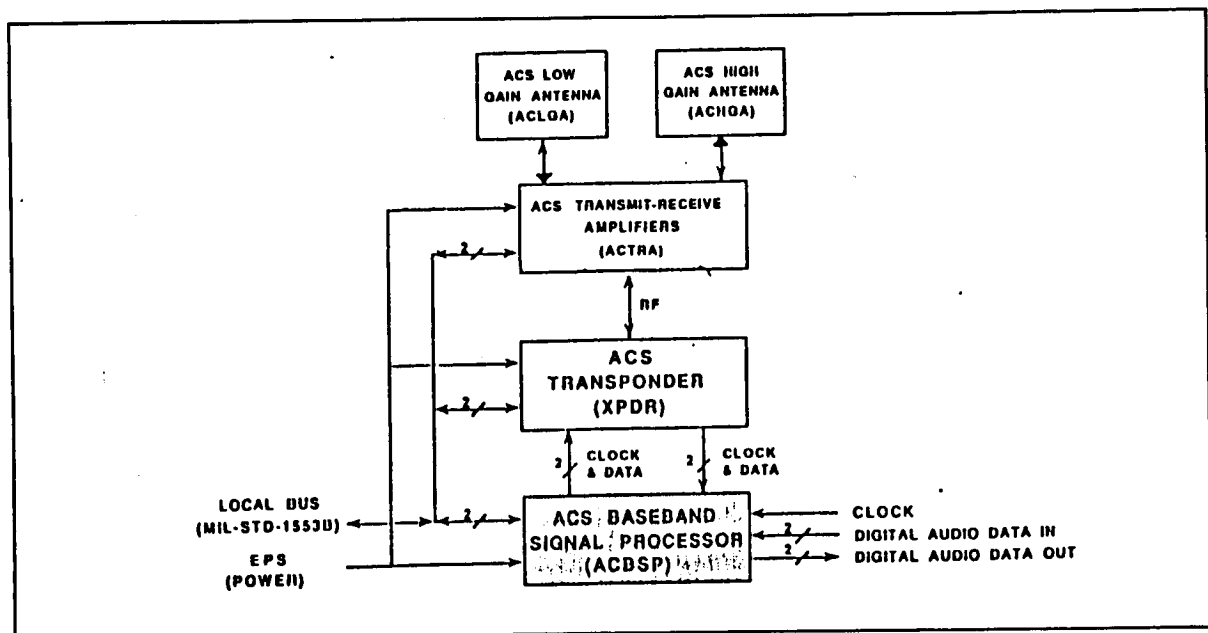


Figure 1 - Block Diagram of the Assembly/Contingency Subsystem (ACS).

Failures in any of the ORUs, or in the external units that provide the connections to both the ground and other subsystems onboard the Station, have the potential to interrupt the link between the Station and the ground, effectively isolating the Station until such time as the linkage can be restored.

The absence of communications, even for extended periods, is under some conditions an inconvenience, while in other cases, such as extra-vehicular activity, it may be life-threatening. In any event, it is clear that some facility must exist to accomplish SIFT in this subsystem, so that permanent communications loss can be avoided. Furthermore, during the early unmanned stages of the Station's deployment, all control will be initiated from the ground, making it difficult to repair any problems on board without communications. Even in the manned operational phase, facilities for manual diagnosis and repair of the ACS will be limited, placing continued importance on the system's ability to detect and recover from failures autonomously.

For these reasons, among others, the ACS design includes a Failure Management subsystem (ACFM) for detecting the loss of communications capabilities in the ACS and providing mechanisms by which communications can be

restored in the event of failure. The ACFM collects equipment status and performance data, and performs failure management functions including equipment and string isolation and redundancy management.

The ACFM itself is composed of three constituent capabilities (also called subcapabilities). Each of these subcapabilities provides a distinct set of services supporting SIFT in the ACS:

- ACFM External Control (ACFXC) is responsible for accepting failure indications from sources external to ACFM and establishes the monitoring mode for failure detection and isolation, determining the behavior of the ACFM in response to detected failures.
- ACFM Failure Detection and Isolation (ACFFDI) collects the raw equipment sensor data and status, and analyzes it to detect and isolate equipment faults within the ACS.
- ACFM Failure Recovery (ACFFR) is responsible for managing the redundant ACS resources at the string level for recovery from a detected failure.

Information used by the ACFM may be generated internally or input from external sources. There are two external sources of information, which may also be updated by the ACFM:

- the Run-Time Object Data Base (RODB), which maintains system management information from the Station subsystems in a common area and is the primary interface with the Tier 1 system, which provides Station capabilities such as the onboard system executive, the crew interface, and the Station control center, and establishes the operating environment for the ACS; and
- the ACS data base, which maintains ACS information that is not needed outside the ACS.

Additionally, the ACFM may exchange information with the ACS Services Management (ACSS) subsystem, the second subsystem in the ACS, both directly and through the RODB and ACS data base.

FAILURE DETECTION AND RECOVERY IN THE ACFM

The overall *modus operandi* of the ACFM can be characterized as "if it ain't broke, don't fix it." Even if external or component faults have occurred and are detected in the system, no failure is declared as long as commands are successfully being transmitted and processed on the Station. The reasoning behind this approach is that it is preferable to have communications (even if failures exist within the system) than to attempt recovery that might not succeed. This approach also handles the problem of incorrect sensors triggering an unnecessary recovery.

Loss of signal

An interruption in the information flow between the Station and ground support is termed loss of signal (LOS). Both the forward (ground to Station) and reverse (Station to ground) communications links pass through the TDRSS. This system consists of two geostationary satellites, TDRS-E(ast) and TDRS-W(est). As shown in Figure 2, the Station passes from the range of one TDRS to the other as it orbits the Earth. Once in each orbit, the interposition of the Earth between the Station and the TDRS satellites, which causes the Station to experience LOS, defines a zone of exclusion (ZOE). Upon emergence from the ZOE, the Station attempts acquisition of the signal (AOS) to reestablish communications.

LOS is detected in the ACFFDI through examination of the TDRS mode and the demultiplexer state, which determines the quality of the forward link. If the TDRS mode is "SEARCH" and the demux state shows "Degraded Frames" continuously for a specified TDRS Search Time, or if the TDRS mode is "ACQUIRED" and the demux state shows "Degraded Frames" continuously for the LOS Declaration Time, then LOS is declared by the ACFFDI. LOS caused by the Station's passing through the ZOE or by other, scheduled competition for use of the TDRSS is not considered a failure in the ACFM. However, LOS in the forward link that is *unscheduled* may result in

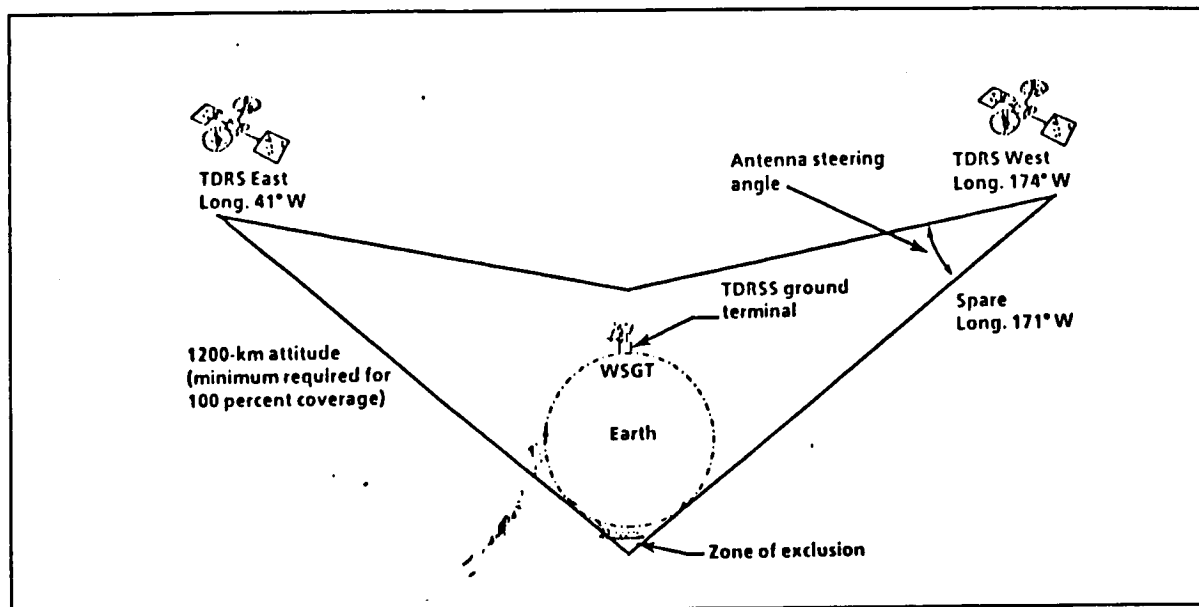


Figure 2 - The Tracking and Data Relay Satellite System (TDRSS).

ACFM declaring a communications failure. A failure may also be declared if the Station is unable to achieve AOS when emerging from the ZOE or passing from the range of one TDRS to the other.

To determine if a detected LOS is scheduled, the ACFM checks an internal TDRS event table, which lists the times when the Station is expected to be in (or near) the LOS period. An uncertainty time is applied to the event times for entering and leaving the period. If no event is found that indicates that the LOS is scheduled, then ACFM declares the LOS to be unscheduled.

Initiating recovery

The detection of an unscheduled LOS initiates additional effort in the ACFM to isolate the cause of the LOS and to recover communications. The conditions that lead to declaration of failure are described in the following section. Recovery in the ACS is achieved by switching from one string of ORUs to another. These strings are functionally identical, but receive and transmit on slightly different frequencies to avoid interference. For this reason, the ACFM must keep track of which of the two strings is the *active* or operating string and which is the backup or *alternate* string.

Recovery from communications loss can take place at two levels. At the string level, operations are switched from the active string to the alternate; if successful, the active and alternate indicators are switched so that the active string becomes the backup and vice versa. Recovery may also take place at the data rate level. If external or internal failures make it impossible to use the High Data Rate (HDR) communications facility (which depends on the steerable high-gain antenna) on either the active or alternate string, another configuration, which uses the omnidirectional low-gain antenna, can provide a Low Data Rate (LDR) link.

To avoid spurious recovery and unneeded reconfiguration, as well as to provide some flexibility in the recovery actions, all ORU component events in the system have an associated *criticality*. Only an event identified as "CRITICAL" will initiate recovery. This status is maintained through a criticality table in the RODB. This table (which can be modified from the ground) determines whether a failure in a given component will initiate recovery or simply be noted. In the following descriptions of the failure modes, criticality will be mentioned where it affects the behavior of the ACFM.

FAILURE MODES FOR THE ACS SUBSYSTEM

Through an informal failure mode and effects analysis, a number of events have been identified as failure modes for the ACS. These events fall into three major categories, which correspond to the recovery strategies incorporated in the ACFM:

- pointing-vector failures,
- hardware failures, and
- extended loss of command link.

Pointing-Vector Failures

Once an unscheduled LOS has been declared, the ACFM will attempt to isolate the source of the problem that has caused the failure. One possible problem is that the Guidance, Navigation, and Control / Propulsion (GNC/P) subsystem is unable to correctly position the steerable antenna with respect to the TDRSS. This condition is detected in the ACFM by an unscheduled LOS and the ACS High-Gain Antenna Management (ACAM) subcapability of the ACSS indicating that the ACAM mode is "POINTING" (i.e., using the steerable, high-gain antenna) but ACAM pointing data is not available (this condition is derived by ACAM from AOS/LOS data in the RODB). Under these circumstances, the ACFM will declare an active string failure in the pointing vectors mode. Recovery from this error will be attempted, as described above.

Hardware Failures

Failures due to faults within components of the ACS are broadly categorized as *hardware* failures. Several combinations of faults can result in a failure of this kind. All active string failures in this mode will initiate an attempt by the ACFM to isolate the failure and recover communication. Examples of hardware failures include errors in physical positioning of the high-gain antenna, failures in the individual ORUs, loss of power, or a lack of activity in the multiplexer output. If any such failure is detected in the active string, then an active string failure in hardware mode is declared, and recovery initiated as described above.

Extended Loss of Command Link

Under some circumstances, such as malfunction in the transmission system or corruption of transmitted data, it is possible for information to be received by the Station that does not correspond to any known command. Occasional occurrences of this type are not a major problem, since the command can be retransmitted once it is known on the ground that the Station did not respond. However, it is essentially impossible for the ACS to differentiate between corrupt and incorrupt data, so the ACS may not detect LOS when, in fact, no usable information is being received. This can cause a long-term interruption of information or commands transmitted through the forward link.

To avoid this problem, the Integrated Station Executive (ISE), the software that performs the centralized coordination of the various command and control subsystems on the Station, includes an "egg timer" that is periodically reset by telecommand. If this timer expires (that is, it is not reset within its expected period), it indicates that the system has experienced extended loss of command link (ELOC). In response, a preset sequence of recovery actions, designed to discover if any path exists through the ACS that will support communications, is initiated. If such a path exists, then communications can be restored.

EVALUATION OF THE ACFM DESIGN THROUGH PROTOTYPING

As in any new software system, evaluation of the design before, during, and after its implementation is important to assure that the system will operate as correctly as possible. In the case of SIFT, this evaluation is even more important, as the protocols in this system must be able to restore functionality to the system they monitor and

control even if other software fails. In a complex system such as the ACS, formal verification of the SIFT algorithms' design is difficult to carry through, while exhaustive testing of the implementation is prohibitively expensive, if not impossible.

One approach for evaluating both design and implementation that is both cost-effective and reasonably thorough is *prototyping*, in which a system design is implemented at a high level in software, then evaluated to determine the effectiveness of its design in meeting the system requirements. This method has been shown to be useful in evaluating software designs, particularly for complex systems that must be highly reliable. Since the cost of reworking an implementation greatly exceeds that of reworking a design, it is clear that costs can be dramatically reduced if evaluation of a system can be carried out prior to the implementation stages.

This approach has been used to evaluate the SIFT algorithms in the design of the ACFM. A prototype of the ACFM design was initiated by the author as part of a NASA/ASEE Summer Faculty Fellowship at Johnson Space Center (JSC) in 1993, and is being continued at the University of Wyoming through support from a JSC Director's Grant. In the following sections, the prototype and the experiments being planned for it are described.

Components of the Prototype

The prototype is written in C and runs under the UNIX operating system. It was originally implemented on a SUN workstation in the Control and Monitoring Systems Development Laboratory (CSDL), operated by the Tracking and Communications Division at JSC. It has since been ported to a DEC workstation in the Distributed Computing Laboratory at Wyoming and expanded. Plans call for the complete system to be re-installed in the CSDL in the summer of 1994 for on-site testing and integration with other software tools.

The prototype is constructed with modularity and testing in mind, using an object-based approach in which the functionality of each of the subcapabilities was encapsulated as a set of source modules corresponding to that subcapability. The interfaces between the modules model the ACS data base and the RODB, and all interaction among the subcapabilities uses data in those two interfaces. Since the interfaces were designed first, the subcapability designs could be carried out independently.

Each module in the prototype contains code to handle the various SIFT activities specified for an associated subcapability in the ACFM. For example, the ACFXC subcapability handles events from external sources, such as calibration and configuration tests that are generated by ACSS, ORU power status changes signaled by the EPS, and so on. The ACFFDI subcapability, the most complex of the three, detects failure and, where appropriate, initiates recovery based on a variety of sensor data, such as the ACS equipment status summary, ORU BIT summary monitors, pointing vector data from ACAM, etc. When either of these two subcapabilities detect an error, a number of messages and other indicators are sent through the RODB to other subsystems.

Recovery in the prototype is initiated when the ACFXC or ACFFDI module declares active string failure. The failure, which will be in either hardware or pointing vectors mode as described above, is then acted upon by the ACFFR module. Given the nature of the failure, reconfiguration at either the active string or current data rate level is attempted. If this recovery does not succeed, or if the detection process is unable to isolate the cause of the failure, then the prototype indicates that ACS communication is not available, a condition that would be handled through the ELOC recovery mechanism. The prototype does not simulate ELOC recovery, however, only its detection.

Implementation and Testing

In addition to the modules for the ACFXC, ACFFDI, and ACFFR subcapabilities, there are support modules for various subtasks not directly related to SIFT, an initialization module, and a number of modules used as test drivers to inject faults into the prototype. These test drivers run as UNIX processes that accept input simulating various internal and external failures like those described previously in this paper. These "failures" are signaled to the prototype by setting a software flag that corresponds to an indicator in the ACS or one of its associated data bases.

When a change in the system state is detected by the prototype in either the ACFXC or ACFFDI modules, action as specified for the ACFM is initiated. The prototype thus detects the injected fault, issues appropriate messages and updates, and initiates recovery when conditions warrant.

With each execution of the prototype, the simulated system responses (which are implemented as character strings sent to a window on the workstation screen) can be collected for validation of the code against the specification. The prototype was tested to branch coverage using this method to ensure that all expected inputs produced the appropriate responses. A combination of black box and white box testing was used to design the test cases. A test plan consisting of 37 tests (some of which tested multiple cases of the same response: failure and restoration of an ORU, for example) was developed and used to carry out testing of the algorithms. Where revisions were made due to tests in later stages of the validation, regression testing has been used to assure that the changes did not affect any previously validated code.

To fully test the effectiveness of the ACFM algorithms, a test profile that accurately simulates communications of an orbiting space vehicle is being developed. Data from previous Space Shuttle missions has been collected that provides frequency and duration of outages, both scheduled and unscheduled. From this data, test scripts are being developed to test the ability of the ACFM prototype in preventing permanent loss of communications. Performance data is also being collected to determine the response time of the system to detected outages.

SIFT IN REAL-TIME CONTROL SYSTEMS

The communications control system for Space Station Alpha is one example of a larger class of computer-based C³ systems that must operate in the physical world; that is, they interact with sensors, gauges, motors, switches, etc., that control physical, electronic, or mechanical action. The industrial world has become increasingly dependent on such systems, usually referred to as *real-time control systems*, which have requirements for availability, reliability, safety, etc., similar to those in space-based systems. For example, the shutdown of a manufacturing line due to the failure of its control system means the loss of the output of the line, as well as the loss of the productivity of its associated workers. The shutdown of a nuclear power plant's control system could mean irreparable environmental damage, even sickness and death. Unfortunately, these systems have typically relied upon proprietary hardware and software, fine-tuned over a number of years through trial and error [1]. As a result, many of the organizations are forced to retain out-of-date technology.

The technology being developed for SIFT in space-based C³ systems can be applied to earth-based real-time control systems. To achieve the high availability required in such systems, redundant components can be used as "spares" to be configured into the system when the SIFT algorithms detect and diagnose a potential failure. While the algorithms for determining the existence and source of a problem vary from application to application, the techniques for controlling the redundancy can be applied almost directly. In a manufacturing control system, for example, redundant controllers at each station of a production line can be switched in when a failure in the "active" controller was detected. The problem here, of course, is being able to maintain a consistent state in both controllers so that the alternate can continue from where the active controller failed. This can be accomplished through the use of multiple processors that exchange state information periodically for detection of failure and reconfiguration on a controller's failure [2].

The backup approach can be similarly applied to field-deployed military C³ systems. If provided with multiple communications links among mobile ground-based stations, and the capacity for moving in replacement stations for any component of the system lost due to hostile action or internal failure, a computer-controlled system can withstand a number of individual component failures before communications are lost [3]. Since reliable communications are essential in the modern battlefield, this capability is extremely important.

Smaller-scale technologies can also benefit from the use of redundant components controlled by software. Heart pacemakers are, in a sense, small autonomous real-time control systems that must be able to function reliably over a long period of time. Most pacemakers today monitor the heart as well as control its beating, and thus are equipped with a telecommunications system that transmits collected data to receivers for analysis of the collected

data. The resemblance to an autonomous space vehicle is clear, and so, therefore, is the applicability of SIFT technologies in assuring reliable operation of the device.

Prototyping, as described above, can be equally well applied for evaluating the design of SIFT systems in these environments. Again, the approach requires an analysis of the potential failure modes of the system, design and implementation of SIFT algorithms that respond to the identified failures, and development of a test profile that will exercise the algorithms in both expected and unexpected (according to the requirements) ways. Prototyping has been used to evaluate the ability of a manufacturing control system design to detect and recover from both software and hardware failures [4].

CONCLUSIONS

While all software aboard a spacecraft like the Space Station is critical, the SIFT software must be more trustworthy than any other, since it is responsible for detecting and recovering failures generated elsewhere in the system. Unfortunately, the processes and mechanisms that can support this level of reliability in software are still inadequately understood. For these reasons, the validation and evaluation of the ACFM and similar SIFT software for space-based communications systems are important tasks that require further study.

Longer range goals involve the exploration of methods by which the reliability, performance, and error coverage in space-based and other C³ systems can be enhanced. In many environments where C³ systems must operate, it is almost impossible to predict what events will and will not occur that may have an adverse effect on the system. Rather than trying to design a system that responds in a predefined manner to a given set of failures, one might consider an intelligent system that chooses its response to a particular event from among several alternatives, using information about the system and environmental state to find the response that has the highest probability of maintaining or restoring the system's functionality. Providing this kind of information to SIFT algorithms might allow a higher degree of reliability than has been previously achievable in computer-based C³ systems.

ACKNOWLEDGEMENTS

The author would like to thank his colleagues in the Tracking and Communications Division at Johnson Space Center, Shayla Davidson, Sally Stokes, and David Overland, for their support and cooperation in this project, and Kent Gaylor of LinCom Corporation for his assistance in understanding and elaborating the requirements of the ACS. This work is supported in part by NASA contracts NGT-44-001-800 and NAG9-722.

REFERENCES

- [1] Sha, L., "Industrial Computing: A Grand Challenge," *Computer* vol. 27, no. 1, January 1994, 12-13.
- [2] Gantenbein, R.E., Shin, S.Y., and Wang, Z., "Software Fault Tolerance in a Distributed Real-Time Control System," *Proc. Fourth ISMM/IASTED Int. Conf. on Parallel and Distributed Computing and Systems*, Washington, DC, October 1991, 61-64.
- [3] Avizienis, A. (Ed.), *Application of Fault Tolerance Technology: Design of Fault-Tolerant Systems*, BM/C3 Algorithm and Processor Working Group Report, Rome Air Development Center, Griffiss AFB, NY, October 1989.
- [4] Gantenbein, R.E., Wang, Z., and Shin, S.Y., "Mechanisms for Supporting Distributed Software Fault Tolerance," *Proc. First Int. Conf. on Computer Communications and Networks (IC³N)*, San Diego, CA, June 1992, 270-274.

32-32
10/10

SIGNAL PROCESSOR DEVELOPMENTS BY PERSONNEL OF THE JSC SIGNAL PROCESSING SECTION

S. Douglas Holland
Tracking and Communications Division
NASA / Johnson Space Center
Holland@TCD.JSC.NASA.GOV

ABSTRACT

The purpose of this paper is to describe systems and components of systems developed by personnel in the Signal Processing Section of the Tracking and Communications Division. The scope of this includes past developments which are in current use in NASA flight operations and future developments which are targeted for upcoming NASA applications. These projects specifically are:

1. NASA High Definition Television (HDTV) Project
2. Video Codecs
3. NASA Electronic Still Camera (ESC) Project
4. Hercules Payload
5. Ku-band Communications Adapter (KCA)
6. Windows Drivers for Satellite Interfacing to Commercial Equipment
7. Advanced Statistical Multiplexers

The methods used to determine what projects should be done in-house as opposed to which should not is based in NASA applications versus commercially available systems to meet those applications. If a commercial-off-the-shelf (COTS) component or system is available which meets the need, the first choice is to use COTS equipment. If it is not, and there is a NASA requirement, it is developed in-house. This results in technology which is being developed which otherwise was not available. Personnel involved in these projects have been contacted by many commercial companies interested in licensing or obtaining the NASA design.

INTRODUCTION

In the systems described below, it will be pointed out what parts are NASA unique and thus have been determined to be research and development candidates. In some cases a commercial item might be available but it may not be able to be used due to size, weight or power concerns. In other cases, the NASA requirements may exceed the commercially available components capability. There may be expectations of commercial development of equipment in the future for some applications. This requires a balancing act between waiting for a commercial equivalent against developing needed items before they become available.

The following describes each of these individual systems in overview. For further information, contact the Project Engineer associated with each development.

NASA HDTV Project

Status - In process

NASA Project Engineering - S.D. Holland, Denise Richards, David Olson

The NASA HDTV Project combines COTS, in-house developed and TBD equipment. The TBD components come from the fact that the North American HDTV standard has not been finalized. This has kept full scale commercial development of HDTV components from starting. NASA is actively interacting with and studying the development of HDTV components and systems. From this it has been determined which parts of the NASA HDTV system will most likely be NASA unique. Figure #1 shows the NASA Shuttle HDTV system.

The projected COTS equipment is: Internal Cameras, HDTV monitors, ground HDTV recorder, the majority of the Ground Station, ground HDTV to NTSC converter, ground HDTV broadcast formatter, Internet network interface. All COTS equipment will still require some modifications for flight application. The projected In-house developed equipment is: the encoder (compression) / decoder (decompression), framer / de-framer, mux / demux, error detection / error correction (edec), on-board HDTV to NTSC converter, Ground Station software, specialized ground image processing hardware, non-standard on-board cameras. The TBD equipment is the external camera and on-board HDTV recorder.

The North American HDTV standard includes provisions for extensibility, scalability and interoperability. These parameters will play an important part in the NASA system. Many image formats, resolutions, spectral responses, frame rates and auxiliary data will be included in the design. The NASA HDTV Transition Plan provides for extended functionality of image formats compatible with the HDTV data format.

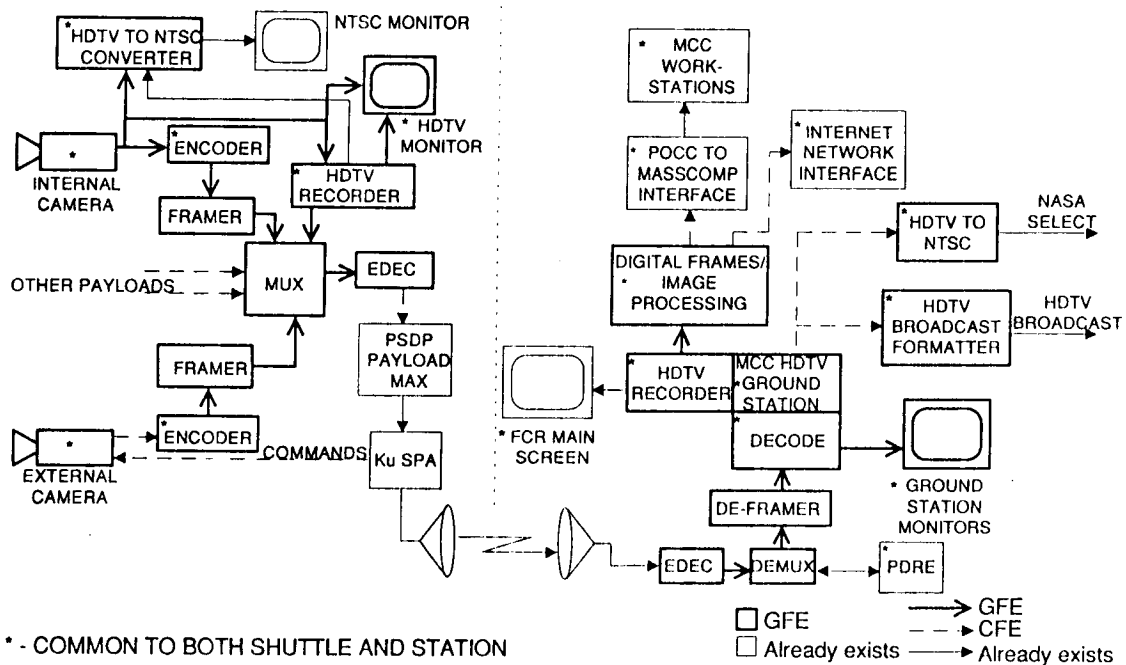


Figure #1

Video Codecs

Status - In process

NASA Project Engineer - S.D. Holland

NASA has some unique requirements for video encoders / decoders (codecs) which cannot be met by ones currently available. There are portable units available that do not have acceptable image quality. The models which do have acceptable quality are excessively expensive and are too large for NASA or other size, weight and power constrained applications. Due to these challenges along with those of HDTV, NASA is in the process of developing and testing codecs optimized for NASA applications.

Figure #2 shows the layout of the test bed and potential future flight video codec system. This design implements Motion Joint Photographic Expert Group (JPEG), Motion Picture Expert Group (MPEG) 1 and MPEG2. The resolution is full CCIR-601 resolution, 704 x 480 or square pixel 640 x 480, 30 frames per second, 60 fields per second. It can support NTSC composite or Y/C component inputs. By way of the multiplexer, multiple channels can be supported along with other payload data. Previously, during Shuttle operations, only one channel of video has been available. This system will provide many channels simultaneously and in digital form. An HDTV implementation will follow upon completion of the NTSC version. It will be compliant with the Grand Alliance specifications concerning HDTV data formats.

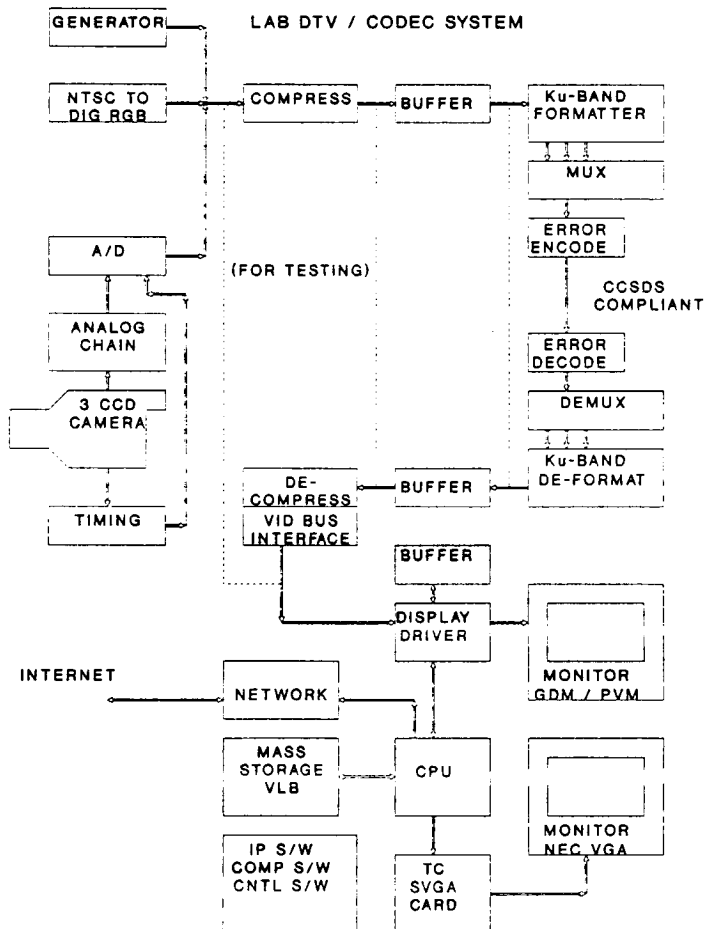


Figure #2

NASA Electronic Still Camera (ESC) Project

Status - Flight System (To date, 8 Shuttle Flights), Patents Pending

NASA Project Engineer - S.D. Holland

The NASA ESC Project was conceived out of the need to be able to downlink high resolution, high signal to noise ratio imagery from manned space flight missions in near real-time. The project was originally started for Space Station applications but has been employed in Shuttle operations.

Referring to Figure #3, the system consists of three components, a handheld battery operated camera, an onboard image processing system and a ground station. The ground station and onboard image processing system (Playback / Downlink Unit, PDU) consist of standard computer image processing, networking and storage capabilities. At the time they were developed there were some innovative features not available commercially, but today those features are showing up in commercial systems. The PDU was based on a laptop computer, modified to accept a removable hard disk containing images taken with the ESC. It was loaded with image processing software which allowed virtual imaging and other features not available at that time for a laptop computer. The software was developed under a NASA Small Business Innovative Research (SBIR) contract, which the developing company now sells as a commercial product.

The camera, on the other hand, was not standard and still retains unique features today not found in commercially available systems. The research, development, design and implementation of the camera was done by civil service personnel at JSC. There is currently a patent pending on this design held by the Project Engineer. Some features of the design cannot be disclosed here due to the patent status but an overview of it can be. The camera consists of two major components, the camera body and the electronics box. These are detachable for different mission objectives. The optical platform used is a Nikon F4 retaining almost all of the original functionality of the F4. The camera employs an astronomical grade charge coupled device, three phase, multi-pinned phase, boron implant with 100% fill factor and spectral response from 400nm to 1100nm. The camera is programmable so that it can support a variety of image sensors and sensor readout strategies under program control. For the shuttle missions to date, the camera has flown in a 1K x 1K configuration. It also easily supports 2K x 2K operation. The ASA equivalent is 330. Earth shots from Shuttle orbital altitudes have typically been shot at 1/1000th sec, f/8. The camera is fully programmable, has a 4 mode selector switch which executes different pre-flight setup programs. It includes both a standard data interface and a Shuttle payload digital downlink interface which runs @ 2 MBPS. The ESC stores its images on a removable hard disk making the number of images taken limited only to the number of hard disks available. When the flight system was designed the maximum capacity in a 2.5" hard disk was 42 MBytes (40 images). Today, disk of the same form factor can store over 500 MBytes.

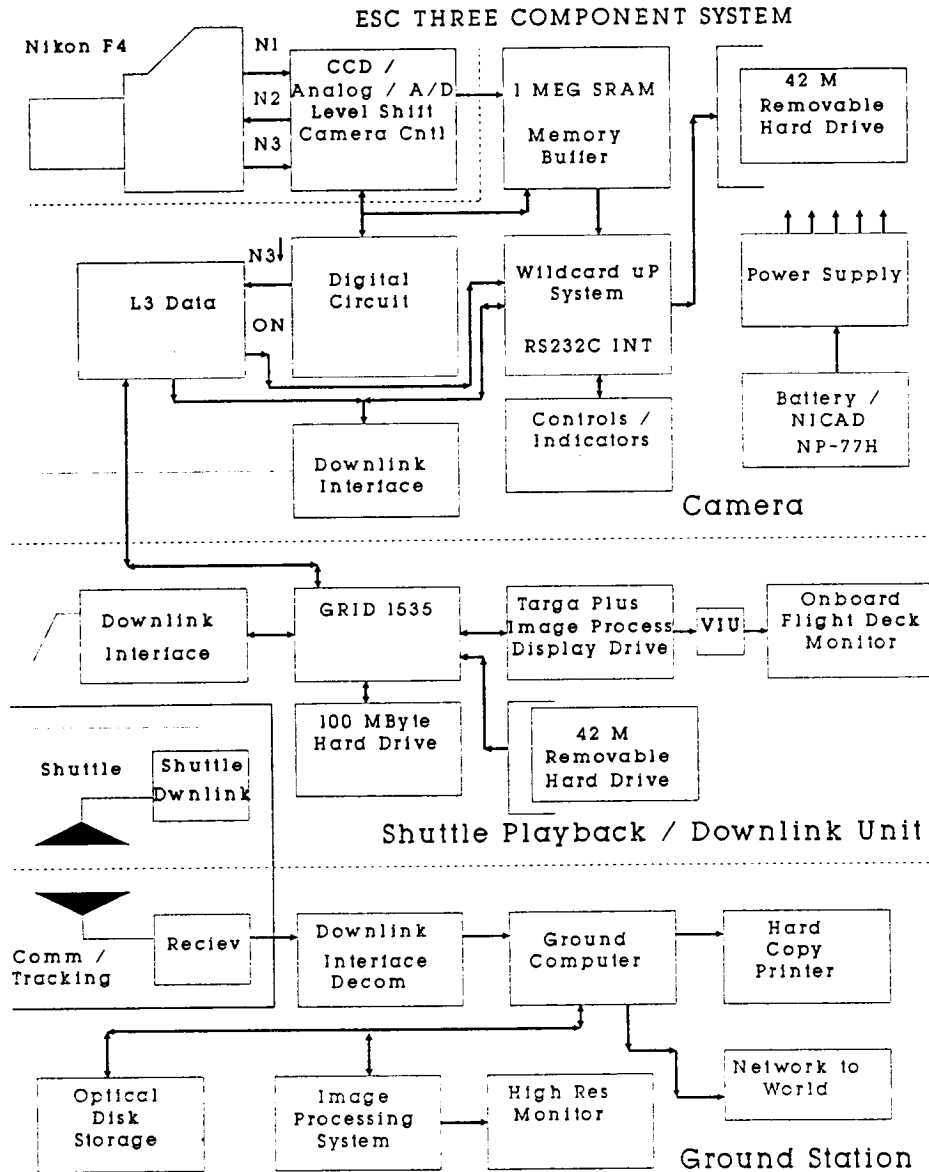


Figure #3

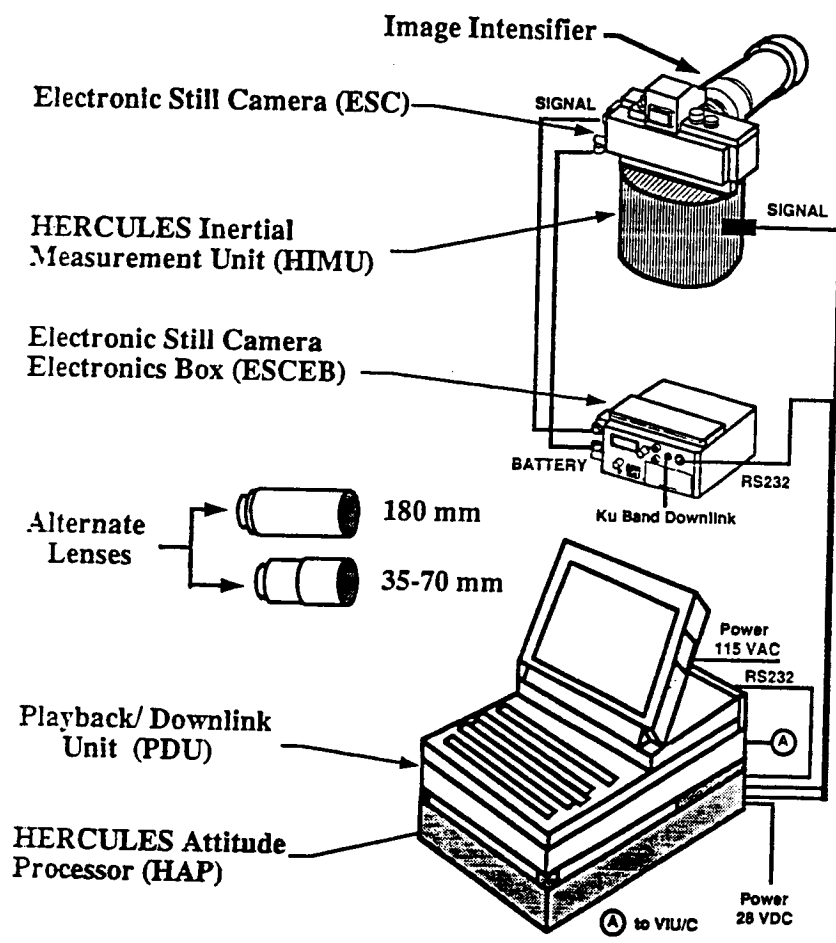
Hercules Payload

Status - Flight System (To date, 2 Shuttle Flights)
 NASA Project Engineer - S.D. Holland

The Hand-held, Earth-oriented, Real-time, Cooperative, User friendly, Location, Targeting, and Environmental System (HERCULES), was a joint NASA / Naval Research Laboratories (NRL) project. Referring to Figure #4, the onboard system consists of the Electronic Still Camera (ESC), Electronics Still Camera Electronics Box (ESCEB) and Playback / Downlink Unit (PDU) developed by NASA at JSC. NRL developed the Hercules Inertial Measurement Unit (HIMU) and the Hercules Attitude Processor (HAP). The ESC Ground Station made up the ground support equipment.

The Hercules Payload gives NASA the capability to have near real-time high resolution digital imagery with latitude / longitude location (L3) data included with each image. The L3 data is stored with each image on the ESC removable hard disk and is downlinked as the pictures are taken or at a later time if desired. The Hercules Payload gives geolocation information typically to within one nautical mile. Prior to this development, all Earth observing imagery had no geolocation data which required imagery analysts post mission to determine the location. With Hercules, this can be done as the pictures are shot.

The calibration of the payload is done by entering an Orbiter state vector and then performing a star alignment. The HAP then calculates location data based on this alignment and the current position of the HIMU. The HIMU employs three Helium-Neon ring laser gyros mounted orthogonal to each other. Inertial angular rates of change of the three mutually orthogonal axes, are measured by the HIMU. The data produced by the HAP at the time of camera shutter firing is sent to the ESCEB by way of the standard data port. This system can maintain an acceptable alignment for several hours before needing to be recalibrated.



HERCULES Configuration

Figure #4

Ku-band Communications Adapter (KCA)

Status - Flight Ready (First Shuttle Flight, STS-62, March 1994)

Project Engineer - Steve Schadelbauer

The KCA is a custom circuit designed by the Project Engineer. The KCA provides payload designers, experimenters, astronauts, and Shuttle flight operators with a simple, commercial interface to the Space Shuttle Ku-band Communications System. The KCA has been built and successfully ground-tested through the Ku-band communications system. In-flight operation of the KCA will be performed during the upcoming STS-62 mission. A block diagram showing the orbit to ground data flow is shown in Figure #5.

Prior to the KCA, Shuttle payload users requiring Ku-band communications had to learn the requirements of the Ku-band Subsystem and design custom circuitry for the Ku-band interface. Quite often the payload users are unable to design this interface correctly, resulting in expensive testing, re-design and possible re-flight of the experiment. The KCA was designed to remedy this problem by performing the processing for the Ku-band interface and providing payloads with a commercial, well-documented, well-understood generic computer interface. The KCA is a single printed circuit board which may be used at either end of the Ku-band system, flight or ground. The KCA provides the following features:

1. Computer interface is the Industry Standard Architecture (ISA) bus. The KCA is directly compatible with most personal computers available today. If the computer containing the KCA is not the end destination for the Ku-band data, it can be transferred from there by a wide variety of commercial means. Some examples are networks, other ports like SCSI, serial ports, parallel ports, floppy disks, removable hard disks, etc.
2. Software can be easily developed to control all aspects of the KCA operation and data exchange. For the STS-62 mission a software package running under Windows has been developed which allow multimedia, audio files, imagery and text to be uplinked and downlinked from orbit to ground in real-time during Shuttle operations.
3. The KCA represents a speed increase of 50 times the corresponding S-band uplink service and 500 times the corresponding S-band downlink service.
4. The KCA includes error detection and correction by way of a high performance Reed-Solomon error correcting code. Data integrity can be further enhanced by implementing computer-to-computer communications protocols, which the KCA supports by virtue of its ISA bus interface and bi-directional communications capability.

The KCA design can be readily adapted for many NASA communications applications. The functions performed in the KCA such as data scrambling to maintain receiver bit synchronization, frame synchronization, coding and automatic fill insertion to maintain constant transmission rates are good telemetry formatting techniques applicable to other RF systems. The KCA will be flown on STS-62 supporting 2 MBPS and 4 MBPS rates. The KCA data rates could be reduced for operation on the Shuttle S-band communications subsystem or it could be adapted for use with the Space Station Ku-band system. By the same reasoning, other RF communications could use the KCA for generic computer interfacing for a variety of needs and applications.

KU-BAND COMMUNICATIONS ADAPTER (KCA) DATA FLOW

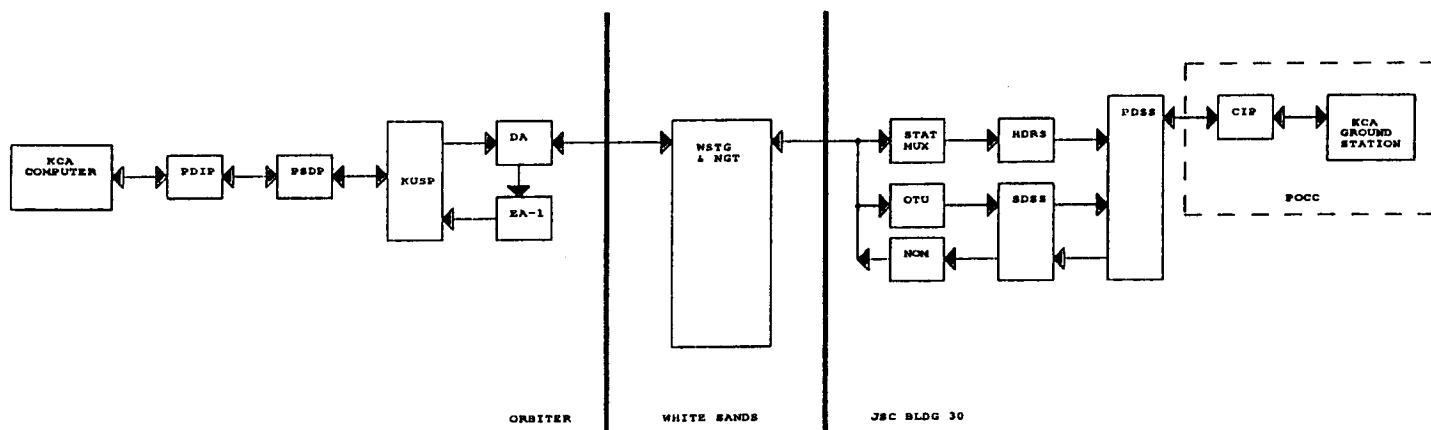


Figure #5

Windows Drivers for Satellite Interfacing to Commercial Equipment
 Status - Flight System (Used on all current Shuttle Flights)
 Project Engineer - Brett Parrish

In an ongoing effort to save money by using more commercial-off-the-shelf (COTS) equipment, NASA is currently flying a commercially available printer developed by Raytheon as an element in the text and graphics system. The system, Thermal Impulse Printer System (TIPS) shown in Figure 6, allows ground controllers to uplink visual images and textual information to an orbiting Space Shuttle during a mission and have it print out on a Raytheon printer. Crew members receive important information such as up to date procedures, mission strategies, goals and Earth observation images via the TIPS.

Ground operators develop messages and images to be uplinked on commonly available commercial personal computers using COTS software packages. Two paths exist for ground controllers to interface to the TIPS, the S-Band System and the Ku-Band System. In both cases, data routed to the TIPS makes its way directed to a card internal to the Raytheon printer. The communication interface board (CIB) constructed by the Signal Processing Section resides in the Raytheon printer, where it receives an interface to the printer's power supply. The CIB communicates to the printer and other devices through an external general purpose interface bus (GPIB). This provision allows onboard personal computers to connect to the GPIB bus and use the Raytheon as a printer. Additionally, data from ground systems directed to an onboard personal computer connected to the TIPS allows the TIPS to function as a one way modem.

The software that makes all of this possible is the Windows Driver developed by the Project Engineer which allows interfacing of the TIPS with the ground computer. This driver is also used on the ground to interface to a ground TIPS for ground testing of the material to be uplinked. This Project Engineer has also developed Windows Drivers for a future virtual instrumentation project. One example which is to be used on an upcoming mission is to allow a ground controller to operate onboard video equipment by way of a Windows representation of the onboard equipment. Software has been developed and is now working in a laboratory form which shows a 'virtual VCR' which looks like the front panel of a VCR with controls and indicators which are a virtual representation of the onboard system's controls and indicators. The user operates the controls by clicking on buttons and reads information corresponding to what the onboard system is displaying on the virtual display. Commercial companies have contacted this Project Engineer to inquire about licensing these developed and proven designs.

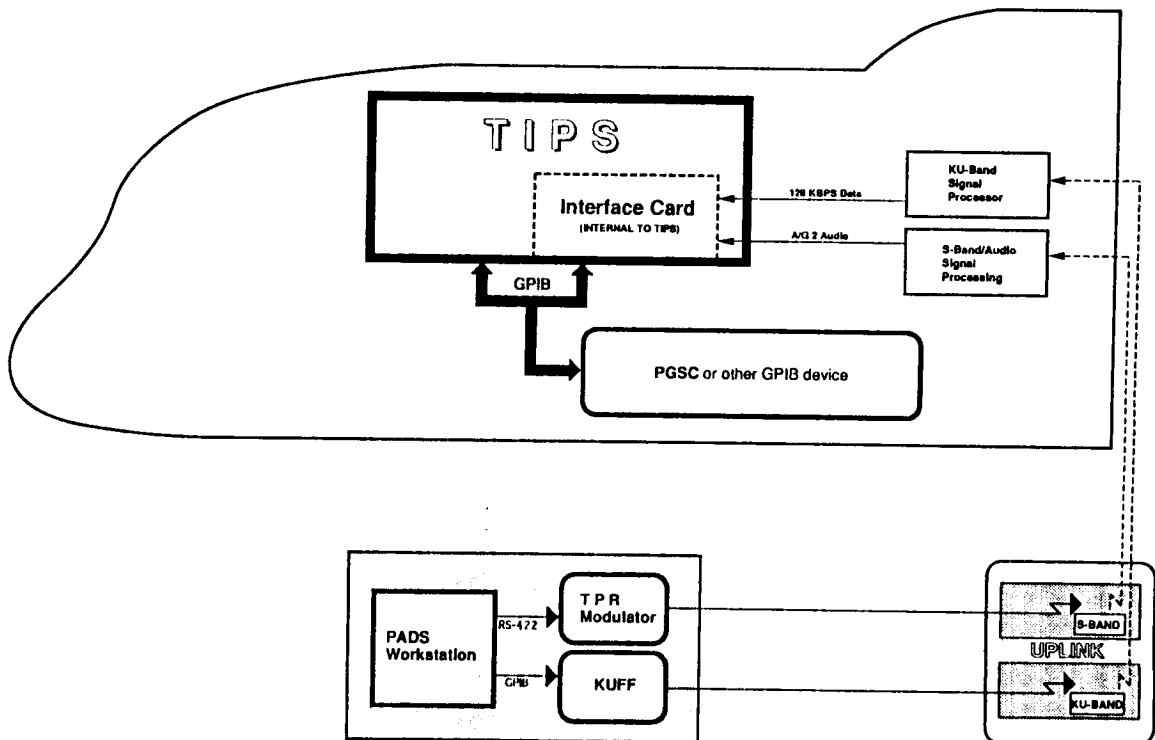


Figure #6

Advanced Statistical Multiplexers
 Status - In process
 NASA Project Engineer - David Olson

Advances in technology such as high rate computers and digital imaging techniques continue to increase the amount of information gathering and processing onboard spacecraft such as satellites, Shuttle and the future Space Station. Increases in information retrieved leads to an increase in the information required to downlink using existing communication channels. To use available digital communications systems without wasting resources, a multiplexing / demultiplexing system is often desirable. For the case of Shuttle, for example, without a multiplexer (mux), only one high rate (2 to 48 MBPS) user has access to the digital downlink at a time. If a mux was added, many users could use the high rate communications system simultaneously as long as the combined resulting available bandwidth is not exceeded.

A multiplexing system is currently under development at JSC which will allow multiple downlink data users to have simultaneous access to the digital downlink system. Refer to Figure #7. This system will allow multiple rate users, burst data and options to prioritize data streams. The output of the system conforms to the Consultative Committee for Space Data Systems (CCSDS), Advanced Orbiting Systems Recommendations. Several mux lab prototypes have been developed in the past and tested. The goal for the current project is to develop a versatile system which can be used for HDTV, DTV and other payloads. As with all NASA flight projects, much attention will be given to size, weight and power constraints.

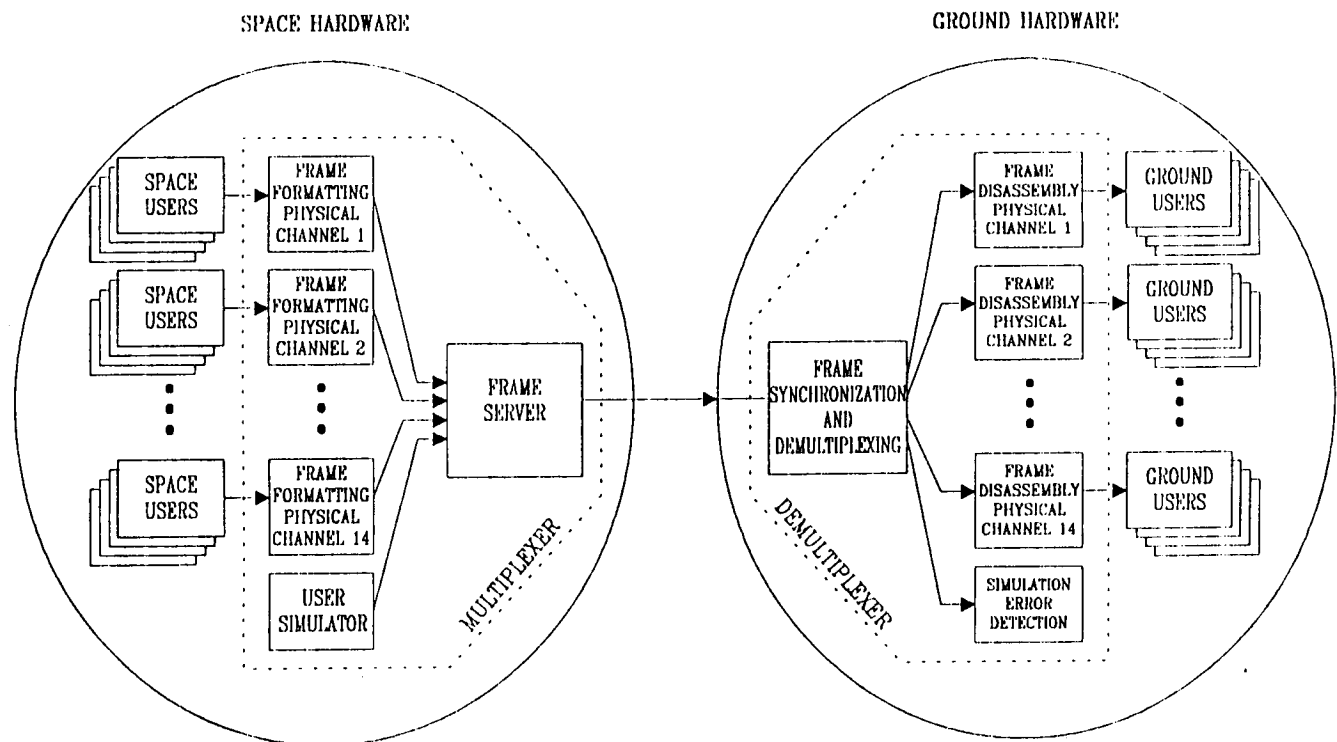


Figure #7

CONCLUSION

The systems and components of systems described above have been shown to have commercial applications and multiple uses. NASA develops required components when they are not commercially available. By definition, items developed would be unique and available for dual-use. The systems and components of systems described here have been developed by personnel in the Signal Processing Section of the Tracking and Communications Division. Further information can be obtained from the respective Project Engineer.

ACKNOWLEDGMENTS

Engineering Staff of the
Communications Processing Branch,
Signal Processing Section

H.A. Vang Branch Chief
R.W. Richards Section Head
K.E. French
L.E. Halley
S.D. Holland
K.D. McClain
D.M. Olson
B.T. Parrish
D.M. Richards
S. Schadelbauer
G.F. Steele

Development Staff of the NASA
Electronic Still Camera Project

Monty Moncrief - Mechanical Design
Rao Linga - Image Processing
Carroll Dodgen - Technician

REFERENCES

Holland, S.D., Electronic Photography and the NASA Experience, IS&T (The Society for Imaging Science and Technology), 46th Annual Conference, May 9-14 1993, Boston, Ma.

Holland, S.D., The NASA Electronic Still Camera System, International Geoscience and Remote Sensing Symposium (IGARRS), IEEE, Vol I, IEEE Cat #92CH3041-1, Library of Congress #91-72810, May 26-29, 1992, page 149.

Holland, S.D., Patent Pending, High Resolution Handheld Digital Electronic Still Camera, Patent Case MSC-21797-1, 1992.

Schadelbauer, S., Invention Disclosure, Ku-band Communications Adapter (KCA), NASA / JSC, 10/5/93.

Chien, Philip, Beam Me Down Discovery, Shutterbug, Titusville, Fl, Vol 21, No. 1, Issue 254, Nov 1991, Pages 54 - 58.

Larish, John, Creating Digital Photographs, Digital Photography Pictures of Tomorrow, Micro Publishing Press, Space Age Photography from the Space Shuttle, April 1992, Page 20.

Holland, S.D., Digital Electronic Still Camera, NASA Tech Briefs, New York, NY, Vol 17, No. 6 June 1993, Page 30.

NASA, Mission Accomplished, NASA Tech Briefs, New York, NY, Vol 16, No. 1, Jan 1992, Page 116.

NASA Office of Advanced Concepts and Technology, Electronic Still Camera, Spinoff 92, Washington, D.C., NP-201, page 32.

NASA Office of Advanced Concepts and Technology, Shuttle Operations, Spinoff 93, Washington, D.C., NP-211, page 46.

NASA Office of Advanced Concepts and Technology, Image Processing, Spinoff 93, Washington, D.C., NP-211, page 71.

Pitts, David, Earth Observations During Space Shuttle Mission STS-45, Mission to Planet Earth, International Journal of Remote Sensing, Hong Kong, Vol 7, No. 4, 1992, Pages 69 - 80.

Isbell, Douglas, Space Shuttle Earth Imagery Catalog at Record Size, Space News, August 10 1992, Page 9.

Isbell, Douglas, Shuttle's Secondary Payloads Also Include High-Tech Camera, Space News, Sept 9 1991, Page 24.

Osborne, Joe, Interfacing High Resolution Software with Existing Display Devices, Electronic Imaging International '91, Boston, Ma, Sept 30 - Oct 3 1991.

Osborne, Joe, The First On-board Digital Imaging System in Space, Electronic Imagery West '92, March 23 -26 1992.

Advanced Imaging, Shuttle Missions Turn to Hands-on Imaging & Image Processing, Page 8, Nov 1991.

McMillan, Tom, Image Processing in Space, Resolution - Professional Applications of Truevision-based Imaging Systems, Vol 3, No. 1, Jan 1992, Pages 41 - 42.

Arnold, H.J.P., Innovations in Imaging, British Journal of Photography, Oct 10 1991, Pages 21 - 23.

NASA, Space Shuttle Integration and Operations Office, Electronic Still Camera, Shuttle Integration Bulletin, No. 19, Oct 1991.

Higgins, R.F., Hercules Attitude Processor: Gyro Data Processing System for Real-time Geolocation of Images Captured by Astronauts, Naval Research Laboratories, Washington, D.C., May, 1993

omit

**Session C4: APPLICATIONS DERIVED FROM
CONTROL CENTER DATA SYSTEMS**

Session Chair: Marvin LeBlanc

S11-62
4-7

THE DASH DATA SHARING SYSTEM

David A. Hasan (hasan@gothamcity.jsc.nasa.gov)
Brooks Slaughter (slaughte@gothamcity.jsc.nasa.gov)

LinCom Corporation
1020 Bay Area Blvd., Suite 200
Houston, Texas 77058

Abstract

The Data Sharing system (DASH) project was recently undertaken at NASA Johnson Space Center to introduce distributed data sharing into the Mission Control Center (MCC). Although the project focused on MCC communications, the solution is a general one. This paper describes that project.

DASH allows applications to share data. It provides callable interfaces for applications wishing to export or import data. The system consists of several processes: *publishers* make exported data available to *subscribers*, which provide it to interested importing applications. A *network registration service* provides network location transparency, allowing the other processes to reside at arbitrary network locations. These processes act as intermediaries between external producing and consuming applications.

DASH has been demonstrated in the MCC where it transmits Shuttle electrical bus data from the Bus Loss Smart System to the Configurable Realtime Analysis System. In addition, the Failure Impact and Procedure Analysis system used DASH to transmit Shuttle remote manipulator data from an expert system to a version of CRANS. DASH is currently being used to integrate a knowledge acquisition application and the CLIPS expert system shell.

1 INTRODUCTION

This paper discusses the DASH data sharing system. This system was developed during 1993 as the result of a joint effort between the NASA Johnson

Space Center (JSC) Software Technology Branch (STB) and the JSC Real-Time Data Systems project (RTDS).¹ The objective of this effort was to apply distributed, cooperative computing technologies developed by the STB to operational software in the control center. Previously developed technologies included architectures for coordinating the *cooperation* of independently developed expert systems and *distributed* application-to-application communications techniques.

This previous work has led to technology prototypes and a number of software tools. A *hierarchical cooperation architecture* was developed which studied practical issues associated with integrating existing expert systems in such a way that by minimally modifying their implementations, they can be applied to a set of overlapping problem domains. In addition, a *peer-to-peer negotiation architecture* was developed which employed market-like mechanisms to coordinate a number of expert systems involved in allocating computational tasks over a network of workstations. The software tools which emerged from this work included a reliable UDP library, a process health monitor, graphical process deployment utilities, and a distributed version of the CLIPS expert system shell.

The importance of these technologies is assessed mainly on their relevance to the primary mission of Johnson Space Center, which is to support Space Shuttle and Space Station flight operations. However, the tools which have been developed and the applications which have been implemented need not be tailored to the specific context of JSC. An effort has been made in the STB to develop the technologies in such a way that they lead to generally useful software tools and generic application software.

The DASH project was conceived in 1993 as a mechanism of transferring some of the existing STB expertise and tools to the control center. The general objectives of the project were to develop a *general purpose framework* to allow one application to share data with others. In a sense, this represented a narrowing of the ambitions of the previous work (since the "cooperative" nature of the work was being reduced to "distributed computing"). The importance of this relatively modest project was that it would exploit previously developed STB technology and significantly benefit the control center, which relies heavily on voice loop communication between disciplines and does not currently exploit the network for interapplication communications.

The specific goals of the project were to

¹The RTDS project has since been renamed the Advanced Control Center Technologies project (ACCT).

- develop a parameter-based data sharing system,
- deploy that system on the existing (flight following) RTDS TCP/IP network of Unix workstations,
- use the system to introduce flight controllers to the advantages of electronic data sharing, and
- design the system and field a prototype in six months.

The initial DASH prototype was chosen to allow Shuttle electrical bus states to be communicated around the control center. Bus states are of interest to most disciplines, since the diagnosis of most other subsystems depends on whether or not those systems are getting power. The existing Bus Loss Smart System (BLSS), a G2 expert system, was chosen as the data-producing application for the DASH prototype. The existing Configurable Realtime Analysis System (CRANS) was chosen as a candidate data consumer, since it is used by many disciplines to analyze cause and effect relationships for the purpose of failure diagnosis.

2 SYSTEM DEFINITION

Since DASH was intended to be a general purpose system, it was designed with the following considerations

- it must be possible to tune the implementation of DASH to adjust its consumption of network and CPU resources,
- it must allow the various processes to run *anywhere* on the network
- it must allow the process locations to be changed without modification of the system or maintenance of cumbersome configuration tables,
- the protocols and application programmer interfaces (APIs) must not be specialized to the control center context, and
- the system should be based on a *client/server architecture*.

The meaning of the term “client/server” here is that only external (client) applications ever initiate communication with DASH. DASH never asynchronously contacts the clients. The motivation for this was that such an approach would simplify the incorporation of the DASH APIs into existing

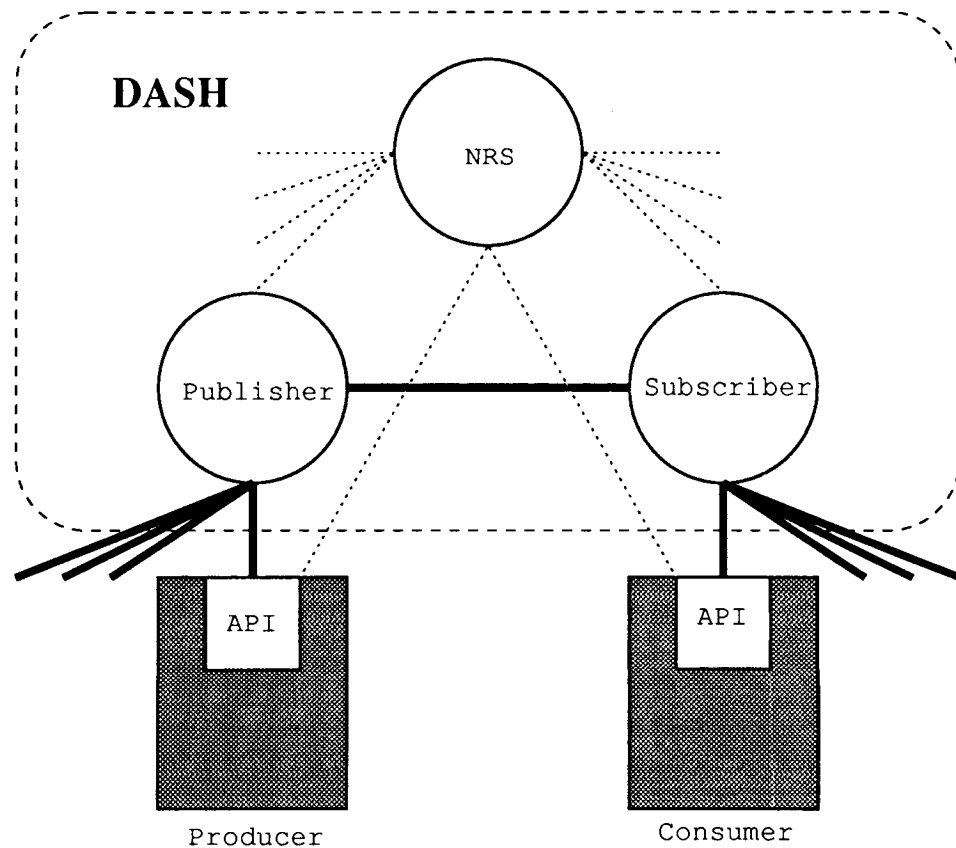


Figure 1: The DASH system design.

applications. Clients need not adopt a particular programming model in order to use DASH. Instead, all client—DASH communications are the result of an explicit call of a DASH API routine by the client code.

These general guidelines lead to the development of DASH as a set of cooperating processes. Thus, although DASH itself may be viewed by clients as a single server, actually it is implemented as several (possibly many) processes. This design is shown in Figure 1. In this figure, the shaded regions denote externally written applications. The white circles denote DASH processes, and the white squares inside the external applications represent the DASH APIs. The dashed rounded rectangle represents the abstract DASH server.

The DASH system includes a *network registration service* (NRS), which maps DASH process roles and names to process locations. The NRS is the mechanism by which network location transparency is achieved, since

interprocess communication is performed in terms of roles/names instead of specific network addresses.

In addition to the NRS, DASH consists of two other types of processes: *publishers* and *subscribers*. There may in fact be many of these. Publishers handle data coming into DASH. They distribute data to appropriate subscribers and serve as a single point of contact for applications wishing to export data. Subscribers handle outgoing data. They collect data from publishers on behalf of applications wishing to import data and act as a single point of contact for those applications.

External applications are classified as either *producers* or *consumers*. Producers generate data which are intended for use by other applications. Upon contacting DASH, producers specify the name of the "service" they will provide. The service consists of a set of named parameters which are available for consumption by other applications. Consumers import data generated by producers. Upon contacting DASH, consumers ask to subscribe to a particular service. Values for the parameters provided by that service are shipped to the consumer as the result of the consumer's calls to the DASH consumer API. This is not implemented as consumer-to-producer polling. The DASH system caches the data internally as the producer generates new values. The consumer API has functions which allow the consumer to determine if any parameter values have changed and to obtain them if they have. Thus the DASH consumer API is a change-based data communication mechanism.

In fact, external applications may be both producers and consumers. Such hybrid applications call both DASH APIs. Since the DASH system was originally conceived as a system for communicating parameter values from a producer to other interested applications, the design of DASH supports "one-way" flow of data from producers to consumers. Hybrid applications can achieve a two-way data flow, however, this is done at the expense of consuming one TCP/IP socket for each direction.

3 INTERFACES

Access to DASH is provided by two distinct application programming interfaces (APIs): one for use by producers and one for consumers. The designs of these APIs were motivated by the following concerns.

- they should present a conceptually simple "model" of interaction with DASH,

- they should leave the client in complete control of resource consumption (e.g., CPU cycles and network bandwidth),
- they should be easy to incorporate into existing applications, and
- they should provide extensive status information to the client in the event of system failure.

As a result of these concerns, the APIs consist primarily of synchronous remote procedure calls, although some asynchronous behavior exists in situations where the client need not be aware of certain error conditions.

Both APIs involve the concept of a “session” which is begun and ended by explicit calls to API routines. During a session, a producer may issue calls to an output routine whenever a service parameter value changes. This triggers a sequence of events in DASH which mark that parameter as changed. A consumer may check to see if any parameter values have changed (since they were last read) and subsequently choose to read them. This is essentially the full extent of the DASH programming model.

The C-language bindings to DASH consist of two header files and associated libraries. The producer library exports three routines: `PROD_Begin`, `PROD_Write` and `PROD_End` with some associated types being defined in the header file, `PROD.h`. The consumer library exports four routines:² `CONS_Begin`, `CONS_Updates`, `CONS_Read` and `CONS_End` with some associated types being defined in the header file, `CONS.h`.

In addition to the C bindings, a G2 interface to DASH was developed for the initial BLSS/CRANS prototype. This interface involved developing a separate process (known as a GSI process) for handling communication with the G2 producer on the one hand and calling the DASH producer API routines on the other hand. Since for the prototype only the producer application was based on G2, no corresponding G2 consumer interface was developed.

4 APPLICATIONS

The initial DASH prototype involved the transmission of Shuttle electrical bus states from an existing G2 expert system to an existing C/Xlib fault diagnosis system. The G2-based Bus Loss Smart System (BLSS) was originally developed to provide electrical power system flight controllers with

²Actually, the consumer API is slightly more elaborate than this.

a high-level, visual assessment of Shuttle bus states based on an analysis of Shuttle telemetry. Minor modifications to BLSS and the development of the G2/GSI process mentioned above extended its capability to include transmission of bus states to any interested DASH consumer. The prototype consumer was based on the existing Configurable Realtime Analysis System (CRANS). Modifications were made to the CRANS source code to invoke the DASH API routines in order to obtain bus state information. In some cases, this replaced independently coded (in some cases different) logic from that in BLSS. In other cases, the automatic provision of bus states replaced the need for manual manipulation of the CRANS interface by flight controllers.

After the development of the BLSS/CRANS demonstration of DASH, a simple DASH consumer was developed which issues a visible and audible alarm when parameter values change. The motivation for this simple consumer was that it uses virtually no resources (in contrast to CRANS) and could therefore easily be used to demonstrate the advantages of DASH and electronic data sharing to flight controllers. This simple consumer has been used along with BLSS during integrated Shuttle simulations.

The success of the DASH prototype led to its use in two additional systems. The first is a Shuttle remote manipulator system (RMS) failure analysis project in which an existing G2 expert system (the RMS Decision Support System - DESSY) transmitted its conclusions to an implementation of CRANS. The DASH libraries and GSI process were used in this project with only minor modification of the GSI required. The second additional application of DASH has been in an STB project involving the integration of the Task Analysis and Report Generation Tool (TARGET) and the CLIPS expert system shell. In this case, the DASH APIs are being used to implement two-way communication in which each program drives the other during different phases of operation.

5 CONCLUSIONS

A number of lessons were learned during the DASH project. Some of these are itemized below.

- TCP/IP protocols provide great flexibility by virtue of their wide availability, but higher level protocols are required for high level interapplication cooperation. It is not obvious that parameter-based protocols are sufficiently general to cover the general needs of such cooperations.

In addition, there are likely to arise situations in which the change-only, event-based approach of the DASH protocols is not appropriate. A *suite* of higher level protocols ought to be available for different and evolving needs of client applications.

- Connectionless protocols, like UDP, decrease the coupling between components (primarily by increasing the asynchronous behavior of the system); however, UDP is inherently unreliable. Thus additional protocol layers on top of UDP ought to be available. This is particularly important in cooperative architectures for which *multicasting* is a reasonable option.
- The need for high level protocols calls for the use of *commercially available tools* such as OSF DCE, Sun RPC, Sun XDR, and various “middleware” products. On the other hand, the use of such tools can severely constrain the appeal of a product if it requires clients to purchase or license additional commercial software.

In the time elapsed since the original conception of DASH, a number of significant developments have occurred which affect the deployment of distributed computing software in JSC control center operations. The new consolidated control center (CCC) design is based exclusively on a TCP/IP network of Unix workstations. There will be no “compute nodes”, and the mission operations computer (MOC) will be decommissioned in the near future — requiring the development of producer applications which run on workstations and export data previously provided by the MOC. User-developed software will play a big role in the CCC. Methods for integrating this software into a cooperative framework are already being deployed. A client/server system called ISP (Information Sharing Protocol) has emerged as the data sharing framework of choice in the CCC. ISP handles change-only telemetry and supports publishing and subscribing clients through a programming model based on Xt callback functions. As a result of the choice of ISP in the CCC, the continued development of DASH for control center use is unlikely. Nevertheless, for applications outside the control center, DASH continues to be a viable communications option. It assumes little about the context of a client application’s operating environment, and it imposes few requirements on the client application developer. Thus DASH is a system with real dual use potential.

ADAPTATION OF A CONTROL CENTER DEVELOPMENT ENVIRONMENT
FOR INDUSTRIAL PROCESS CONTROL

5/2-61

Ronnie L. Killough
Research Analyst
210/522-5557
RKillough@swri.edu

James M. Malik
Senior Research Analyst

Southwest Research Institute
P. O. Drawer 28510
San Antonio TX 78228-0510

ABSTRACT

In the control center, raw telemetry data is received for storage, display, and analysis. This raw data must be combined and manipulated in various ways by mathematical computations to facilitate analysis, provide diversified fault detection mechanisms, and enhance display readability. A development tool called the Graphical Computation Builder (GCB) has been implemented which provides flight controllers with the capability to implement computations for use in the control center.

The GCB provides a language that contains both general programming constructs and language elements specifically tailored for the control center environment. The GCB concept allows staff who are not skilled in computer programming to author and maintain computer programs. The GCB user is isolated from the details of external subsystem interfaces and has access to high-level functions such as matrix operators, trigonometric functions, and unit conversion macros. The GCB provides a high level of feedback during computation development that improves upon the often cryptic errors produced by computer language compilers.

An equivalent need can be identified in the industrial data acquisition and process control domain: that of an integrated graphical development tool tailored to the application to hide the operating system, computer language, and data acquisition interface details. The GCB features a modular design which makes it suitable for technology transfer without significant rework. Control center-specific language elements can be replaced by elements specific to industrial process control.

ADAPTATION OF A CONTROL CENTER DEVELOPMENT ENVIRONMENT FOR INDUSTRIAL PROCESS CONTROL

In the control center, raw telemetry data is received for storage, display, and analysis. This raw data must be combined and manipulated in various ways by mathematical computations to facilitate analysis, provide diversified fault detection mechanisms, and enhance display readability. These mathematical computations must be encoded in computer programs which retrieve the required data, perform the specified calculations, and then return the results to the data acquisition system. Historically these computations have been implemented in various ways ranging from high-level language implementations to computations implemented with the aid of a text-based computation builder.

NASA identified the need for a development tool which would provide a consistent platform for the development and maintenance of computations while taking advantage of up-to-date graphical programming technology. Southwest Research Institute was awarded a research grant to develop a proof-of-concept prototype for a new computation builder. A prototype development tool called the Graphical Computation Builder (GCB) was implemented as a result of that effort. Currently, Southwest Research Institute is under contract with Loral Space Information Systems to produce an operational version of the GCB as part of the Consolidated Control Center (CCC) development.

THE GCB DEVELOPMENT ENVIRONMENT

The Graphical Computation Builder provides an environment for the development and testing of programmatic implementations of mathematical computations. The GCB provides a language that contains both general programming constructs and language elements specifically tailored for the control center environment. The GCB is based on a technology referred to by various terms, including graphical programming, pictorial programming, iconic programming, and diagrammatic programming⁷. Graphical programming systems utilize graphical icons connected by control and/or data flow lines to represent the structure of the program. In graphical programming systems the diagram does not merely represent the program: the diagram is the program.

The GCB consists of the following primary components:

- Editor - facilitates the creation and maintenance of computations.
- File Manager - facilitates the storage, retrieval, duplication, and deletion of computations.
- Parser - defines a hybrid graphical/textual computation language.
- Function Library - provides powerful operators and functions not available in ordinary high-level languages.
- Code Generator - translates computation language into portable ANSI C.
- Compiler - invokes platform-native C compiler to produce executable computation.
- Tester - generates user-specified test data to exercise the computation.

Together, these components qualify the GCB as a complete development tool. The more salient components and their features are discussed below.

Editor

The GCB Editor consists of a canvas-like work area in which the computation developer constructs the graphical computation (see Figure 1). Through interaction with the context-sensitive palette, a developer can quickly construct an entire computation using only the mouse or other pointing device. Additionally, the developer has the option of also providing a plain-English equivalent of the expression contained in each symbol. The symbols in the work area can then display either the mathematical expression or the textual explanation of it (compare the flowcharts in Figure 1 and Figure 2). The GCB language is a hybrid graphical/textual language in that a computation consists of a flowchart composed of connected graphical symbols which contain textual expressions. Both the graphical flowchart and the textual expressions can be constructed using the mouse, except initial use of new variables. The GCB Editor provides common editing features such as cut, copy, and paste of any portion of the work area, but also provides the uncommon feature of automatic propagation of changes to variables (including changes in scope, data type, and variable name) throughout the computation. Computations are composed of one or more "elements" (subroutines), allowing the development of arbitrarily large computations in a structured fashion.

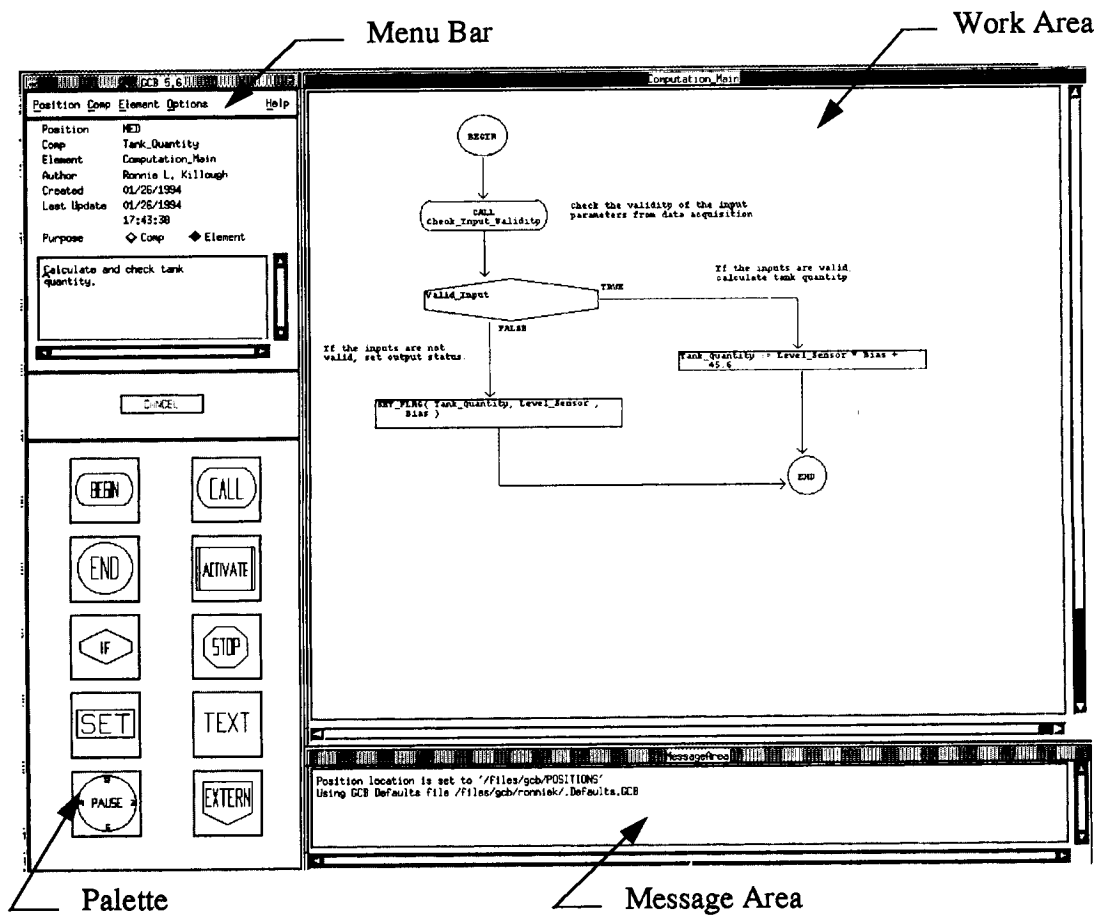


Figure 1 Graphical Computation Builder Interface

Parser

High-level computer language compilers, and particularly C, are well known for their cryptic error messages and also for their inability to detect many types of semantic errors (although some strongly-typed languages such as ADA have dealt with this latter problem). Additionally, an error in a given program statement is not reported until the whole program is compiled, not when the erroneous statement is entered. Therefore, even engineers who are very skilled in computer programming are unable to quickly diagnose the cause of compile-time errors in their programs resulting in time-consuming edit-compile-edit-compile iterations. The GCB Parser provides immediate and specific feedback whenever an expression contains syntactical or semantic errors. If the developer does not understand the error produced, on-line help is available to further explain the condition. This is particularly useful when operators which perform complex operations are involved, such as operators which solve systems of equations or which perform matrix multiplication. The Parser also maintains the state of each button on the palette as a visual indicator to the developer as to what language tokens are currently valid based on the current contents of the expression field (see Figure 2).

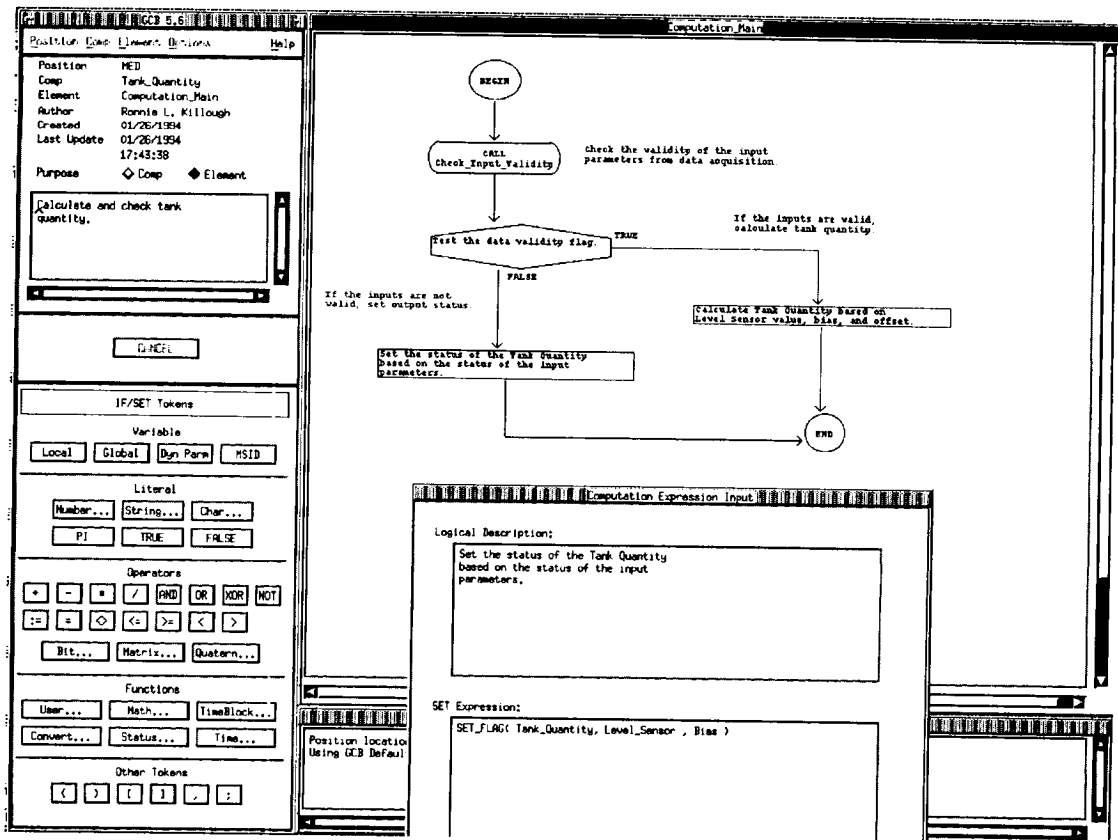


Figure 2 Context-Sensitive Palette and English Text Display Option

Function Library

Most high-level languages provide standard arithmetic, relational, and trigonometric functions and operators. The GCB extends this list by integrating additional functions into the computation language which perform powerful operations. These additional functions and operators include

matrix operators, unit conversion macros, quaternion operators, and status functions. In addition to the obvious advantage of availability to the developer, providing these additional functions allows more complex computations to be represented using less "real estate" without sacrificing (and in fact enhancing) the clarity of the program itself. Some of the additional functions and operators provided have specific applicability to the control center, others are of more general use.

The GCB also provides the capability to define a time block of data to maintain a user-specified number of values for a given input or calculated output. A suite of functions are provided which allow the developer to then calculate the average, minimum, maximum, least squares root, or standard deviation over the full range or specified sub-range of the time block.

The GCB provides extensibility features through user-defined functions. This feature allows computation developers to link externally developed object files or function libraries with their computation, and invoke those external functions from within the computation.

Tester

Many graphical programming systems and graphical display-building tools provide the capability to test the program or display by feeding randomly-generated data provided by the tool to the program or display. However, use of entirely random or even playback data is not adequate for testing the behavior and integrity of a computer program. To supplement this method of testing, a developer needs to be able to specify the input so that the output can be predicted. Actual results can then be compared to those predictions. The GCB provides a method by which a user can define the manner in which test data should be generated for a given input. This includes literal sequencing with repeat values, incremental/decremental sequencing with repeat values, and random data within a specified range.

SUITABILITY FOR TECHNOLOGY TRANSFER

Before discussing existing and potential applications of the GCB in the industrial setting, it is useful to evaluate how suitable an existing CCC product such as the GCB is for technology transfer to the industrial setting once a need is identified.

The GCB features a modular design which makes it suitable for technology transfer without significant rework. As previously discussed, the GCB "language" contains both general purpose constructs and language elements specifically tied to CCC needs. These CCC-specific language elements include both graphical symbol tokens and textual function and operator tokens, as shown in Figure 3. Because of the modular design of the GCB, these application-specific elements can be interchanged as the GCB is applied in various environments. Similarly, GCB-generated computations must interface with external interfaces specific to the CCC and which could also be replaced with other external interfaces.

In addition to modularity, the GCB has been implemented in an open systems environment. The GCB itself was implemented under the UNIX operating system in ANSI C with an X Windows/Motif user interface, and is currently supported on the SPARC and DEC/MIPS platforms with the DEC/Alpha soon to follow. Moreover, the GCB generates portable ANSI C source code which is currently compilable on 4 platforms - Sun/SPARC, DEC/MIPS, DEC/Alpha, and Concurrent/Masscomp. The code generation module of the GCB has also been designed in such a way so that changing the source code language generated by the GCB does not require extensive design alterations to the GCB.

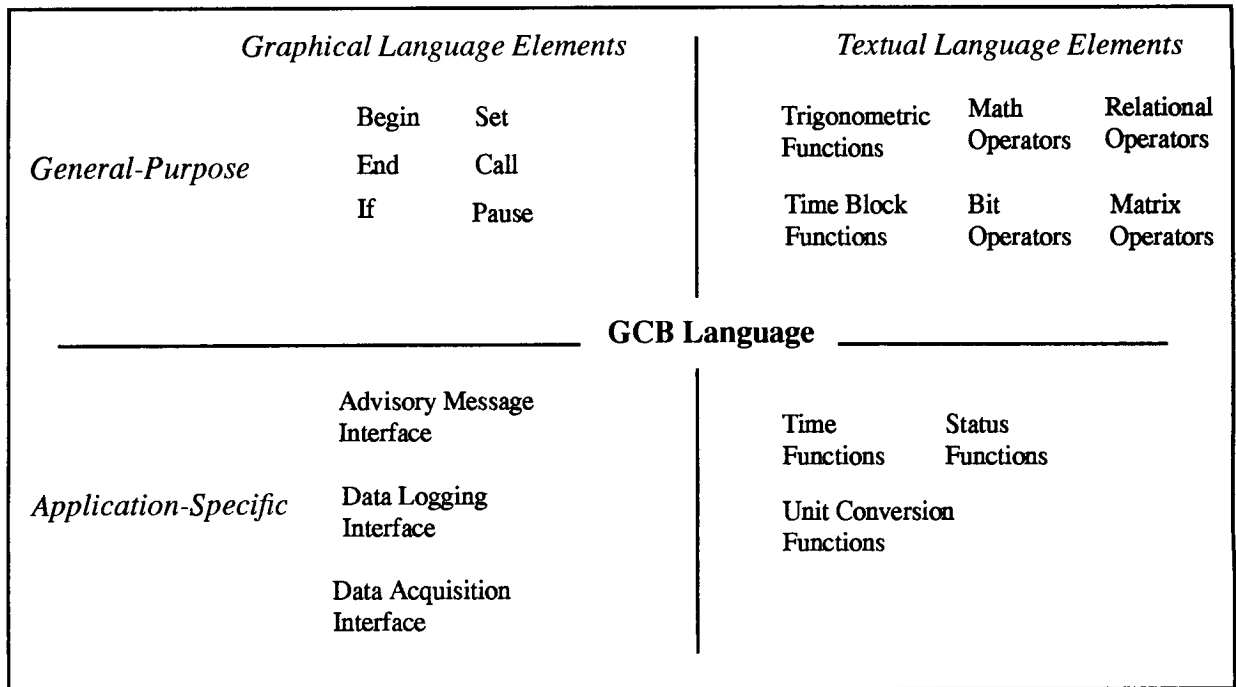


Figure 3 Modular Architecture of the GCB Language

APPLICATIONS FOR GRAPHICAL PROGRAMMING SYSTEMS

The benefit of graphical programming systems in the industrial setting can hardly be disputed. Fully 20% of the papers published in the 1993 *Proceedings of the ESD International Programmable Controllers Conference & Exhibition* dealt directly or indirectly with the application of graphical programming systems to the manufacturing process⁵. In a 1986 article summarizing the features of one such system, the author explained:

Once scientists or engineers must double as programmers, the multitude of programming languages and operating systems often obscures the orderly identification of software requirements, tools, and trade-offs. Essentially, engineers must depart from intuitive block diagrams and work instead in a detail-oriented way that the computer can understand, say, the sequential steps of a particular programming language.⁸

A sampling of the applications of graphical programming systems in the last several years are summarized in the following paragraphs.

Programmable Logic Controllers

A graphical programming technique called Grafset (also known as Sequential Function Charts or SFCs) has been proposed as an alternative to ladder logic and other common techniques in programming Programmable Logic Controllers (PLCs). SFCs are part of the International Electrotechnical Commission's draft standard for programming languages for programmable controllers (IEC standard 65B). SFCs are currently specified for use as the structuring tool for the representation of programs written in other PLC languages; however some propose to extend the SFC specification so that SFCs will be powerful enough to be used to directly program PLCs¹.

Device Controllers

Some companies manufacture products which contain embedded controllers which can be reprogrammed by their customers. One such company which produces variable speed motors developed a graphical programming system to program the motor controllers. This graphical programming system is made available to those customers who need to reprogram the motor controller. A graphical model was adopted so that those knowledgeable in the application of the controller could reprogram it even if they were not skilled in computer programming⁷.

Desktop Data Acquisition and Control Systems

Desktop PC-based data acquisition and control systems have been steadily migrating to graphical programming models in the last several years. The availability of PC operating systems incorporating graphical user interfaces and icons such as Macintosh System 7, Microsoft Windows, and OS/2 has helped fuel this trend. These systems have not so much replaced alternative control systems and languages as they have provided a practical alternative to manual operation of small-scale processes³. These systems ease the specification of control algorithms and data flow and also automate (and therefore hide) the intricacies of interfacing with data acquisition hardware such as GPIB⁸. Examples of application-specific functions included in these systems are PID (proportional-integral-derivative) control functions and device-specific controls common in small-scale process control applications.

DIFFERENTIATING FEATURES OF GCB FROM EXISTING SYSTEMS

The previous section summarized current applications of graphical programming systems. This section examines the features of the GCB which differentiate it from existing systems.

As discussed above, most existing systems apply graphical programming to desktop data acquisition and control of relatively small processes, or are used to program specific components such as PLCs or motor controllers. While the GCB could be used in these environments as well, its features make it most applicable to larger, distributed information and control systems. As automated manufacturing matures, distributed control of processes must merge with information systems to provide a seamless flow of information throughout an organization, allowing automation systems to be "physically distributed but act as a coordinated whole."² Moving beyond the PLC-level program, then, graphically-developed computations could gather data from various locations throughout a manufacturing plant and perform tasks as simple as computation of values for display (distributed information systems), to complex tasks such as combining the data in algorithmic fashion to control the behavior or speed of a component in another part of the plant or even another plant altogether (distributed control systems).

The GCB is in many ways a general purpose tool, implemented in an open systems environment, which produces independent executable programs that do not require the presence of the development environment to execute them. Hence, computations may, if desired, be replicated in a distributed environment without replicating supporting software and accompanying licenses. This characteristic distinguishes the GCB from most of the desktop graphical programming systems. Additionally, because of the demonstrably portable code produced by the GCB, GCB-generated computations would function well in the heterogeneous client/server environments which are likely to infiltrate large-scale process control².

CONCLUSION

The GCB concept allows staff who are not skilled in computer programming to author and maintain computer programs. The GCB user is isolated from the details of external interfaces and has access to high-level functions which allow more complicated computations to be encoded in a smaller representation while enhancing clarity. The need and desire for graphical programming environments such as the GCB is clear, and many such systems have already been implemented. However, most existing systems are either focused on programming specific classes of devices or are integrated components of small-scale desktop systems. The GCB, while developed specifically for the control center environment, features a modular design which would facilitate transfer to other applications. Additionally, its open systems foundation and the fact that it produces independent processes differentiates it from existing systems and positions it for use in larger distributed process control environments.

ACKNOWLEDGEMENTS

The authors of this paper wish to acknowledge the contributions of Ms. Lore Rice of NASA JSC, Mr. Tim Barton of Southwest Research Institute, and Mr. Jerry Ratner, formerly of Southwest Research Institute. Ms. Rice envisioned and championed the computation builder concept. Mr. Barton and Mr. Ratner conceived and implemented the prototype version of the GCB which served as the baseline from which the current version of the GCB emanated.

REFERENCES

1. Brandl, Dennis L., "Life After IEC65B - Real World Programming Using Sequential Function Charts", ESD TECHNOLOGY, Detroit MI, Vol. 54, No. 3, March 1993, pp. 28-32.
2. Davis, Don H. Jr., "New Strategies, New Realities In Customer-Driven Manufacturing", ESD TECHNOLOGY, Detroit, MI, Vol. 54, No. 3, March 1993, pp. 9-12.
3. Graham, Glenn, "Closed-Loop Computer Control", INTECH, United States, Vol. 36, No. 9, September 1989, pp. 78-81.
4. Lloyd, Mike, "Graphical Function Chart Programming for Programmable Controllers", CONTROL ENGINEERING, Nice, France, Vol. 32, No. 10, October 10, 1985, pp. 73-76.
5. PROCEEDINGS OF THE ESD INTERNATIONAL PROGRAMMABLE CONTROLLERS CONFERENCE & EXHIBITION, Detroit, MI, April 5-8, 1993.
6. Russi, Alan, "Graphical Programming Environments For Process Control", PAPER NO. 89-0512, ISA/89 International Conference & Exhibit, Philadelphia, PA, October 22-26, 1989.
7. Simonsen, Warren, "If Programming Is An Art, Why Aren't We Drawing Pictures?", PROCEEDINGS OF THE ESD INTERNATIONAL PROGRAMMABLE CONTROLLERS CONFERENCE & EXHIBITION, Detroit, MI, April 5-8, 1993, pp. 289-297.
8. Wolfe, Ron, "Block Diagrams and Icons Alleviate The Customary Pain of Programming GPIB Systems", ELECTRONIC DESIGN, United States, Vol. 34, No. 9, April 17, 1986, pp 125-132.

ADAPTATION OF CONTROL CENTER SOFTWARE TO COMMERCIAL REAL-TIME DISPLAY APPLICATIONS

S13-61

Mark D. Collier
Manager (Acting)
Southwest Research Institute
Post Office Drawer 28510, 6220 Culebra Road
San Antonio, Texas 78228-0510
voice: (210)522-3437
email: mark@trident.datasys.swri.edu

ABSTRACT

NASA-Marshall Space Flight Center (MSFC) is currently developing an enhanced Huntsville Operation Support Center (HOSC) system designed to support multiple spacecraft missions. The Enhanced HOSC is based upon a distributed computing architecture using graphic workstation hardware and industry standard software including POSIX, X Windows, Motif, TCP/IP, and ANSI C. Southwest Research Institute (SwRI) is currently developing a prototype of the Display Services application for this system. Display Services provides the capability to generate and operate real-time data-driven graphic displays. This prototype is a highly functional application designed to allow system end users to easily generate complex data-driven displays. The prototype is easy to use, flexible, highly functional, and portable. Although this prototype is being developed for NASA-MSFC, the general-purpose real-time display capability can be reused in similar mission and process control environments. This includes any environment depending heavily upon real-time data acquisition and display. Reuse of the prototype will be a straight-forward transition because the prototype is portable, is designed to add new display types easily, has a user interface which is separated from the application code, and is very independent of the specifics of NASA-MSFC's system. Reuse of this prototype in other environments is an excellent alternative to creation of a new custom application, or for environments with a large number of users, to purchasing a COTS package.

INTRODUCTION

NASA Marshall Space Flight Center (MSFC) utilizes the Huntsville Operations Support Center (HOSC) to support various missions. A new Enhanced HOSC is currently under design to support current and future spacecraft missions. The Enhanced HOSC is based upon a distributed processing computing architecture utilizing high-speed digital networks and high-performance graphic workstations. The system is based upon industry software standards such as POSIX, X Windows, Motif, TCP/IP, and ANSI C.

One of the primary capabilities required by the Enhanced HOSC is a real-time data display application. The purpose of this display application is to allow any of several types of users, such as programmers, flight controllers, scientists, and other users, to easily build real-time displays. This is a critical capability, because a significant amount of the effort performed during mission control involves using displays to review engineering parameters and scientific data in real-time. Without such an application, users would be required to either buy expensive display packages or generate X/Motif-based displays from scratch.

The Display Services application will provide this function for the Enhanced HOSC. The Display Services application consists of the following two programs:

- Display Generation - a program which allows a user to interactively build a data-driven display. The display will consist of a collection of simple graphics, input objects, and output objects to be driven with telemetry data. The primary function supported by this program is the editing capabilities used to position, size, and set object attributes.
- Display Operation - a program which accepts a display defined by Display Generation, connects to telemetry streams, and drives the dynamic output objects in a highly efficient manner. Display Operation also processes the input objects and performs the appropriate control actions.

Figure 1 presents an illustration of the basic architecture and data flow between the Display Generation and Display Operation programs.

The Display Services application is being developed in ANSI C, UNIX, and X Windows/Motif. The TeleUse User Interface Management System (UIMS) is being used to develop much of the Graphical User Interface (GUI). Although the application is a primarily custom product, Commercial-Off-The-Shelf (COTS) "widgets" are used to provide the majority of the complex graphic objects such as plots, bar charts, and virtual instruments. A widget is an X Windows standard used to implement a user interface object such as a push button, menu, slider, or plot.

Southwest Research Institute (SwRI) is currently responsible for developing the prototype for the Display Services application. The purpose of the prototype is two-fold:

- The prototype is designed to completely identify and resolve user interface requirements. By rapidly developing new user interface functions and providing frequent user evaluation opportunities, NASA-MSFC hopes to completely define the desired look-and-feel of the application.
- The prototype is being used to exchange design approaches and concepts with the prime contractor responsible for the final implementation of the system.

The goal of NASA-MSFC and SwRI is to develop a prototype which identifies all user interface issues and provides a solid baseline which will become the foundation for the operational system.

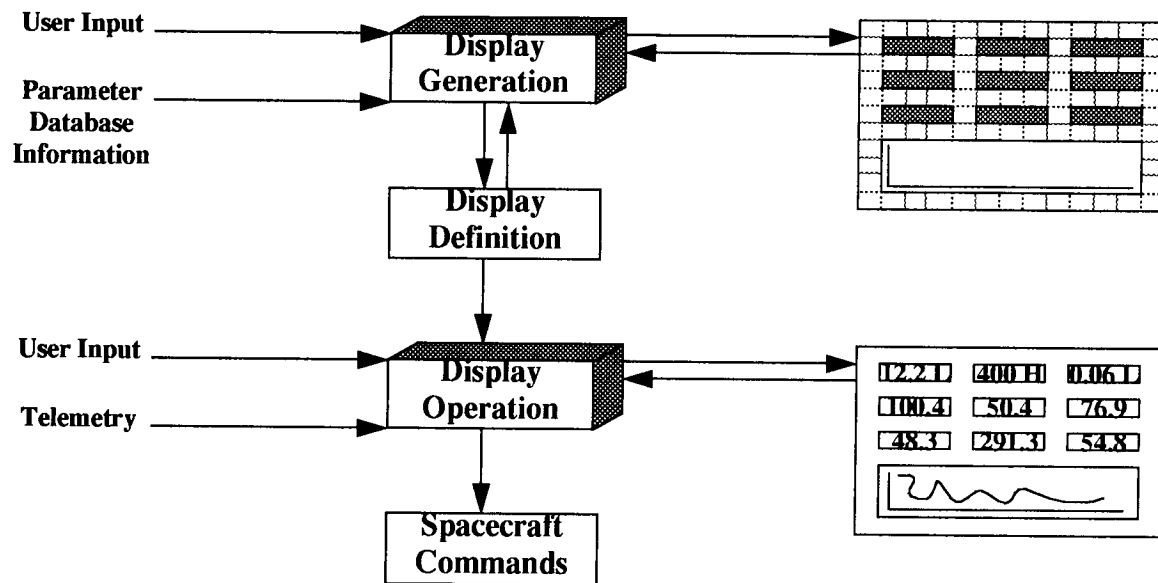


Figure 1 - Architecture/Data Flow Overview

Display Services Design and Capabilities

Before reuse issues can be discussed, it is necessary to describe the capabilities of the Display Services prototype. The following sections summarize the design and capabilities of the prototype.

COTS Versus Custom

Before the prototype effort began, SwRI and NASA-MSFC evaluated COTS data display products. Several COTS packages were identified, including SAMMI, DataViews, SL-GMS, TAE, and VAPS. Unfortunately, most of these packages are extremely expensive when full building capability is required on hundreds of workstations. This problem is compounded if licenses are required for off-site users such as scientists using the control center for monitoring of experiments. NASA-MSFC has also had some negative experiences with COTS tools not being supported well on selected architectures. Finally, most of these tools are very flexible, but this results in reduced performance and/or redundant capability. The very flexibility of these tools can also make it more difficult for users to complete simple, frequently required tasks.

A significant amount of the functionality of a real-time display application exists within the graphical "objects" which provide various forms of data display. Time-oriented plots, X/Y plots, bar charts, virtual instruments (meters, dials, gauges), pie charts, and other objects are complex and require a significant amount of code. To balance the COTS/custom content of the application, SwRI and NASA-MSFC have decided to base the application almost entirely on COTS widgets. By using Motif for most input objects and third party widgets for most dynamic objects, the cost of development has been dramatically reduced. The widgets currently in use include:

- Motif widgets are used for all input objects, including Push Buttons, Text Fields, Command Fields, and Sliders.
- Motif widgets are also being used for several simple output objects, including Toggle Buttons (for binary discretets), Radio Boxes (for multi-state discretets), and Sliders.

- Third party COTS widgets are being used for the majority of the sophisticated output objects, including Time-Oriented Plots, X/Y Plots, Pie Charts, Bar Charts, and Virtual Instruments (Meters, Dials, Gauges, etc.).
- One SwRI-designed widget is being used to support Dynamic Text Parameters and Static/Dynamic Primitive Graphics (Lines, Circles, Polygons, etc.).

SwRI developed the custom widget because no COTS widget could be found for this purpose. Development of a custom widget also allowed optimization of the critical text and graphic display capabilities.

The prototype currently uses the DynaGraphX data widgets from the V.I. Corporation. The DynaGraphX data widgets are basically a widget-compatible wrapper developed on top of the graphic objects provided by V.I. Corporation's DataViews package. This set of widgets was selected over competing widget sets because it provided a wider range of widgets and is well suited to real-time use.

Although the prototype is currently using DynaGraphX widgets, it will be possible to add new graphic widgets from another source or even to replace DynaGraphX if users do not feel that the widgets adequately satisfy requirements. The prototype implements a library layer which allows widgets from any source to be manipulated as generic objects. This isolates the specifics of a given widget set within one major area of the prototype.

Software Architecture

The prototype is completely based upon ANSI C, POSIX, X Windows, and Motif. No non-standard interfaces are used. The prototype depends upon the TeleUse UIMS as its GUI-building tool. TeleUse is used to build all of the static GUI screens and is used to interface the GUI with the application C code. The prototype also depends upon the DynaGraphX data widgets.

The prototype is being developed on Sun workstations. The target platform will be Silicon Graphics "Indy" workstations. SwRI plans to work with the NASA-MSFC's prime contractor to port the prototype to the SGI in early February.

Development Methodology

The prototype is being developed jointly by NASA-MSFC, SwRI, and NASA-MSFC's prime contractor. The division of responsibility is as follows:

- NASA-MSFC - Responsible for the static presentation of the GUI. NASA-MSFC uses TeleUse to build all "screens" and therefore define the basic "look-and-feel" of the application. This approach has the advantage in that NASA-MSFC personnel are best suited to determine the look-and-feel requirements of local users.
- SwRI - Responsible for the dynamic operation of the GUI. SwRI takes the static presentation developed by NASA-MSFC and develops the code necessary to bring the application to life. All code is being developed in ANSI C and with the Dialog ("D") language provided by TeleUse. D is an interpreted language which simplifies the user interface development process.
- NASA-MSFC Prime Contractor - Responsible for generation of the detailed design. NASA-MSFC, SwRI, and the prime contractor meet frequently to insure that the prototype code is well-synchronized with the design. The goal is to insure that once the prototype is complete, the code will become the baseline for the operational system.

NASA-MSFC has a strict user interface standard (MSFC-1956) which describes the basic static and dynamic operation of all applications in the Enhanced HOSC. As the prototype evolves, a Common User Interface Review Board (CUIRB) reviews the prototype and requests changes which insure that the prototype is consistent with other Enhanced HOSC applications.

To insure that the prototype will satisfy the subjective requirements of users, NASA-MSFC schedules frequent user evaluations. These user evaluations allow users "hands-on" use of the prototype, without the aid of NASA-MSFC or SwRI assistance. The evaluations are recorded and comments are filtered back into subsequent versions of the prototype. NASA-MSFC has completed two user evaluations, with six more scheduled prior to the final design review.

Features

Most of the features within the prototype consist of the editing and display capabilities required for this class of tool. The real measure of the prototype is the cohesive operation of editing features within Display Generation and how effectively these functions can be used to build displays.

Display Generation presents a main control screen with palettes for all objects and common attributes (colors, fonts, etc.) which can be applied to objects. Also present is a work area in which the user constructs the display. A rich set of editing functions is available from pull-down menus, pop-up menus, and keyboard short cuts. See Figure 2 for a sample Display Generation screen.

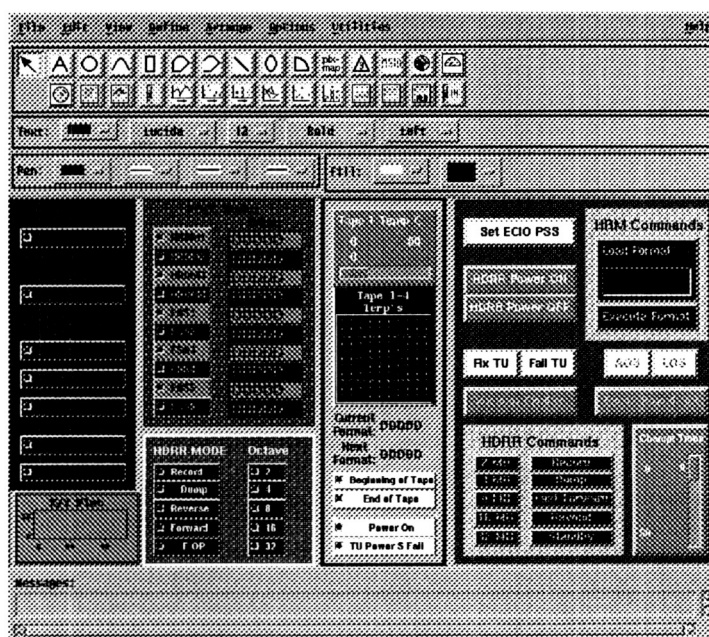


Figure 2 - Display Generation Sample Screen

Following is a partial list of the functions available within Display Generation. The functions shown in bold are not yet implemented.

- File Functions - New, Open, **Retrieve** (from shared storage), Save, Save As, **Store** (to shared storage), Print Window, **Print Display Information**, and Exit.
- Edit Functions - Move, Resize, Undo, Cut, Copy, Paste, Delete, Select, Select All, Deselect, and Reshape (for graphics).

- View Functions - Clear Messages, Control Message Window, and Refresh.
- Define Functions - Define Display Settings (Grid, Snap, Display Parameters, etc.), Edit Object Attributes, **Define User Pull-down Menus**, and **Define Start-up Comps**.
- Arrange Functions - Front, Back, Forward, Backward, Group, Ungroup, and Align.
- Options Functions - Fit Display, Show Drawing Tools Palette, Pixmap Editor, **Auto Generate Objects**, and **Validate**.

A typical user scenario would be to create a new display, set user preferences, use the palette to select and create new objects, use the editing functions to place and size the objects, use the object attribute functions to set object-specific parameters, and then save the display. Most of the user's time is spent in the work area using the editing functions and setting object-specific attributes to achieve the desired behavior for each object. Each type of object (a push button, simple rectangle, or time plot) has its own attribute screen, each containing a unique set of attributes.

Display Operation is a simple application which is primarily responsible for opening a pre-defined display and driving the display with data. Figure 3 shows an example Display Operation screen.

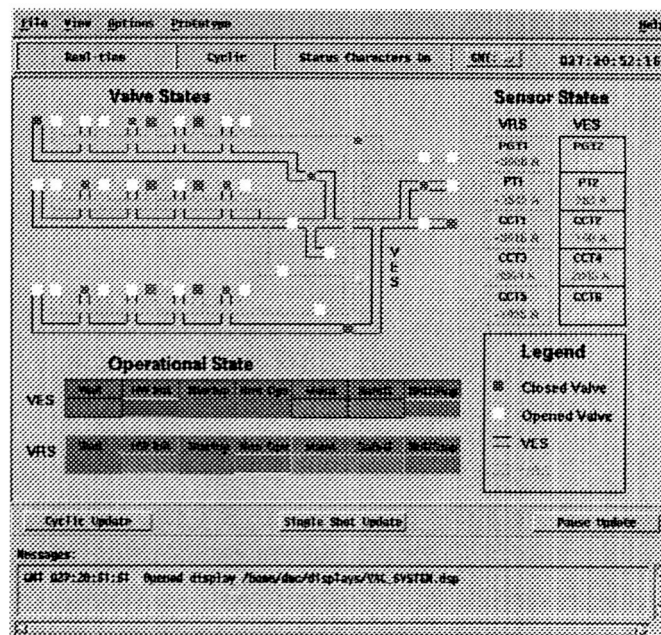


Figure 3 - Display Operation Sample Screen

Following is a partial list of the functions available within Display Operation. The functions shown in bold are not yet implemented.

- File Functions - Open, Print Window, and Exit.
- View Functions - Clear Messages, Control Message Window, Fit Display, and Refresh.
- Options Functions - Control Status Characters, Set Update Type, Set Update Rate, Set Data Source, **Validate**, and **Print Display Information**.

The prototype also provides a data simulator which drives objects with random, incrementing, or "pseudo-nominal" data. This simulator can be controlled by the user to best stimulate different types of objects.

Advantages Of Reuse

The Display Services prototype is now and will continue to evolve into, an excellent tool for building real-time data display applications. This is true within both NASA environments and other environments requiring display of data in various graphical formats. The prototype is designed to be extremely easy to use and is being verified through extensive user evaluations. The prototype is efficient and can effectively display large amounts of data without adversely loading a system.

Because the source code is available to NASA and any government organization, the prototype could be obtained and then modified to support another environment. For commercial organizations, the source code would be obtained directly from SwRI. Before the end of this year, SwRI will have developed a fully functional prototype, which would be one possible baseline for reuse. Another option would be to delay and obtain the final, documented source code and design from NASA-MSFC when the production product is complete. This product however, will not be complete for approximately 2 years and will be highly dependent on the specifics of the Enhanced HOSC.

Although the prototype is being developed on Sun Workstations, use of standard Application Programming Interfaces (APIs) such as POSIX, X Windows, Motif, and ANSI C will insure that the prototype will be portable to other environments supporting these standards. This will be demonstrated by the end of February 1994 by porting to a Silicon Graphics workstation.

Obtaining the prototype and modifying the source code for a particular environment will be a cost-effective solution for systems requiring full display building/operation functionality on many workstations. The prototype will be most effectively reused in environments characterized by the following:

- Flexible and/or dynamic display formats - a critical ability of the prototype is the ability to quickly build and modify display formats. The prototype will be most effectively used in environments where many different formats are required and the ability to quickly modify formats is necessary. The prototype is also useful in other environments characterized by limited or static displays, but the abilities of Display Generation will be less critical.
- User-level generation of displays - Display Generation is designed to empower users with the ability to build and modify displays. The prototype will most effectively be used in environments where there are many users and these users require the ability to build their own displays. This is an instance where a custom tool has an advantage over a COTS tool because each instance of a COTS display builder will carry a license fee.
- Flexible graphic data formats - the prototype is designed to allow new data display objects to be easily integrated. This is critical for environments requiring new data display formats, such as special plots or 3-dimensional data display objects.
- Performance - Display Operation is designed to be a real-time display system and is optimized for data manipulation and display of common object types.

As described, the prototype treats all widgets as generic objects within the application. This isolates the details of the widgets within a common functional area. To add new or replace existing widgets is a straight-forward process. Adding a new widget normally involves updating a few tables, adding a small amount of new code to support unique widget attributes, and then building a new

attribute screen for Display Generation. Display Operation may also need to be modified to properly support update of the object, however, this is not a difficult process.

The prototype also supports an easily modifiable design for adding new attributes for existing objects. To add new attributes for an existing object, an internal table for the object would be updated, the attribute screen would be modified, and perhaps a small amount of code would be required to support the attribute if it is new to the application.

Because the prototype utilizes TeleUse, the GUI is highly separated from the application code. The entire static user interface is represented in separate files which are not intermixed with the code. Therefore, profound changes can be made to the appearance of the GUI without impacting the operation of the code. The static user interface also makes heavy use of "templates", which are reusable class objects within the GUI. Components of the user interface which are Enhanced HOSC-specific are for the most part represented in templates. By changing a few templates, the dependence on "MSIDs" and other Enhanced HOSC-specific constructs can be eliminated and the modification replicated throughout the application.

The prototype is strictly compliant with the Motif Style Guide. One of the functions of the NASA-MSFC CUIRB is to enforce this compliance. This insures that the prototype will work well with other custom and COTS Motif applications.

Because the prototype goal is to concentrate on the user interface, most of the Enhanced HOSC-specific interfaces have not been addressed in detail. There are very few environment-specific functions which will have to be replaced in order to reuse the code in other environments.

Examples of Reuse in Other Environments

Some environments in which the prototype can be effectively reused include the following:

- Commercial control centers - large, industry-based control centers responsible for monitoring of commercial equipment (such as satellites).
- Process control - applications such as plant floor, oil/gas refinery operation, and conventional/nuclear power generation.
- Real-time data acquisition - graphic display of seismic and oil exploration data.
- Real-time simulation - display of real-time simulation data originating from the simulated component or system.

Conclusions

The Display Services prototype is a powerful, user-friendly environment designed for efficient generation and operation of real-time data-driven displays. Although developed for the NASA-MSFC Enhanced HOSC environment, there are no major technical problems in reusing the software as the basis for equivalent applications in other environments. This includes other NASA centers, other control centers, process control applications, and data acquisition systems. With a modest level of effort, this prototype and its future generations could be transitioned to other environments in a cost-effective manner.

2011

**Session G3: INTEGRATED VEHICLE
HEALTH MANAGEMENT**

Session Chair: Wayne McCandless

DUAL-USE ASPECTS OF SYSTEM HEALTH MANAGEMENT

by
P. R. Owens, B. J. Jambor, G. W. Eger, and W. A. Clark
Martin Marietta, POB 179, Mail Stop T330, Denver, CO 80201

514-38

ABSTRACT

System Health Management functionality is an essential part of any space launch system.¹ Health management functionality is an integral part of mission reliability, since it is needed to verify the reliability before the mission starts. Health Management is also a key factor in life cycle cost reduction and in increasing system availability. The degree of coverage needed by the system and the degree of coverage made available at a reasonable cost are critical parameters of a successful design.

These problems are not unique to the launch vehicle world. In particular, the Intelligent Vehicle Highway System, commercial aircraft systems, train systems, and many types of industrial production facilities require various degrees of system health management. In all of these applications, too, the designers must balance the benefits and costs of health management in order to optimize costs.

The importance of an integrated system is emphasized. That is, we present the case for considering health management as an integral part of system design, rather than functionality to be added on at the end of the design process. The importance of maintaining the system viewpoint is discussed in making hardware and software tradeoffs and in arriving at design decisions.

We describe an approach to determine the parameters to be monitored in any system health management application. This approach is based on Design of Experiments (DOE), prototyping, failure modes and effects analyses, cost modeling and discrete event simulation. The various computer-based tools that facilitate the approach are discussed. The approach described originally was used to develop a fault tolerant avionics architecture for launch vehicles that incorporated health management as an integral part of the system.

Finally, we discuss generalizing the technique to apply it to other domains. Several illustrations are presented.

INTRODUCTION

All complex systems use some kind of health management. Space launch systems are in many ways the epitome of complex systems. Most of this paper will deal with techniques we are using in evolving an avionics system for space launch vehicles that integrates health management functionality. As the discussion of the techniques progresses we will indicate other applications that could benefit from the techniques.

Current launch vehicle systems have added health management functionality on an ad hoc basis. The usual practice has been to add sensors to the vehicle during the testing phase. Common practice has been to leave those sensors in place as transition is

¹We define system health management as functionality that:

- Finds out system status and performance, captures that information, and delivers that information;
- Analyzes the status and performance information with respect to design requirements, captures that information and delivers that information; and
- Enables the system operators to use the information to prevent and correct problems or to improve the system.

made to the operations phase. This approach is not optimal. What is needed is a total systems approach. Such an approach would ask and answer two questions:

- **How much health management coverage is optimum?**
- **How should system health management be structured?**

The answers to these questions can only be derived through considering the system as an entity that includes both the launch vehicle and the ground system. Considering the launch vehicle alone will only yield a local optimum. Certainly, subsystems must be monitored on the launch vehicle. However, there are numerous ground subsystems whose malfunction can stop a launch or cause a launch to fail. Consequently, the whole system must be considered.

As one considers the whole system, it is clear that there are so many potential problems to monitor that to monitor them all is impractical. We could load the system with so many sensors that the vehicle would have terrible performance and the whole system would be too expensive to build and operate. Additionally, each additional sensor is a potential false alarm source and tends to reduce the system reliability and/or availability. Part of the problem is that a systematic way of choosing individual parameters and combinations of parameters to measure has been lacking. Historically, subsystem experts have given informed opinions about what parameters are important. There has been no way to optimize the parameter set. Consequently there has been no way to optimize the sensor suite.

This problem of deciding what parameters are important is not limited to launch vehicle systems. Examples of this problem can be found in semiconductor processing systems, other manufacturing systems, chemical production plants, power production and distribution systems, and transportation systems (including rail, ship, automotive traffic and aircraft). All of these systems are complex and have many parameters that might be important in controlling the system.

Currently, launch vehicle system health management information is a mixed bag of digital and analog signals that is taken from the umbilical or monitored in flight via telemetry. There is little integration to the health management functionality. Almost all the attention is on the vehicle. On current systems, health management functionality is difficult or impossible to use in the early activities of receipt-to-launch processes, an important fraction of the life cycle cost. An integrated systems approach is called for. The next section deals with such a systems approach to launch vehicles.

RESULTS OF THE INTEGRATED FAULT TOLERANT AVIONICS (IFTA) STUDY

During the period March 1990 through December 1993 Martin Marietta studied the feasibility of developing a fault tolerant avionics² system for expendable launch vehicles (ELV's). The study objective was to select a cost effective, high reliability, single fault tolerant avionics system architecture for ELV's.

The architecture selected had the following characteristics and components.

- It is a distributed avionics architecture with subsystems communicating via a network of nodes.
- Each subsystem forms a self-contained fault tolerant zone.
- It uses a pentad or hexad inertial measurement unit.
- Pairs of self checking processor pairs form the basis for processor fault tolerance.

²Avionics as used here means anything on the launch vehicle with an electrical connection.

- Nodes reside on Stage 2 and communicate with Stages 0 and 1 via Data Communication and Distribution Units (DCDU's)
- Point to point FDDI connections are used.
- Electrohydrostatic or electromechanical actuators are used.
- It has a centralized power system with distributed control.
- Rechargeable silver zinc batteries are used.
- Health management is an integral part of the architecture.

The advantages of the architecture are as follows.

- The system can accommodate multiple faults (one per fault containment region).
- The health management functionality supports ground operations.
- Redundancy management is transparent to flight software.
- Fault recovery is predominately hardware based.
- The system is easily upgraded by exchanging older technology boxes for boxes employing newer technology.
- The architecture is suitable for expendable launch vehicles, upper stages, and payloads.
- The system has reduced susceptibility to lightning, electromagnetic interference and radio frequency interference by isolating subsystems with optical cables.
- The architecture makes maximum use of standard hardware and software.

Two important points can be derived from the above account. First, a thorough systems approach can resolve requirements that have historically tended to clash. For instance, fault tolerance has frequently resulted in an avionics design that masked faults. Such designs have made health management difficult. However the IFTA architecture *facilitates* health management by not masking faults but by flagging them. Further, the IFTA architecture confines the faults to a specific subsystem and corrects them in real time. The fault detection mechanism is used by health monitoring functionality without interfering with fault recovery, thereby integrating the fault tolerant and health management functionalities. This integration of health management functionality with avionics can be extended into the entire vehicle. It is realized by distributing the classical notion of avionics into the vehicle structures and propulsion subsystems using advanced micro-controller technology in conjunction with advanced sensors. These 'smart sensors', when integrated with the main-line avionics, provide a built-in-test capability encompassing virtually the entire vehicle. This test capability enhances the early detection and correction of a faulty vehicle component or subsystem before it escalates into a catastrophic or costly failure.

Second, the design sought to not only make it easier to do health management, but also to use the health management functionality of the vehicle in streamlining the launch site (and conceivably the manufacturing) test and checkout processes. These have historically used independent equipment and software as opposed to using capability already built into the vehicle. For instance, in the proposed design, the network processors check themselves and signal when a processor pair has failed. The same health management functionality that checks the vehicle in flight and reports the failure can be used during the ground checkout of the avionics network.

The two points discussed above facilitate a major system benefit—providing for more autonomous and integrated operational test and checkout. This capability increases availability and decreases ambiguity of test results (allowing optimal use of support

personnel), resulting in a lower cost launch system with higher probability of mission success. It must be stressed that the on-board health management system must be integrated with the ground operations system to realize these benefits.

We suspect that many complex systems have been designed without a balanced consideration of the health management needs of the system for checking itself out as it is being assembled and integrated and as it is being operated and maintained.

HOW MUCH HEALTH MANAGEMENT COVERAGE IS OPTIMUM?

The approach to specifying the sensor suite for health management is shown in Figure 1. The diagram and the process description that follow are simplified and leave out many details. However, they serve to illustrate what is unique about the IFTA approach. Prototype hardware and simulations are used to firm up the design. Failure modes and effects criticality analyses (FMECA), and fault tree analyses are all used at the system and functionality level to analyze the design for problems. The system is designed to be fault tolerant to simplify and shorten FMECA cycles. So far this is standard practice.

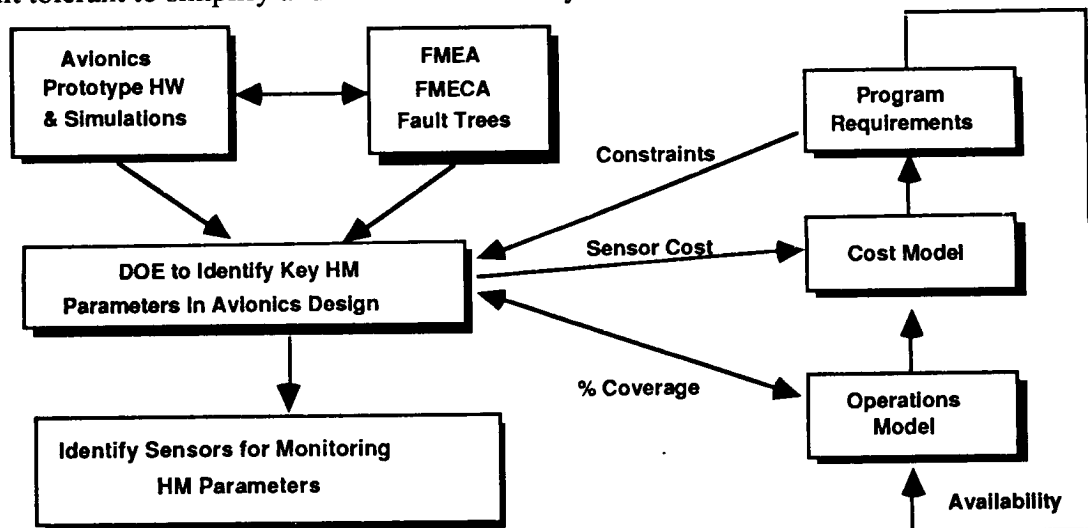


Figure 1. Avionics and Health Management Design

The departure from standard practice comes with using the method of Design of Experiments (DOE) to identify which system parameters give how much coverage. Using DOE entails starting with an expert designer's list of parameters that need to be measured. A test bed is instrumented to monitor the parameters. If part of the test bed must be simulated, a very accurate and rich simulation is required. Experiments are then run varying the parameters and collecting data from the instrumentation. The DOE technique tells how to vary the parameters and collect useful information with the minimum number of runs. The data are then analyzed. The analysis tells what the fault coverage will be and the confidence level of the coverage with variations of parameters.

This fault coverage information is useful in three ways. First, it can tell us if some parameters are not really needed and what combinations of parameters are important. Second, it can be used to help define an operations model that incorporates monitoring a given combination of parameters. Third, the fault coverage information can be used to determine a sensor suite whose associated costs are input to a cost model. Notice that the cost model allows for weighing the costs and benefits of various sensor suites as they allow various operations concepts. For instance, the additional cost of adding a particular

sensor to the suite may be more than overcome by the benefit derived from the simplification of the operations it enables. Alternatively, the benefit may be derived solely from the additional fault coverage guaranteed and the likelihood of avoiding a catastrophic failure.

The interplay of determining coverage, operations concept, and cost modeling can take place iteratively until a satisfactory solution is arrived at. This process is not automated. But there are computer-based tools available in the form of DOE structuring and analysis, cost modeling, and operations modeling tools.

The tools we have been using are as follows:

- RS/1 Series (from BBN Software Products Corp.) is used to do the DOE experimental setup and parameter variation and for analyzing the results of the experiments;
- MacProject II or MacProject Pro (from Claris) is used to do operations flow precedence networks;
- AutoMod (from AutoSimulations) is used to do discrete event simulations using the networks developed;
- SES Workbench from Scientific and Engineering Software, Inc. is used to simulate subsystems not available in the testbed; and
- The cost model is a Martin Marietta proprietary program that is being evolved using Independent R&D funding.

This approach is applicable to the design or modification of the health management functionality for any complex system. For existing systems, the testbed already exists—it is the system itself. Consider an aircraft transportation system. Such a system consists of the aircraft and its subsystems, the manufacturing facilities and people, and the operations and maintenance facilities and people. Start with the existing health management instrumentation and the data that has been generated by it. DOE techniques can be used to analyze the existing health management information to determine which parameters currently being measured are important. This assumes that all the truly important parameters are being measured or can be derived from the data. The "testbed" can be incrementally modified with sensors and experiments run to optimize the system costs in conjunction with cost and operations models. Likewise existing manufacturing operations can be used as testbeds.

For systems in the design phase, The technique can be used as described above. There will be physical testbeds built with modeling filling in parts that aren't actual hardware and software. The Intelligent Vehicle Highway System (IVHS) is a good example of such a system. Here there will be many parts of the system that can be tested separately first and then in conjunction with other parts. For instance, the system that governs "platooning" (grouping of cars nearly bumper to bumper on a high speed stretch of highway) could first be tested independently of an urban highway net control center. As the design progresses the platooning system (which occurs on the edge of the urban net) could be tested integrated with the net control system.

HOW SHOULD HEALTH MANAGEMENT CAPABILITY BE STRUCTURED?

In our Integrated Fault Tolerant Avionics approach, health management is treated as a system-wide problem that has an information management solution.

System Problem

The vehicle is not the only component of the system that requires health management. There are numerous pieces of support hardware and software that can cause launch delay, or hazards to people or equipment, or mission failure if they operate improperly and if this improper operation is not detected. Consequently, when one considers health management, one must consider the whole system.

Sensors and built-in-test on sub-assemblies can indicate a pass-fail result only. High coverage demands a proliferation of circuitry, which adds complexity but does not resolve existing ambiguities. At the functionality level, one can use the redundancy and self-checking properties of the fault tolerant design to let the system perform tests on itself instead of using external circuitry. An example of this is the use of the self-checking network node processors themselves to tell if they are healthy. The approach is to add visibility into the system by using the system itself to check itself.

This approach of treating the system as a whole implies two things. First, the approach to health management on the vehicle should be compatible with health management of the ground-based components of the system. Second, standard components (both hardware and software) will be useful in creating an affordable and workable structure for health management.

Vehicle and Ground Health Management Compatibility

For the health management ground segment to be compatible with the vehicle segment they must be able to talk. Being able to talk requires a common network and a common communication protocol. Compatibility does not necessarily imply that the ground segment has the same degree of fault tolerance as the vehicle segment. Consequently, the ground segment architecture will be different. But the connectivity and ability to transfer information must be there.

Standard Components

Use of standard hardware and software components simplifies the task of designing, building, and testing the health management functionality. Especially for the ground component, some equipment will be bought as commercial products with whatever degree of health management exists in the product. However, it will be possible to take health management information from such equipment, translate it, if necessary, and integrate it into the system health management functionality. The translator interfaces will be custom modules of software (and possibly hardware). However, the rest of the system can use a relatively small set of hardware and software building blocks. In addition, the translators are minimized if standard layered protocols are used. This will have the advantage of reducing design, programming, maintenance, and logistics time. At the same time, use of standard hardware building blocks will decrease the cost per component of a given building block.

Information Management Solution

The key to the solution is a combination of two things. First, the raw data is transformed into meaningful information, directly usable by the proper management functionality as close to the source as possible. Second, the information is stored in a distributed set of Management Information Bases (MIB's) accessible by the on-board network as well as the ground communication network.

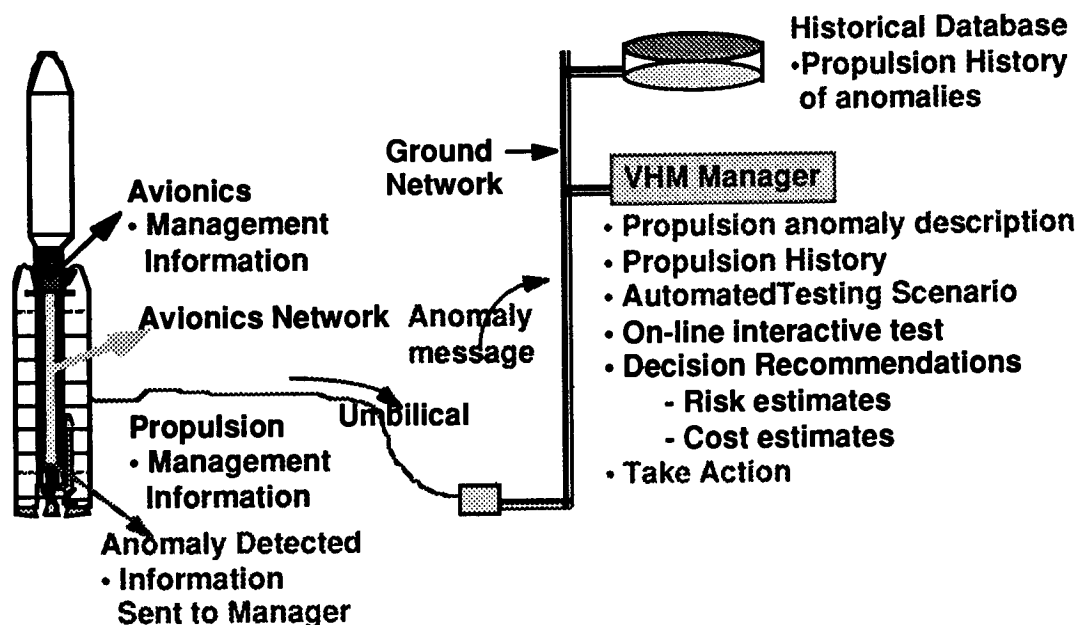


Figure 2. Vehicle Health Management Interactions with the Ground Segment

The on-board and ground segments of vehicle health management are structured so that action is taken at the lowest possible level. For instance, the vehicle network node processors are pairs of self-checking pairs. If one pair detects a fault in itself, it takes itself off-line and the other pair picks up the processing. More complex actions may be directed by the (ground-based) VHM Manager. It periodically polls the each of the vehicle MIB's for high level status. The MIB's each monitor a fault containment region. The VHM Manager can run analyses on the high level status information. If the analysis suggests an impending problem, the VHM Manager can query one or more MIB's for more detailed information. In response to this information the VHM Manager can act. Additionally the MIB's can be programmed to send interrupt messages to the VHM Manager if a designated condition is detected. Again, the VHM Manager can act or request additional information from the MIB's.

The VHM Manager's location (off the vehicle) has two advantages. First, its software does not have to be flight qualified. Second, an off-vehicle location is appropriate for using the Manager's capability as an integral part of vehicle assembly, integration, test, and checkout. Next, we turn to the health management of the ground segment of the launch system.

The health management of the ground part of the system can be represented as shown in Figure 3.³ There is a hierarchical layering of networks. At the highest level the facilities and the vehicle are all linked together. It is at this level that the VHM Manager resides. As one goes inside a given facility there is an organization of subsystems specific to each facility. An example is the Vehicle Integration Facility where the launch vehicle is integrated with the Mobile Launch Platform and the

In the case of the Vehicle Integration Facility, the subsystem network organization is shown in the bottom half of Figure 3. Note that each of the subsystems has a health management component. The database used for facility-specific health management information is maintained in the Data and Information Subsystem. This information is

³This is a generic facility set derived from the National Launch System. Facilities may be added or deleted.

accessible from other facilities (Launch Control, for instance). The Net Management and Services subsystem manages requests for information and transfer of messages and information into and out of the facility. The Data and Information Subsystem is also the home of the executive health management functionality for the facility as a whole. Each subsystem has some monitoring and reporting functionality and may have emergency safing capability. However the Data and Information Subsystem maintains the "big picture" for the facility as a whole. Actions concerning more than one subsystem or coordinating inputs from more than one subsystem will be handled at the facility level by the Data and Information Subsystem. Likewise, the Launch Control facility will maintain the "big picture" for the launch complex as a whole. The Data and Information Subsystem within the Launch Control Facility coordinates all the health management functionality for all of the launch complex. However, if the Data and Information Subsystem in the Launch Control facility is disabled, the Data and Information Subsystem in one of the other facilities can assume the executive role.

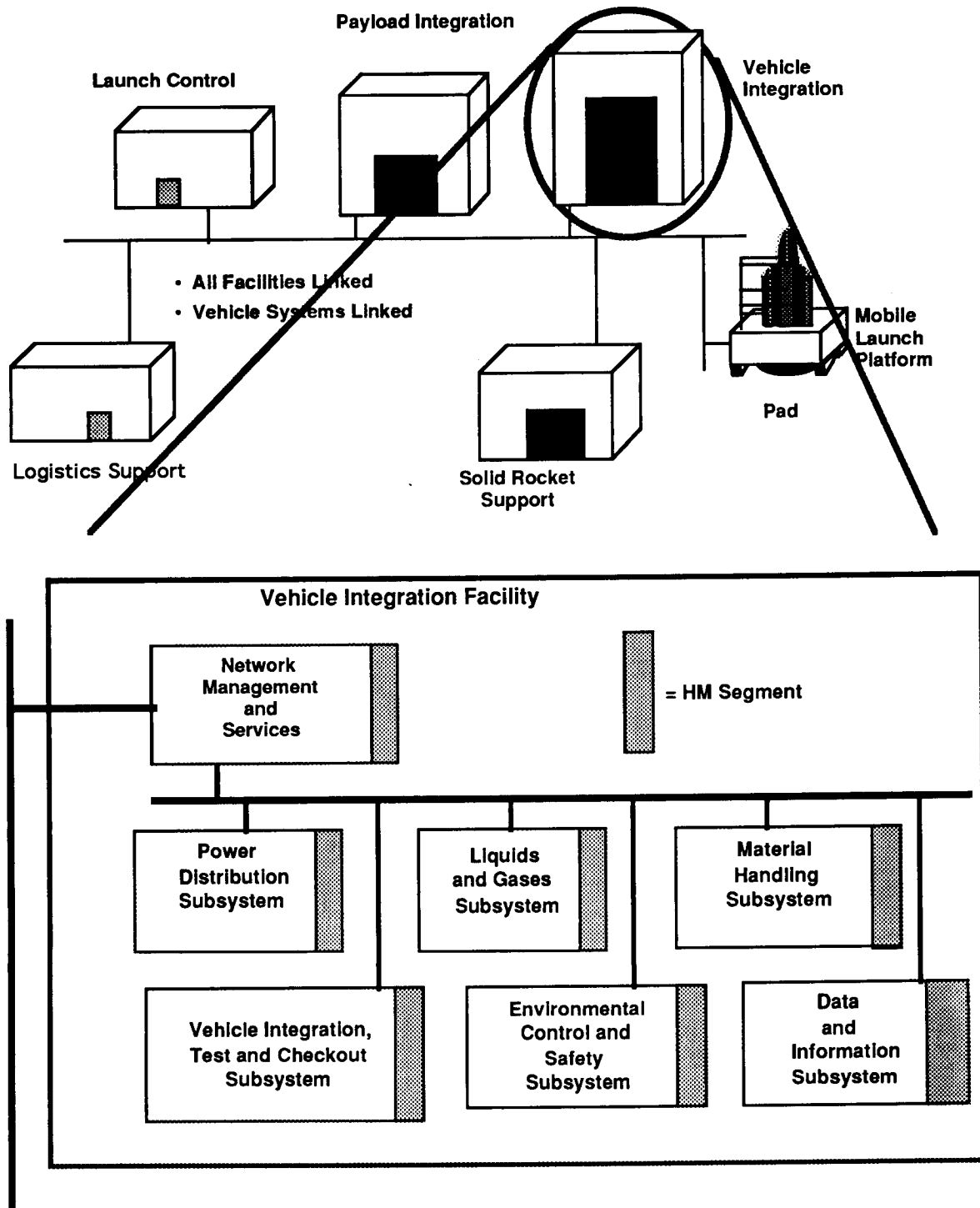


Figure 3. Ground Segment Health Management Architecture

The architecture described above is general enough to be applied to other applications than ELV ground segment health management. It can be equally well applied to commercial aircraft ground facilities, manufacturing plants such as petrochemical or semiconductor production, power production and distribution systems, air traffic control networks, nuclear waste cleanup systems, and rail transportation systems.

To explore the application of the architecture in slightly greater depth, consider the case of power production and distribution. There are one or more generating facilities and there are substations in the distribution network. Operators or maintenance people at the substations need to be able to tell locally what the status is of each subsystem in the substation. So there needs to be a local part of the health management capability to handle local problems. However, the system manager needs to be able to get system status. When a piece of the system goes down the manager should have high level information available at once. The system needs to act locally to safe the failed part of the system. The system manager will then want more detailed information about what failed and what the options are for working around the problem until the failed equipment can be repaired or replaced. The maintenance person at the substation will want detailed information about the equipment being repaired or replaced. The architecture outlined above can easily be adapted to accomplish these tasks.

Protocol for Health Management

Part of the solution of the problems of managing distributed information resides in using an appropriate network management protocol. We have adopted the Simple Network Management Protocol (SNMP) as particularly well suited to this task. The first advantage is that the Protocol is a standard that is used in many places, including such distributed systems as the Internet. The second advantage is that much of the industry standard computer hardware available supports this software standard. The third advantage is that SNMP uses Abstract Syntax Notation One (ASN.1) an International Standards Organization (ISO) standard notation that is very flexible and is particularly useful for expressing health management information. The organization of data is extensible. This feature is particularly applicable to systems that must be reconfigurable and to those that are changed over time. SNMP supports automatic configuration.

Generalizing from Space Launch System Experience

We have been discussing a particular set of approaches for doing system health management for space launch systems. Along the way some examples have been given to indicate other application areas. The next section further discusses the generalization of these approaches to doing health management for other systems.

ADDITIONAL DUAL-USE NOTES

Clearly this approach can be directly applied to any system that has (a) control and sensing elements distributed geographically and (b) needs fault tolerance. Some examples come to mind.

Example Systems

We list in this section classes of systems that should benefit from the approach described above to integrate fault tolerance with health management.

Power Generation and Distribution Systems

Clearly, electric power generation and distribution systems are both distributed systems and need fault tolerance. The health management for such systems may measure generator, transformer, voltage, current, and other parameters. The health management must be so set up that it works in concert with the fault tolerant aspects of the system. This is another case where the fault tolerance must not mask the health management pinpointing of failed components. Power companies want to deliver uninterrupted service **and** find out quickly and accurately what equipment needs repair or replacement. While there are no pieces of the system that are jettisoned in this case, it is important to be able to isolate pieces from the distribution network; and multiple levels of authority and recovery must be possible.

Manufacturing Process Management

This approach to health management can be used for designing, deploying, and improving any number of continuous flow and batch flow processes. There are many cases now where sensors are used to measure either direct or indirect process parameters and this information is used to control the process and where intelligence is distributed rather than centralized. Further, distributed control and monitoring will become the rule rather than the exception in the future. This approach, in particular, will apply to

- petrochemical plants,
- pharmaceutical plants
- semiconductor processing plants,
- steel and aluminum mills, and
- factories where products are assembled or fabricated with some degree of automation.

The consequence is that health management will be an integral part of the control problem for both economic and human safety reasons.

Intelligent Vehicle Highway System (IVHS)

There are several parts of the IVHS that require fault tolerant systems that also have health management capabilities. Traffic control systems in sectors that have platooned vehicle movement certainly need both fault tolerance and need rapid and accurate fault isolation. Sectors that control traffic flow in dense traffic areas will almost certainly use distributed computer resources networked together. These will need both fault tolerance and fault detection and isolation. In fact, these characteristics are true of mass transportation systems of all kinds. Moving humans and/or cargo safely and reliably provides opportunities for employing fault tolerance and health management in an integrated way.

Communications Systems

All kinds of communications systems are candidates for the benefits of integrated fault tolerance and health management. Since such systems are by nature distributed in their computing resources this will be a natural fit.

An Approach, not a Prescription

What we are proposing here is an approach to combining health management and fault tolerance. It is not a one size fits all design. We outline below the elements of the approach.

Take a Systems Viewpoint

Look at all parts of the system and at all parts of the life cycle. Ask whether and where health management can help in design, manufacturing, operations and maintenance. Is fault tolerance also needed in the system concept? If both are required, then an approach must be selected in which fault tolerance and health management play mutually supporting rather than antagonistic roles.

Develop a Concept for Health Management

Assuming health management will be a necessary or enhancing part of the system, determine what elements are necessary. This must be done in conjunction with the manufacturing, operations and maintenance concepts and must be modeled as an integral part with the manufacturing and operations models. Will built-in-test be adequate. Is health monitoring all that's needed or is health management (which implies some sort of capability for corrective action) necessary? Determine if this concept fits with the rest of the design at each stage of the iterative design process

Determine Cost Effective Coverage

Find out what is cost effective. This may be done by using fairly detailed modeling, a hardware and software testbed or a combination of the two. The sensor suite can be determined by Design of Experiments, modeling (if enough is known for a high fidelity model that will enable decisions to be made on how many sensors are needed, what types, and what the benefit is. Or parts of the system may be modeled and experiments run on the rest in a testbed.

Use Standard Hardware and Software Where Possible

Any time the system can use commercially available hardware and software it will save cost over doing custom hardware and software. Likewise, using multiple copies of hardware or software building blocks in the system will tend to reduce costs over designs that use different building blocks for doing similar functions.

Unify the Way Distributed Elements Communicate

In our case SNMP was chosen, but there are probably other good choices available for particular situations. In any case, there must be a consistent way for parts of the system to talk to other parts of the system.

Wring It Out

Prototyping and testing under stress conditions is necessary to uncover what has been overlooked. Design so that the prototyping and testing processes can be conducted rapidly and corrections can be made relatively easily. Here Design of Experiments techniques can be very valuable.

MIT

Dual-Use Aspects of System Health Management

P. R. Owens
B. J. Jambor
G. W. Eger
W. A. Clark

3 February 1994

MARTIN MARIETTA

Overview

- Introduction
- Study Results
- Optimum Health Management Coverage
- Structuring Health Management
- Additional Dual-Use Notes

Introduction

System Health Management – Functionality that

- Finds out system status and performance, captures and delivers that information**
- Analyzes status and performance with respect to design requirements, captures and delivers that information**
- Enables system operators to use the information to prevent and correct problems or to improve the system**

Introduction (cont.)

Complex systems need health management

- Space Launch**
- Other Transportation**
- Utilities**
- Manufacturing**

Total system approach needed

- How much is optimum?**
- How to structure?**

Integrated Fault Tolerant Avionics Study

Architecture characteristics and components

- Health Management an integral part**
- Distributed**
- Each subsystem is a fault containment zone**
- Pentad or Hexad inertial measurement unit**
- Pairs of self-checking pairs for node processing**
- Nodes are on top stage**
- Other stages communicate via Data Communication and Distribution Units**

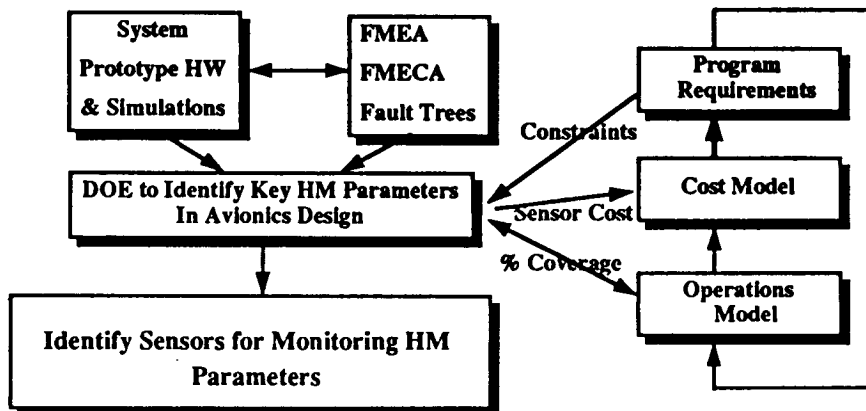
Integrated Fault Tolerant Avionics Study

Advantages of the architecture

- System accommodates multiple faults**
- Health Management supports ground operations**
- Redundancy management is transparent to flight software**
- Fault Recovery mostly hardware based**
- Easily upgraded – box exchange**
- Maximizes use of standard hardware and software**

Optimizing Health Management Coverage

System and Health Management Design



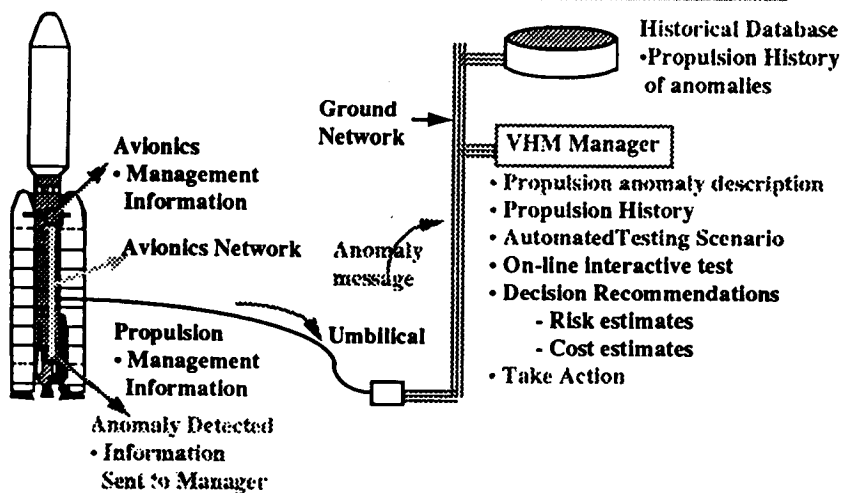
Tools

- **RS/1 Series from BBN**
 - Design of Experiments (DOE)
- **SES Workbench from SES**
 - Discrete event simulation with DOE
- **MacProject Pro from Claris**
 - Operations flow precedence networks
- **AutoMod from AutoSimulations**
 - Discrete event simulation for operations flows
- **Cost Model (Martin Marietta proprietary)**
 - Flexible launch system cost modeling

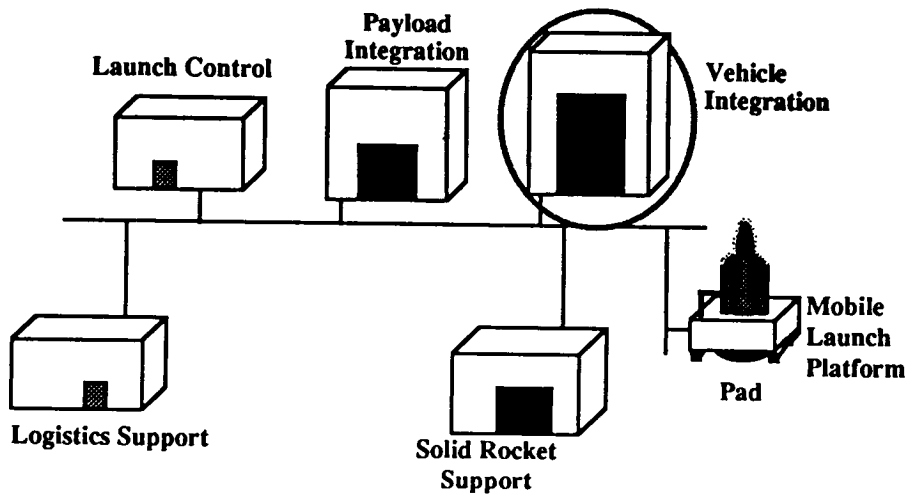
Structuring Health Management

- *System problem*
 - Vehicle and Ground
 - Standard Components
- *Information Management solution*

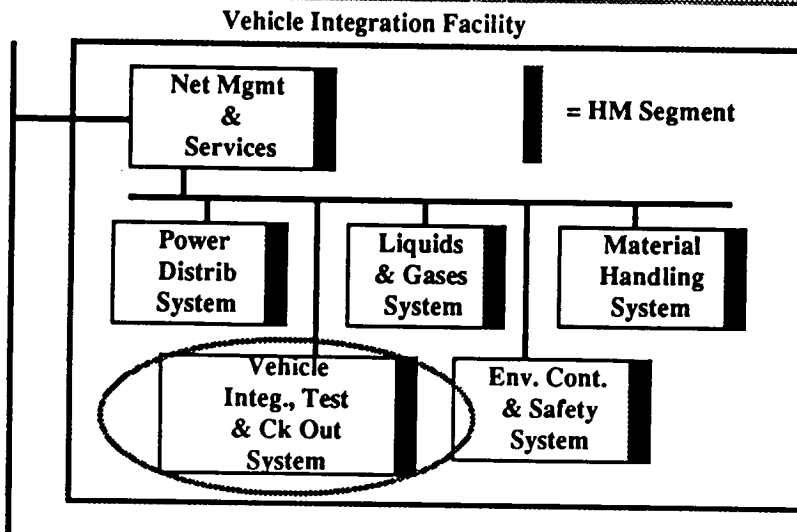
Health Management Interactions



Health Management Information System

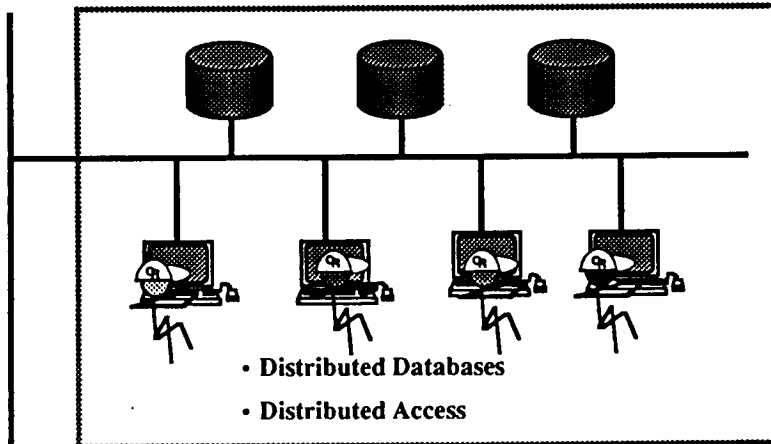


HM Information System



HM Information System

Vehicle Integration, Test & Checkout System



Protocol for Health Management

Simple Network Management Protocol (SNMP)

- Designed for network management
- Uses Abstract Syntax Notation One (ASN.1)
 - > Organization of data is extensible
 - > ISO Standard

Application Areas

- **Power Generation and Distribution**
- **Manufacturing Process Health Management**
 - **Petrochemical**
 - **Pharmaceutical**
 - **Semiconductor**
 - **Steel and Aluminum**
 - **Factories with some degree of automation**
- **Intelligent Vehicle Highway System**
- **Communications systems**

An Approach

- **Take a systems viewpoint**
- **Develop a Health Management concept**
- **Determine cost effective coverage**
- **Standard hardware and software**
- **Unify communications**
- **Wring it out**

REACTION JET DRIVE ELECTRONICS

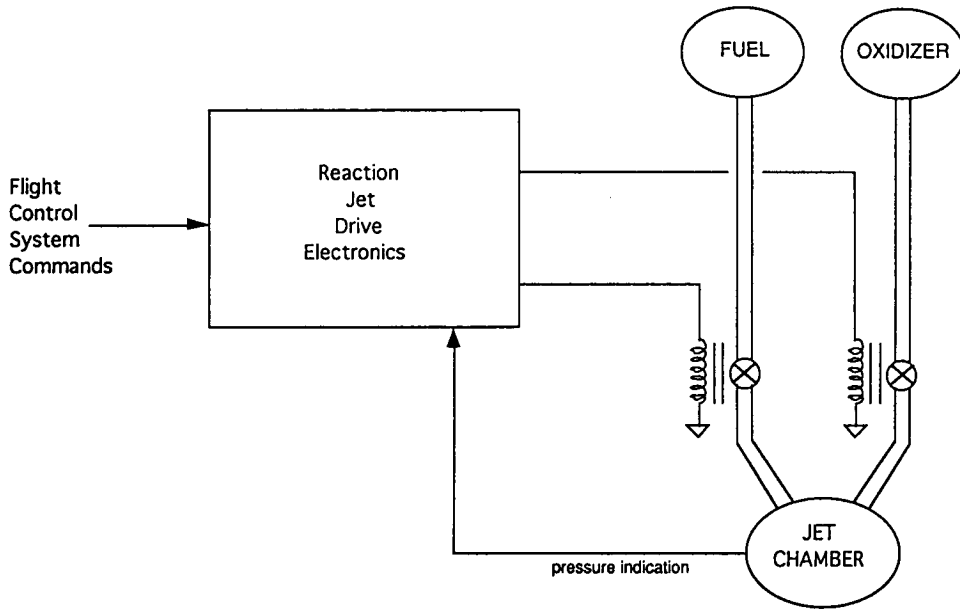
Dual Use Conference 3 Feb 94

Reaction Jet Drive Electronics Project Goals

- Compatible with current Orbiter subsystems
- Incorporate/Demonstrate Integrated Vehicle Health Management (IVHM) concepts
- Modular design, two channels per card, multiple cards per chassis
 - Capable of supporting functional growth
- Conceptual design by end of June, 1993 - Lockheed Sanders funded prototype functional before end of year

Dual Use Conference 3 Feb 94

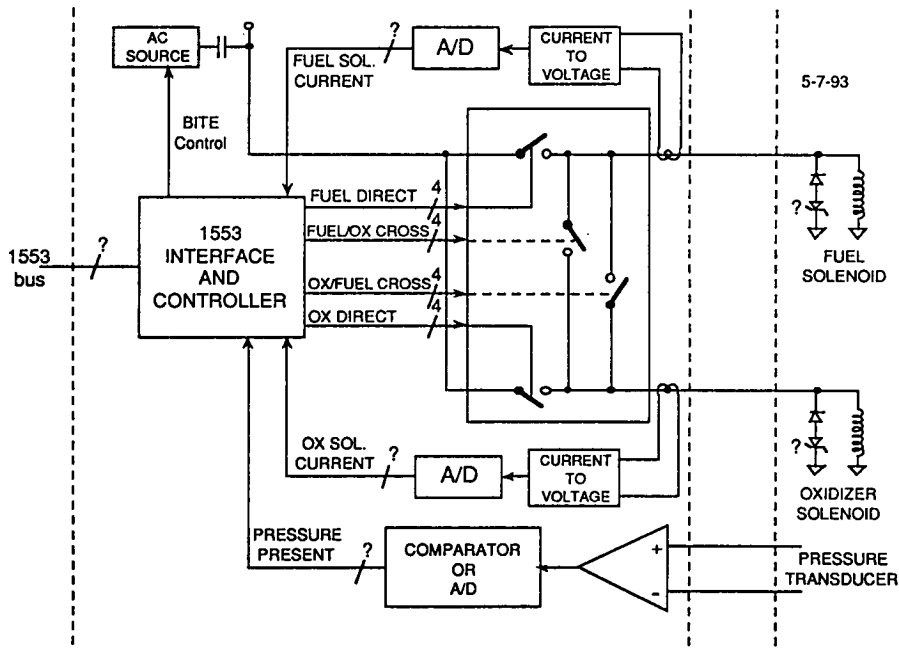
Reaction Jet Drive Electronics



Reaction Jet Driver System

Dual Use Conference 3 Feb 94

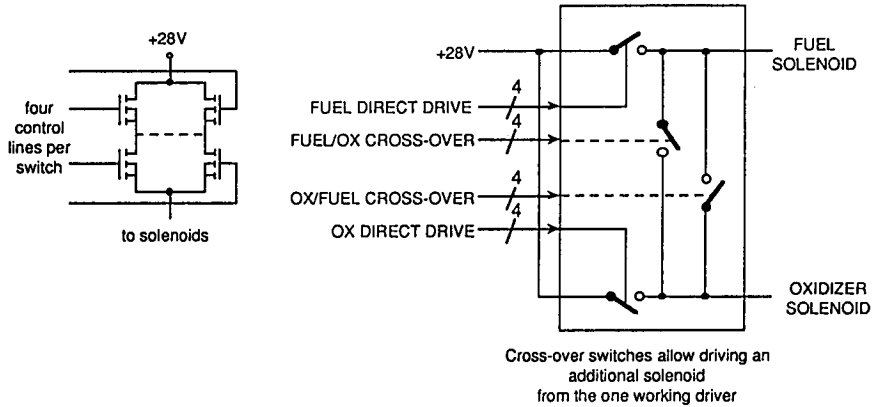
The Block Diagram of the conceptual design



Dual Use Conference 3 Feb 94

The "Smart" Switch Structure

5-7-93

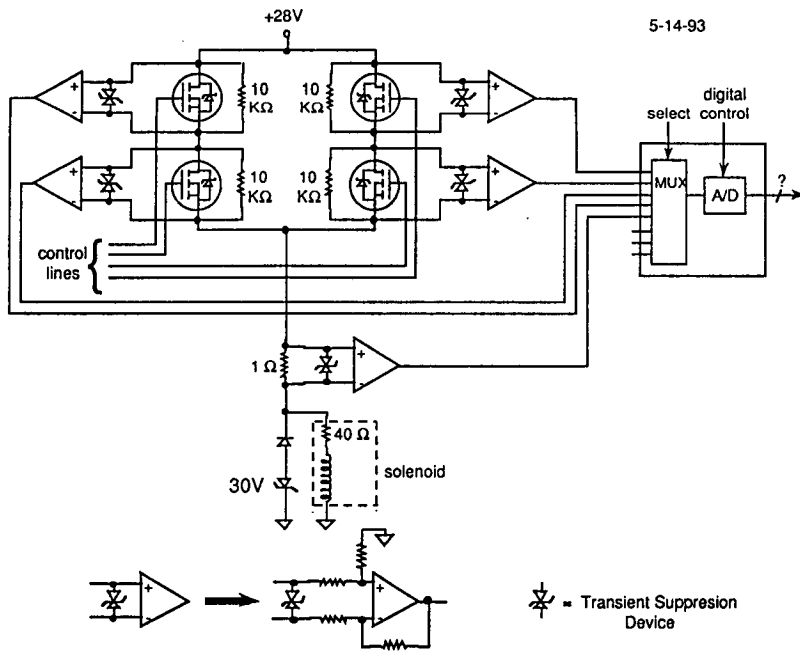


RJD Driver Circuitry

Dual Use Conference 3 Feb 94

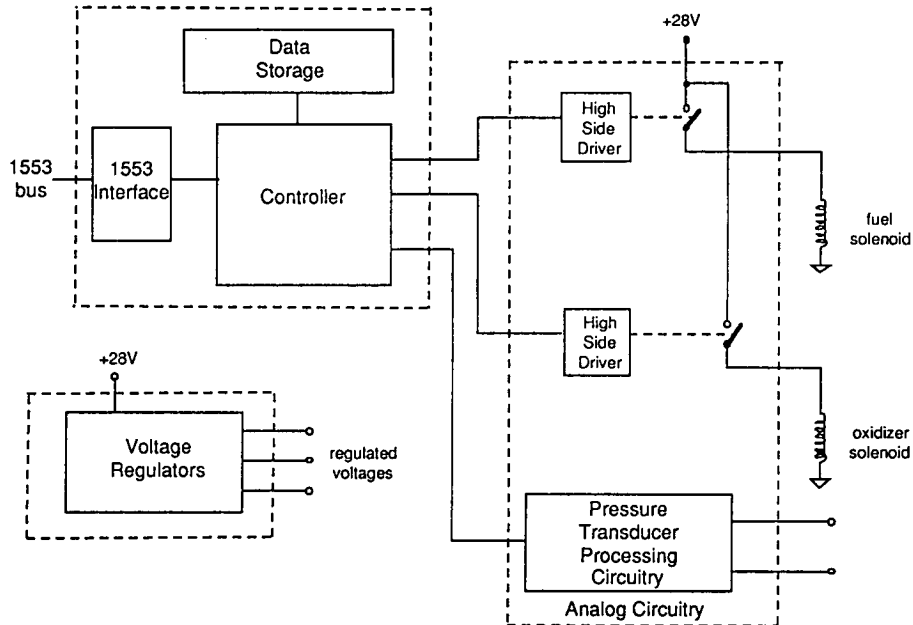
The "Smart" Switch

5-14-93



Dual Use Conference 3 Feb 94

The Digital Circuitry



RJD Hardware Components

Dual Use Conference 3 Feb 94

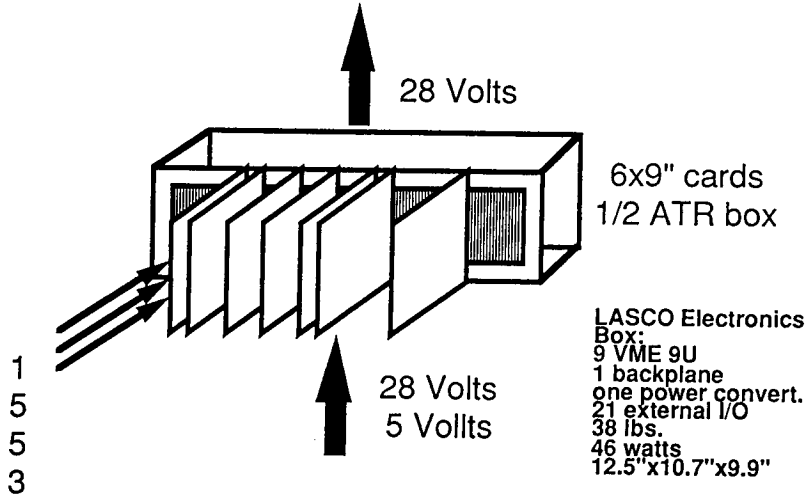
REACTION JET DRIVER (RJD)

Our View of the Future

- **A Digital Interface between the Flight Control System and other Avionics subsystems**
 - Reduce the number of unique interfaces which need verification
 - Reduce the amount of knowledge required in the FCS computers
- **Avionics subsystems should be composed of modular components based on the standard interfaces and definitions of Open Architectures**
 - Simplify the operational and logistics requirements of maintenance
- **Subsystem should be the repository of knowledge about the subsystem**
 - Reduce the requirement that the FCS computers must know the unique characteristics of each subsystem.

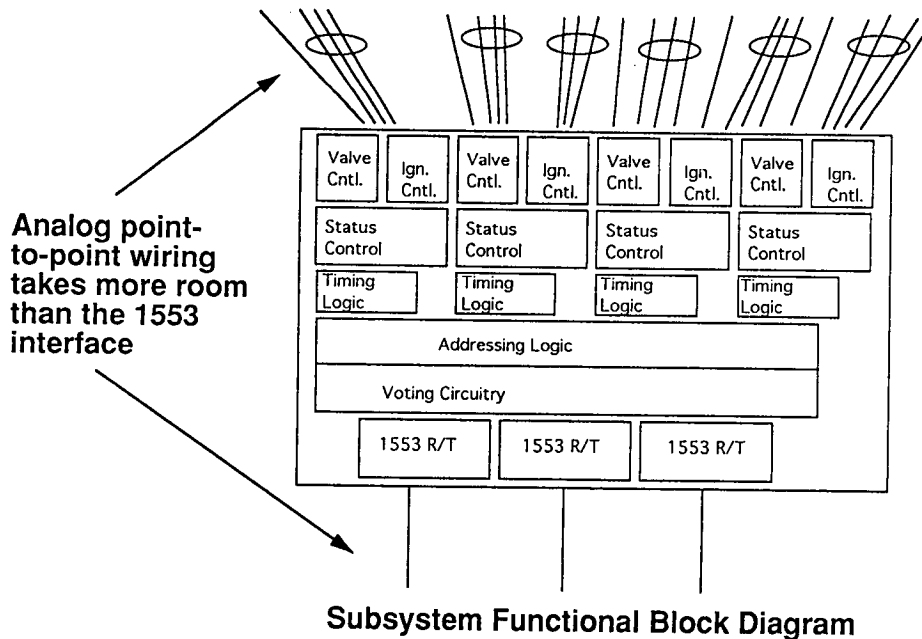
Dual Use Conference 3 Feb 94

Initial Notional Concept of RCS controller realization



Dual Use Conference 3 Feb 94

REACTION JET DRIVER (RJD)



Dual Use Conference 3 Feb 94

REACTION JET DRIVER (RJD)

Results

- Built prototype in less than eight months (concept to hardware)
- Prototype is Fail-off, fault(s)-tolerant
- Prototype is capable of Fault Detection, Isolation & Recovery
- Prototype is modular, utilizing standard interfaces of Open Architecture
- Current activity underway to:
 - Integrate the prototype into the JSC JAEL
 - Further 'productization' of the RJD controller
 - Characterization of the solenoid end-of-useful-life behavior

Dual Use Conference 3 Feb 94

REACTION JET DRIVER (RJD)

Summary: RJD Features and Benefits

- Independent control of fuel/ox valves (F, P)
- S/W selectable minimum fire duration (P, F)
- Accumulation of total jet fire time data (H & S)
- Solenoid residual life estimation (H & S, R)
- Dynamic jet priority adjustment (F, R)
- MIL-SPEC-1553 interface (F, C)
- Elimination of single point fail-on mode (R)
- Improved back EMF suppression (P, R, F, C)
- Modular construction (F, C)
- Increased internal redundancy (R)
- Multivehicle application (F, C)
- Growth capacity (F, C)
- Supports distributed optimal control (P, R, F, C)

Enhancements: H & S = Health and Status Monitoring, P = Performance,
R = Reliability, F = Flexibility/Adaptability, C = Cost

Dual Use Conference 3 Feb 94

A Multichannel Data Acquisition Module with State-Based Feature Recognition for System Health Management

Thomas G. Edwards

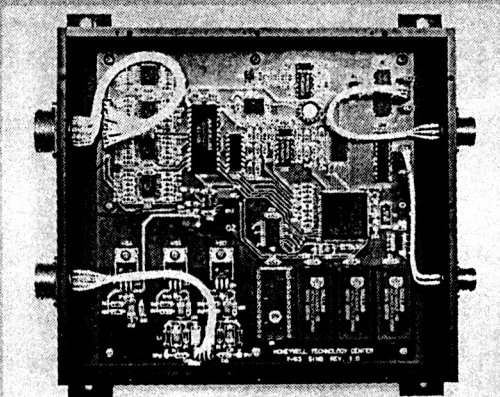
Honeywell Technology Center
10701 Lyndale Avenue South
Bloomington, Minnesota 55420

Honeywell

Honeywell Technology Center

B940076-03

Sensor Interface Module System (SIMS)



*SIMS—Developed for Martin Marietta as Part
of Integrated Fault-Tolerant Avionics Program*

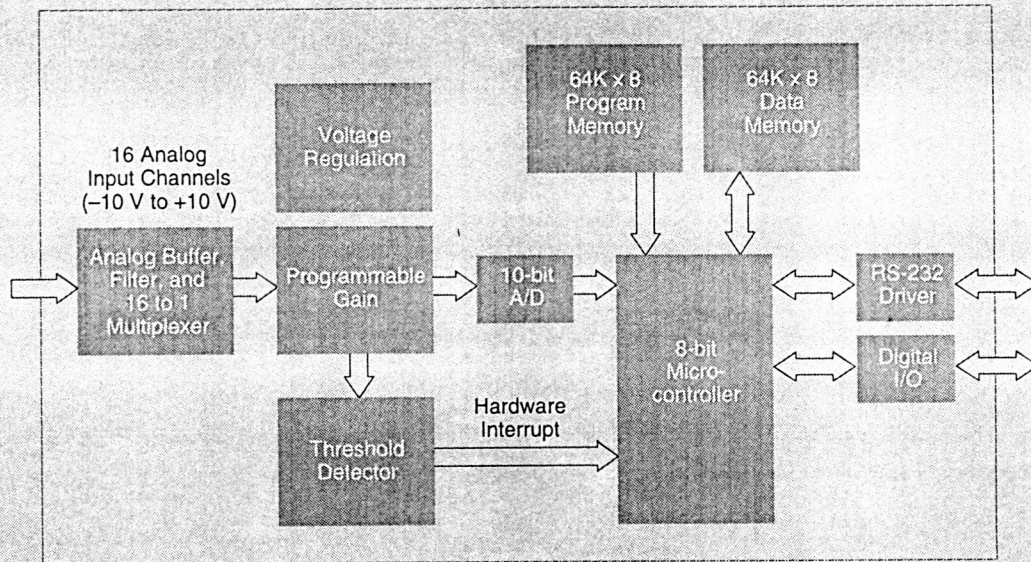
- Microcontroller based
- 16 channels configurable data acquisition
- Digital I/O
- Analog output
- Hardware interrupts
- RS-232 communications
- Built-in Test (BIT)
- Data compression
- Time stamp
- Custom State-Based Feature Recognition (SBFR) algorithms

Honeywell

Honeywell Technology Center

B940076-05

SIMS Block Diagram

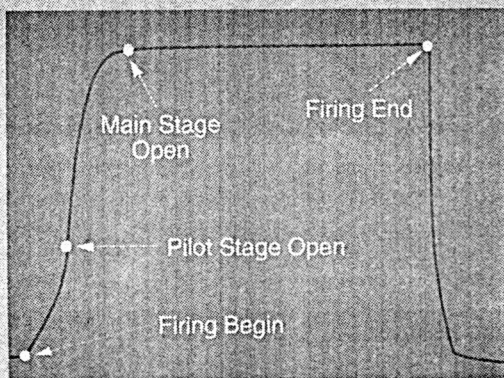


Honeywell Technology Center

Honeywell

B040078-06

State-Based Feature Recognition (SBFR)



SBFR algorithms extracting features from Hall-effect sensors, which sense position of pilot-operated valves of the Space Shuttle's Primary Reaction Control System (PRCS)

- State-based algorithms provide trend detection, pattern recognition, and feature extraction
- User-generated algorithms for custom applications
- Algorithms generated in Sparc-station-based State Machine Development Environment (SMDE) and State Machine Development Language (SMDL)
- SBFR algorithms extract multiple features from a single data channel or a single feature from multiple inputs

Honeywell Technology Center

Honeywell

B040078-07

SIMS Multitool Approach to System Health Management

Data Acquisition	Data Interpretation
<ul style="list-style-type: none">• Reconfigurability for application-specific requirements• No. of channels, gain, sampling rate• Data logger to record environment histories• Data compression• Data filtering• Intelligent I/O	<ul style="list-style-type: none">• Custom user-generated SBFR algorithms• SBFR algorithms interact with acquisition and I/O• Detect trends and critical events from a single or combination of all data paths• Alert host or reconfigure SIMS when events are detected

SIMS Supporting Sensor Information Flow in Distributed Avionics

- Provides a window into component and subsystem performance and health status
- Distribute intelligence to the lowest level of data collection, extending the Built-in Test (BIT) directly into the monitored subsystem
- Support distributed avionics system architectures, providing distributed intelligence at the data collection site; communicate only necessary information

Smart Sensor Heritage



Environmental Stress Monitoring Device for Histogramming Commercial Avionic Cabinets

- Original smart sensor developed under Air Force programs for monitoring avionic environments (Micro-TSMD)
- Smart sensor technology, developed for DoD applications, is being used in commercial applications
 - Space/military avionics systems
 - Commercial aircraft
 - Rotary-wing aircraft
 - Industrial vibration monitoring

SIMS Commercial Applications

- Industrial maintenance
- Failure prediction
- Warranty verification
- Shipment monitoring
- Hazardous waste monitoring
- Long-term environmental histograms
- Aging aircraft
- Failure analysis
- Intrusion detection
- Product improvement
- Process trend monitoring
- Equipment usage monitoring
- Vehicle Health Management (VHM)

Commercial System Requirements

- Provide complete, multiparameter operating environments
 - Single event
 - Trend
 - Histogram
 - Stress correlation
- Reduce maintenance cost—predictive maintenance
- Provide alert if hazardous environment is present
- Reduce design costs by verifying operating environment
- Small size, low power, in situ monitoring

SIMS Future

End users helping to define operating requirements for next-generation SIMS

- 32-bit microcontroller/DSP
- High-speed, 12-bit acquisition
- Open system communication protocols
- SBFR
- Packaging
 - MCM/HDI
 - Application-specific format (ASCM)
 - Stand-alone module

Conclusions

The SIMS, by executing a user-configurable operating/acquisition system and hosting user-defined SBFR algorithms, is capable of meeting the evolving needs of DoD and commercial system health management applications.

cmst

Session G4: ADVANCED AVIONICS

Session Chair: John Ruppert

AUTOMATED LAUNCH OPERATIONS

January 1994

ABSTRACT

Future cost-competitive, Government, and commercial launch systems will depend heavily on highly automated operations. Intelligently managed payload and vehicle build, assembly, test checkout, and integration processes followed by fully automated launch operations will be the key to lowering launch costs up to an order of magnitude below those typical of today. Honeywell is an aerospace industry leader in vehicle health management system technology and is unique in the application of industrial automation and control systems to automated launch systems. Use of this commercial off-the-shelf (COTS) automation technology - in build, assembly, and factory checkout, as well as at the launch site - provides a total launch vehicle system data base that effectively supports on-pad vehicle health management, system diagnostics and maintenance, and highly automated launch operations. Leveraging these modern automation and control approaches is vital to providing more cost-competitive, future U.S. launch services.

AUTOMATED LAUNCH OPERATIONS

Dr. Roger F. Block and Larry Brown
Honeywell Space Systems

INTRODUCTION

Current U.S. launch systems consist of the Space Shuttle and essentially three major expendable launch vehicle (ELV) systems. These vehicles, with their launch infrastructures and operations, are based on diverse 1960s/70s technology, use manually intensive procedures and are very costly to monitor, maintain, and repair. Annual U.S. launch activity costs are currently over \$7 billion. ELVs (Delta, Atlas, and Titan) are essentially converted 1960s ballistic missile systems requiring large on-site launch crews of 300 to 1000 people and 25 to 100 days at the pad¹. By comparison, the modern European Ariane 4 uses a launch crew of 100 and an average on-pad time of 10 days. The associated launch base and range operations for Cape Canaveral or NASA KSC employ 10,000 to 18,000 people, while Kourou Space Center (Guyana) uses only 900.

These metrics graphically explain why U.S. launch systems are more costly to operate (~ \$4500 to \$6000/lb-LEO) and why the U.S. share of worldwide commercial launch business has plummeted from 100 percent in 1972 to less than 30 percent today! Ariane dominates the commercial launch market with a 60-percent share. Russia/China have 10 percent.

Over the last 10 years, selected upgrades have been incorporated into the U.S. ELVs and Shuttle to improve mission success and safety, but these efforts have had little effect on reducing the operations costs. The improvements were focussed on isolated, vehicle based elements that did not attack the "root issue"; out-of-date, manually intensive operations concept and launch support infrastructure.

As long as the U.S. depends on outdated launch systems, our space-launch systems business will continue its decline, even with investments in sporadic and isolated technology enhancements for each vehicle. A new paradigm shift in operations culture and new way of doing business is required to achieve at least a 50 percent reduction in

launch costs. Recent analyses indicate an order of magnitude in cost reductions is possible using single-stage-to-orbit (SSTO) operational vehicles and new automated operations.

Future, Cost-Competitive Launch Operations

To significantly change this disturbing trend, the U.S. must make a firm commitment to invest in a modern space launch system, a new launch infrastructure, and a drastically different way of operating the new system. This decision must be made *immediately* for a new system to be developed and operational just after the turn of the century. After all, real cost reductions are possible. Ariane IV and V demonstrate and validate the impact of modern technology and launch practices.

The new launch system (Shuttle II, a modern ELV, or an SSTO vehicle) must be designed to meet operations requirements far beyond today's accepted environment:

- A highly available ground and vehicle system concept capable of launch on demand and launch on time so as to contain and support consistent launch costs
- Small, dedicated launch site crew (50 to 100)
- On-pad time < 10 days for payload and launch vehicle (deliver and launch)
- Modern vehicle avionics and subsystems capable of automated test, checkout, and vehicle health management
- Autonomous flight control and GN&C
- Fully supportable and maintainable vehicle and ground systems
- Maximum application of commercial/industrial standards and practices

- Highly automated vehicle/payload assembly, test, integration, and launch operations.

The largest contributor to achieving significantly lower launch systems cost will be application, implementation, and acceptance of commercial technology, standards, and practices in vehicle build, assembly, and checkout, as well as the highly automated launch site ground operations system.

Automated Launch Operations

Honeywell Space Systems has been studying and demonstrating automated launch operations for over three years with Government agencies and aerospace launch primes. We are convinced that intelligently applied, modern automation and controls systems technology transferred from Honeywell's worldwide industrial controls business can be cost effectively applied to a next-generation vehicle and its highly automated launch operations. After all, the U.S. industrial automation and controls business is market-dominant, worldwide, and very cost competitive — a sharp contrast to the U.S. commercial launch industry.

A state-of-the-art, Honeywell manufacturing automation supervisor (MAS) is representative of technology forming the heart of over 6000 major controls systems worldwide in factories, refineries, process plants, and *launch complexes* — 365 days a year (Figure 1). MAS features (Figure 2) are just what the automated launch system of the future will require. Table I summarizes capabilities and benefits of modern automation and controls technology to future launch operations.

The Vision Is Being Applied

Honeywell Space Systems has now applied industrial/commercial automation and controls products into Government contracts and is demonstrating the features and bene-

- c. Development/modification of a commercial launch facility and launch controls center at Cape Canaveral Air Force Station, Launch Complex 46.

These early demonstrations are validating many of the features *and* benefits listed in Table I. If one envisions an example for a future, highly automated launch control center using COTS products, the system may look like that shown in Figure 3.

In this launch system infrastructure, COTS products will be configured and integrated at the launch pad for local control and monitoring of vehicle, pad, and ground support systems. Fault tolerant controls will be applied where critical launch and safety requirements demand. At a remote site (5 to 10 miles) a generic, standard launch control center will be implemented with similar or common COTS products to provide highly automated monitoring, test and checkout, data archiving, vehicle subsystem performance trending, anomaly resolution, maintenance scheduling and support, and launch and mission decision support for both payload and launch vehicle systems. Standard interfaces to the launch pad, range safety, and families of launch systems will reduce the complexity, time, and cost of launching both Government and commercial launch vehicles.

The new launch system will be designed to support a complete lifecycle cost process that allows for total system data history, informed launch decisions, and ever improving system quality and performance.

- During the design and production phases, each subcontractor is provided the equipment needed to establish a data base on his hardware and software that will be delivered with each subsystem. In addition, launch personnel would do the same for items designed and developed at the launch site.
- In each of the phases for the ground systems (development, assembly, test, evaluation, integration and verification), information is stored for use at the subcontractor's plant and at the launch site.
- As the vehicle is assembled, the production data and factory information system is made available such that future launch operations can obtain data on the launch vehicle all the way back to the design phase.
- While the launch vehicle is on the pad, continuous test results are compared to previous tests performed on the vehicle and payload and even to previous launches.
- Post mission flight performance analysis results are used to provide a continuing update to the launch factory data bases for statistical quality control and statistical process control to assure continuing launch system quality improvement.
- Ultimately, financial data on the launch system is compiled from the distributed data that exists throughout the launch system life cycle, from the receiving dock to on-orbit. Engineering data on failures, repairs, and refurbishment are combined with financial and safety data at the factory to aid in management decisions.

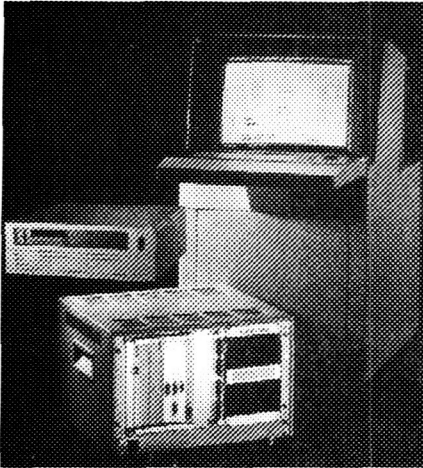


Figure 1. Honeywell Manufacturing Automation Supervisor

fits listed in Table I. MAS and Local Control System (LCS) elements are being implemented in three application areas:

- a. Test management, inertial measurement unit automated calibration, and statistical quality control analysis for the THAAD missile system avionics.
- zb. Facility management, performance monitoring, and system simulation and operator training for a next-generation USAF space systems ground operations facility.

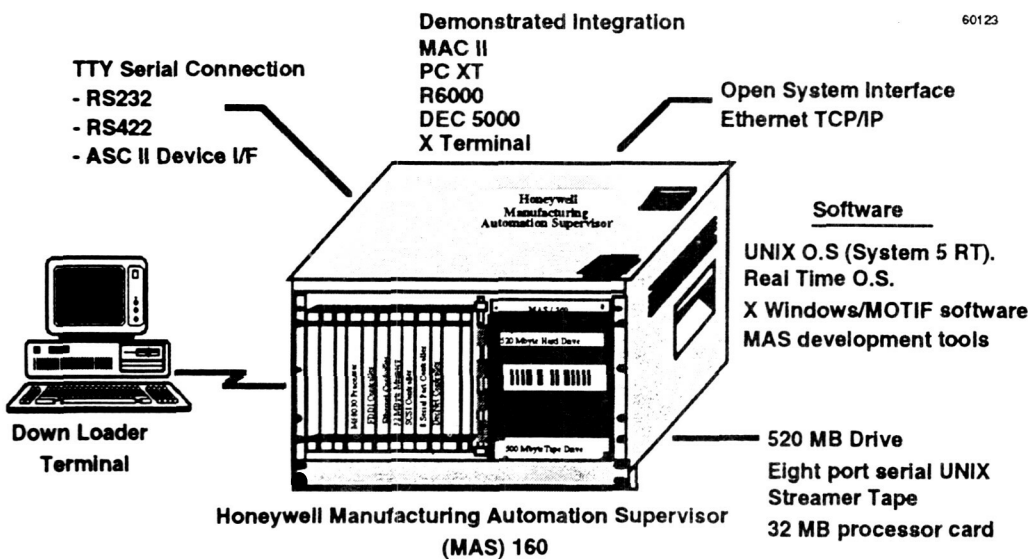


Figure 2. Honeywell MAS Supports Open Systems Architectures

*Table 1. Modern Automation and Controls
Features and Benefits*

Industrial Automation and Controls	Benefits for Automated Launch Operations
State-of-the-art hardware/software technology, continually upgraded in a "transparent" manner to stay competitive. Moving to multi-media, expert systems, and other leading edge technologies.	Products are proven in worldwide customer complex for over 30 years. Reaps benefits of ever increasing hardware/software performance and tools essentially paid for by the much larger commercial industry. The equipment is on the technology bow wave and is never obsolete.
Products are manufactured in large quantities, making them readily available and relatively low cost.	Up-front investment low. Development consists of applying the products to your controls requirements.
Products support open-systems/modular architectures and widely accepted industry standards (hardware, interfaces, software, protocols, development environment and procedures).	Multiple vendor products are interoperable, allowing the system to be expansive to evolving automation needs without large integration risks or costs. Government continuous acquisition and life cycle system (CALS) requirements are achievable.
Rugged industrial control products to meet near-Mil-Spec environments and demonstrate proven fault tolerance for critical processes.	Compatible with launch site environments. Fault tolerance can be tailored to meet high operational reliability and availability for system critical areas.
Commercial/industrial automation and controls are designed to be reliable (MTBF/MTBR) and supportable with ease of fault diagnosis, maintenance, and replacement.	Controls system downtime is minimized, availability is maximized, and unit replacement is readily available. Industrial customers are typically guaranteed less than 30 minutes down time per year for an entire system.
The software tools and extensive libraries that are required to automate a process with today's industrial automation equipment have been formatted to be extremely "user friendly". They are object oriented using symbology and icons that are familiar to process personnel just as object oriented programming languages using icons and menus have an interface that is familiar to office and management personnel.	The designers, users, and operators have the tools available, are validated, and allows a controls system strategy to be implemented, upgraded, and evolved on a cost effective basis.
Industrial systems support a total factory build, test, and on-site test data base with on-line SPC/SQC analysis, performance trending and a history data base for rapid operations decision making.	Provides a natural/cost effective capability that supports health management system implementation anomaly resolution, and launch/mission decision making.
Using industrial/commercial automation and controls is risk-adverse since they are used 365 days a year by worldwide users whose business and success depend on the technology.	The Government is just another user standing on the experience base of the industrial/commercial automation and controls world.
Industrial systems are designed for embedded simulation and operator training on-the-job in the controls environment.	Provides more effective training and more up-to-date problem resolution procedures.

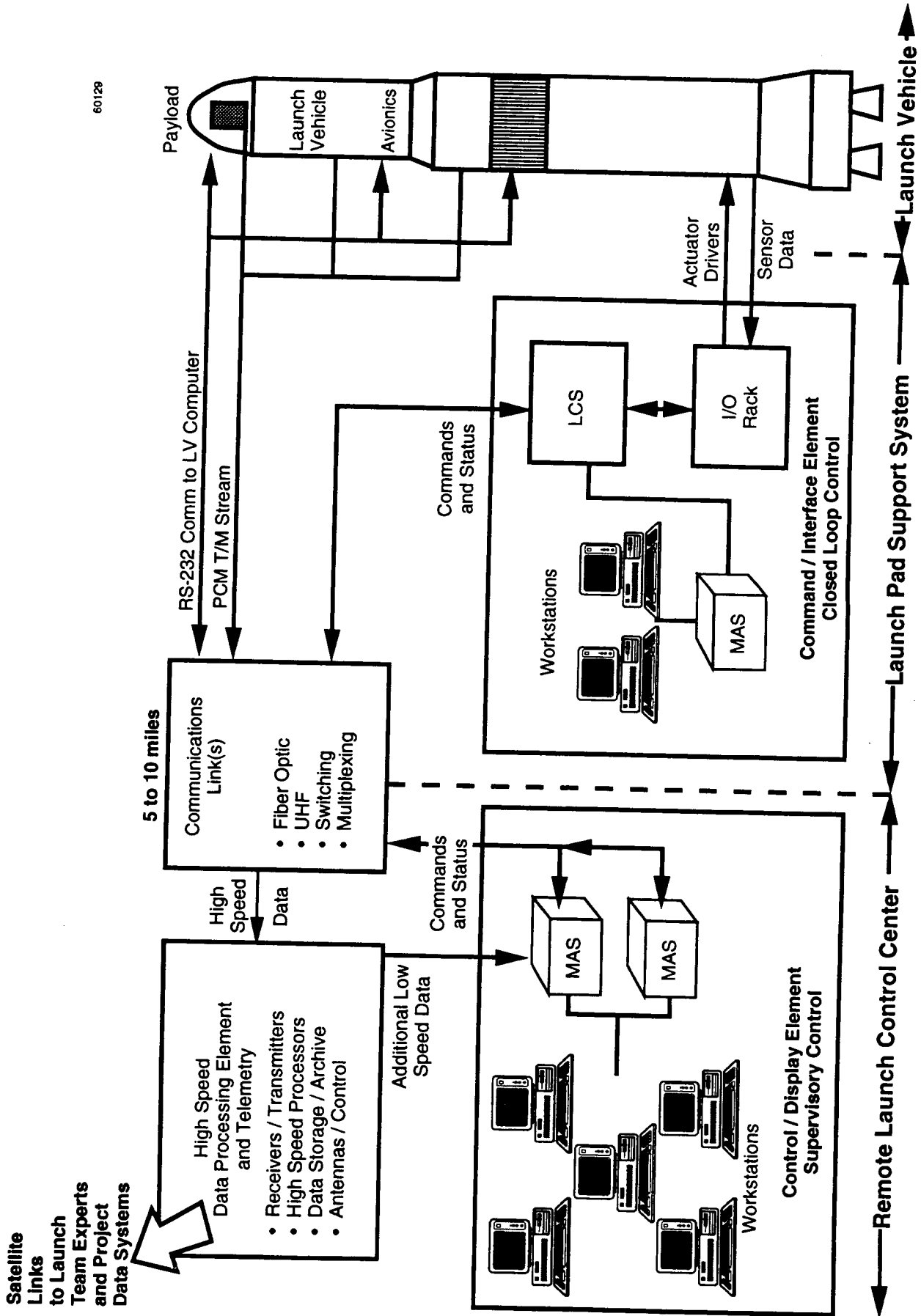


Figure 3. An Approach for Future Automated Launch Control

The remote launch control center (LCC) will contain the major data acquisition, information archiving, subsystem performance monitoring, and "user" remote command and control capability. All real-time launch operations data entering the LCC on the fiber optic link will flow to multiple data acquisition system (DAS) units where a "user" defined data subset will be transmitted to and archived on the terminals. The DAS units will function as the real-time data base for the LCC operations. A complement of standard Manufacturing Automation Supervision (MAS) processors will be used to gather real-time launch operations data and selected historic data and provide graphics, and data formats and displays to all the terminals at the block house locations. In the LCC the terminals will be configured to simultaneously support both data monitoring and remote command and control capability. Of course, all terminals throughout

the distributed LCC system will be capable of data monitoring.

The launch control center and pad will be linked via satellite to payload and launch vehicle primes and their key subcontractors. The launch site will have electronic data access to all component test data and design/parts information at the involved companies, while individual company experts will remain in their normal jobs at the company, and brought in electronically to review test data and support anomaly and problem resolution if it occurs (the true payoff for the vision of CALS).

Sensor and avionics on the vehicle and payload, distributed in the modern ground system, will perform health management, monitor performance trends, and (through on-line expert systems) support the new system's 10 days on-pad integration, checkout, and launch using a small launch crew (100 or less persons). The launch control and op-

erations center will be evolvable in technology, interoperable with a majority of COTS products, and be developed, maintained and managed with ever-improving software tools using object-oriented programming and expert systems. The system will cost effectively use the growing national "information highway and infrastructure." Its cost of ownership or Life Cycle Cost will be low and all launch pad and system users will reap the considerable benefits.

Yes, this vision needs to be validated and its cost benefits demonstrated, but these approaches are in development today. Full realization will be possible by the turn of the century. However, this positive picture depends on this nation making a near-term commitment to the next-generation launch system and its modernized infrastructure.

REFERENCES

1. Congressional Space Launch Review, Fall of 1993.

omit

**Session H1: HUMAN FACTORS
TECHNOLOGY DUAL USE**

Session Chair: John Schuessler

HUMAN FACTORS ENGINEERING: CURRENT AND EMERGING DUAL-USE APPLICATIONS

G. O. Chandlee, Lockheed Engineering and Sciences Company, 2400 NASA Road One, MC C81, Houston, TX, 77058, and B. S. Goldsberry, Lockheed Engineering and Sciences Company, 2400 NASA Road One, MC C44, Houston, TX, 77058

ABSTRACT

Human Factors Engineering is a multidisciplinary endeavor in which information pertaining to human characteristics is used in the development of systems and machines. Six representatives considered to be experts from the public and private sectors were surveyed in an effort to identify the potential dual-use of human factors technology. Each individual was asked to provide a rating as to the dual-use of 85 identified NASA technologies.

Results of the survey were as follows: nearly 75% of the technologies were identified at least once as high dual-use by one of the six survey respondents, and nearly 25% of the identified NASA technologies were identified as high dual-use technologies by a majority of the respondents. The perceived level of dual-use appeared to be independent of the technology category.

Successful identification of dual-use technology requires expanded input from industry. As an adjunct, cost-benefit analysis should be conducted to identify the feasibility of the dual-use technology. Concurrent with this effort should be an examination of precedents established by other technologies in other industrial settings. Advances in human factors and systems engineering are critical to reduce risk in any workplace and to enhance industrial competitiveness.

INTRODUCTION

Human Factors Engineering is a discipline in which information pertaining to capabilities, limitations, and other identifiable human characteristics is used in the development of systems and machines for efficient and effective use. Human factors includes the development and application of science and technology in fields such as biology, psychology, medicine, and engineering to enhance compatibility of people, machines, and environments to ensure safety and effective performance. The growth and recognized benefits of human factors is directly related to an increasingly

greater appreciation of the methods and techniques as they become better understood and the technology is applied in an ever expanding arena. The development of an approach to dual-use of human factors technology currently defines a new frontier that will result in even greater benefits to the public and private sectors.

Evidence of the growing interest in human factors engineering as a dual-use technology was revealed from a survey of selected individuals with knowledge concerning specific industries. Deliverables were identified that currently result or could potentially result from programs NASA is currently conducting. Examples of these deliverables are the Man-Systems Integration Standards (NASA Standard 3000) (NASA, 1967), and recording tools such as the Operational Experience Database (Callaghan, Gosbee, and Adam, 1992; Chandlee and Woolford, 1988). Other examples of deliverables include incident reporting tools such as those developed for the FAA (Armstrong, 1989) that could be developed as dual-use application for aerospace applications.

Traditional technology transfer takes place when an application in one discipline is applied to non-allied disciplines in new and unplanned ways. Technology transfer results after development work occurs and after applications in other industries are identified. A dual-use technology approach, however, refers to a more deliberate approach because the shared technology needs of unassociated industries are identified before development work proceeds or is completed (NASA, 1993). Identification of shared needs leads to defining a generic technology that results in development and finds an even broader use. This dual-use approach reduces the probability of redundancies because opportunities for shared technology are identified before actual development. Thus, substantial cost benefits can result.

METHODS

DEVELOPMENT OF POTENTIAL DUAL-USE TECHNOLOGIES

A total of 7 major technology categories were identified and 85 individual technologies were classified by major technology category. The 85 technology deliverables were selected as being current or planned products of NASA programs. The 7 major categories and the 85 technology deliverables were developed as part of the process of writing the document entitled NASA's Space Human Factors and Engineering Program Plan dated August 27, 1993 (NASA, 1993). The purpose of the classification is to assign each individual technology to an easily

identified activity or process. The classification is informative because it helps describe products of current NASA programs and makes identifying areas of potential dual-use technology easier. Examples of the 7 major categories are as follows:

I. Memoranda of understanding

Establish a formal NASA policy to utilize and effectively coordinate Agency human factors resources to address problems of risk assessment and accident avoidance.

II. Guidelines, requirements, standards, and criteria

Guidelines and standards for working with automated systems that effectively partition automated and human control functions.

III. Procedures, protocols, strategies, and processes

Standardized reporting system for critical incidents in space missions and ground control.

IV. Studies and assessments

Human factors research such as design of restraints and handholds.

V. Data and databases

Compilation of data such as a comprehensive human factors operational experience data base.

VI. Reports, documents, manuals, and plans

Various technical reports such as an updated and distributed anthropometric source book.

VII. Hardware and software

Hardware and software for the management, planning, and distribution of technologies such as tools for cost-benefit analysis.

RATING DUAL-USE TECHNOLOGY POTENTIAL

Six independent judges, considered to be experts, were contacted by either telephone or facsimile communications to rate the 85 technologies for dual-use potential. The individuals were thought to represent a fairly approximate cross-section of the government sector, private industry, and academia.

Each individual was provided with the list of 85 NASA Human Factors technology deliverables and was asked to assign a rating as to the

perceived dual-use of the deliverable. Each respondent was asked to assign to the deliverable a designation of High Dual-Use, Medium Dual-Use, or Low Dual-Use. The small number of individuals contacted, 6, was due to time and budgetary constraints, but should not effect the overall general conclusions. Skewing of results due to the effects of a single respondent are possible, but these effects are thought to be negligible.

RESULTS

The results of the survey were as follows: 75% of the identified deliverable items were indicated as high dual use at least once by at least one respondent. These 75% were broken down as follows: 25% of the technologies were marked as high dual-use technology once; 26% were marked twice; and 24% were marked as a high dual-use technology more than twice. The significance of these findings are that 75% of the technologies were identified as high dual-use at least once by at least one of the six survey respondents. A majority of the six survey respondents identified nearly one-quarter of the identified NASA technologies as high dual-use technologies.

The distribution of responses by individual technology category (Categories I - VII) is shown in Table I. As the percentages in the table indicate, respondents indicated that substantial percentages of technologies within all 7 major categories had high dual-use potential. These findings indicate agreement and consistency of thought because a significant proportion of NASA technologies were identified as dual-use and substantial internal coherency appears to exist among respondents.

CONCLUSIONS

The free flow of technical ideas from one domain to another non-allied domain inevitably results in innovation and in unforeseen and untried applications. Encouraging the flow of information from one individual and organization to another facilitates the effectiveness of the free flow process. For example, evidence indicates that aerospace engineers use the SAE Handbook of ground vehicle standards and often read Automotive Engineering (Rumbaugh, 1993). In addition, individual users of the Man-Systems Integration Standards, NASA Standard 3000, represent environments as diverse as the aerospace, automotive, and academic worlds. This technology "cross fertilization" will result in unanticipated applications and benefits. Parallel paths are being pursued in the realm of policy issue identification and formulation (see Auger and Facktor, 1993 for an example).

Table 1. Distribution of responses for technologies identified by survey participants as high dual-use candidates.

Technology category	Number and percentage (%) of technologies within individual categories	Number and percentage (%) of technologies within each category (relative to the entire number identified as high dual-use)
I. Memoranda of Understanding	2 (2.3%)	2 (100%)
II. Guidelines, Requirements, Standards, Criteria	32 (37.6%)	25 (78%)
III. Procedures, Protocols, Strategies, Processes	14 (16.5%)	12 (86%)
IV. Studies, Assessments	5 (5.9%)	2 (100%)
V. Data, Databases (e.g., Risk and Feasibility analyses)	2 (2.3%)	2 (100%)
VI. Reports, Documents, Manuals, Plans	3 (3.5%)	2 (67%)
VII. Hardware and Software system (tools, models, aids)	27 (31.2%)	18 (67%)
Totals	85 (100%)	63 (100%)

Successful identification of dual-use technology requires expanded input from relevant industries. This input could be gathered in an appreciably enlarged survey of individuals representing a cross-section of all public and private industrial sectors and academia. Such an in-depth survey should be conducted in two-steps: (1) identify and classify industrial sectors which could benefit from human factors technology; and (2) survey representatives from each of the sectors for space technologies that have potential dual-use. As an adjunct, cost-benefit analysis could be conducted for each of the sectors to identify the effectiveness of actual utilization of the dual-use technology. Concurrent with this effort should be an examination of precedents established by other technologies and in other government or private industrial settings. Given current trends, the application of dual-use human factors technologies to the commercialization of space is one application that could find immediate interest.

SUMMARY

The major highlights presented in this report are as follows:

1. Nearly 75% of the technologies were identified at least once as high dual-use by at least one of the six survey respondents.
2. Nearly 25% of the identified NASA technologies were identified as high dual-use technologies by one-half of the six survey respondents.
3. Substantial percentages of technologies within all 7 major categories had high dual-use potential. Results indicate that the level of perceived dual-use is independent of the technology category.

ACKNOWLEDGMENTS

Acknowledgments are due to the members of the NASA Headquarters Team responsible for the draft document entitled NASA's Space Human Factors and Engineering Program Plan dated August 27, 1993. The work reported here acknowledges the matrix of deliverables set forth in the document, the individuals responsible for identifying and enumerating those deliverables, and the individuals responsible for permitting use of the matrix of technologies and deliverables in the survey.

REFERENCES

1. Armstrong, M. E., "Human factors in incident investigation," Proc. Human Factors Soc. 33rd Ann. Mtg., 1989, pp. 1024-1028.
2. Auger, R.N., and Facktor, D. D., "Policy Issues in Space Analogues," IDEAA Two Conference, October 24-27, 1993, 7 pp.
3. Callaghan, T., Gosbee, J., and Adam, S. C., "A database containing operational experience in spaceflight," Paper no. 911499, Proc. 21st International Conference on Environmental Systems, 1991.
4. Chandlee, G. O. and Woolford, B., "Previous experience in manned space flight: A survey of human factors lessons learned," Proc. Human Factors Soc., 32nd Ann. Mtg., Anaheim, Ca., 1988, 49-52.
5. NASA, "NASA's Space Human Factors and Engineering Program Plan; An Integrated Research and Technology Program to Ensure the Safety, Reliability, and Success of NASA's Space Flight Missions," Preliminary Draft, Washington, D.C., August 18, 1993.
6. NASA, "Man-Systems Integration Standards (MSIS)," NASA Standard 3000, Vols. I and II, 1987.
7. Rumbaugh, M. "Technology transfer: A random function?," Aerospace Engineering, July, 1993, p. 6.

**RESEARCH ON
PERSONAL PROTECTIVE EQUIPMENT
FOR DUAL-USE TECHNOLOGY
AND TECHNOLOGY TRANSFER**

**Written by:
Donald C. Driggers
Lockheed Engineering & Sciences Company
February 1, 1994**

RESEARCH ON PERSONAL PROTECTIVE EQUIPMENT FOR DUAL-USE TECHNOLOGY AND TECHNOLOGY TRANSFER

317-54

4-17

ABSTRACT

The National Aeronautics and Space Administration (NASA) places highest priority on the safety of its astronauts and support personnel. Because this is so, and to ensure the continuation of this safety, the agency has undertaken to thoroughly research and develop and provide personal protective equipment (PPE) and individual life support systems (LSS) in support of manned spaceflight.

It is probable that technology developed for manned spaceflight in the field of PPE and individual LSS can be utilized in certain industrial/commercial endeavors. In an attempt to determine these other uses for this PPE and individual LSS, the Space Suit Systems Branch of the NASA JSC Crew Systems Division initiated a research project designed to assess potential common technology that could benefit industry. Such dual-use technology transfer could eventually involve a joint effort by Government and industry.

The research project took place over several months and involved discussions with various manufacturers/suppliers/users, as well as regulatory agencies and industries, of PPE and individual LSS. Research data was compiled and evaluated and a summary of significant findings is presented for identifying and establishing opportunities for future cooperation between Government and industry in the field of PPE and individual LSS.

INTRODUCTION

The list of critical/hazardous environments, shown in the appendix, were identified as those most likely to require personal protective equipment (PPE) or an individual life support system (LSS). Representatives of industries and/or regulatory agencies, shown in Table 1, were contacted for information regarding types of PPE or LSS that are currently in use.

Table 1 - PPE and LSS STANDARDS/CERTIFICATION GROUPS

- National Institute of Occupational Safety and Health (NIOSH)
- Occupational Safety and Health (OSHA)
- Safety Equipment Institute (SEI)
- Mine Safety and Health Administration (MSHA)
- Texas Research Institute (TRI)

REGULATORY AGENCIES

National Institute of Occupational Safety and Health (NIOSH)

NIOSH is the agency of the Government that does the research, etc., for the various safety standards that are in effect. This agency holds hearings and develops and publishes the Criteria Standards.

NIOSH does health hazard evaluations in the field for which respiratory protection is needed. Field personnel must work around several types of materials, (e.g., asbestos, fiberglass, and other abrasives). These personnel use a self-contained breathing apparatus (SCBA) with a compressor and air hoses (or umbilicals).

The compressors are big, bulky, and hard to carry, as well as hard to use. NIOSH would very much like to have smaller, more compact, and easier to use SCBAs. In addition, the SCBAs are not conducive to very long work periods at one time, and an SCBA with a longer use-life would be welcomed.

Occupational Safety and Health Administration (OSHA)

OSHA publishes the final rules regarding safety standards for industry in the United States, based on research done by NIOSH.

OSHA is responsible for the safety rules regarding toxic or hazardous environments. OSHA is also responsible for the safety standards for safety equipment, such as SCBAs, respirators, etc. The materials used for the most part in these environments are Nomex (trade name used for a flame-retardant material) and Tyvek (trade name for a fluid-resistant fabric).

Fabrics that are lighter weight and cooler than the presently available fabrics and that would form a barrier between hazardous material and the skin would be in much demand. A fabric with more strength and durability than Tyvek would be in demand, as Tyvek tends to tear on the seams very easily. Nomex is very heavy and hot and a more lightweight and cooler flame-retardant material would be a big improvement.

Safety Equipment Institute (SEI)

This agency certifies garments for use by firemen, etc., because the NFPA requires additional testing after NIOSH certification. The standards by which the certification takes place are revised every 5 years.

A standards development meeting for personal protective equipment will take place yearly.

Mine Safety and Health Administration (MSHA)

MSHA certifies or approves safety equipment for underground use by using regulations from the Code of Federal Regulations. The codes are updated regularly by MSHA engineers, technicians, and inspectors.

Texas Research Institute

This company specializes in research and development as well as certification of PPE and LSS for chemical protective clothing and barrier-resistant garments.

ENVIRONMENTAL AND HEALTH GROUPS

Lockheed Environmental Sciences and Technology Co. (LESAT)

For personal protective equipment, LESAT uses mostly coveralls, gloves, and booties along with full-face respirators. They also use radiation protection suits. Cooler and lighter-weight garments would be preferred. Tyvek is used extensively but its cost has caused LESAT to use reusable cotton in some areas. The cotton garments are sent for laundering to one of several companies that use sophisticated filter cleaning systems for laundering contaminated clothes.

The SCBAs that are worn for protection are heavy, expensive, and restrictive such that no one likes to use them. Improvements in SCBAs (e.g., lighter weight, cooler, less restrictive) are desired.

Southdown Transportation Storage and Disposal Facility

Breathing equipment is used for all opening and closing of containers in which hazardous equipment has been stored. No cooling suits are used, but this company would like to see something on the market. Nomex and Tyvek suits are used extensively throughout the facility.

FIRE SERVICES

NASA Fire Department

Breathing equipment is used on a regular basis. No cooling suits are used. For isolation from biological/chemical hazards, suits used are designed for use in spills and cleanup of hazardous materials. Each firefighter is issued a suit for his/her duty, including boots, helmet, coat, and pants.

National Fire Protection Association (NFPA)

NFPA would like to see lighter weight garments that would reduce heat stress on the wearer without reducing any of the protective properties of the garment. In addition, it is important that each item of the ensemble interface with other items in the ensemble as well as with other protective items, i.e., headgear and respirators.

As protection against hostile environments, NFPA uses mostly Nomex garments. In working with hazardous materials and in rescue efforts, umbilicals are used; however, they are not used during firefighting.

Red Adair Firefighting Services

Cotton overalls with Nomex underwear are worn to fight oil fires. Water is used as a curtain to protect against heat. Sufficient, breathable air is not a problem in most cases, so breathing equipment is not used. This organization would be interested in some type of suit to protect their firefighters from chemicals.

MANUFACTURING GROUPS

E.I. DuPont de Mours & Company

This chemical company has use for two types of garments - those used in non-fuming acids and those used with fuming acids. For non-fuming acids, garments manufactured of rubber/Tyvek are used. Fully encapsulated suits are used for fuming agents (e.g., SO₃, liquid chlorine, anhydrous ammonia) and are made of a heavy plastic that is very heavy and stiff. These types of acid spills can be very cold and the material sometimes cracks or does not move with the wearer. A major use of the encapsulated suits occurs in cryogenics (e.g., the cleaning up of transportation spills).

DuPont would like to see a material that would stand up to the cold and that could withstand the corrosive chemicals and still remain flexible.

Bethlehem Steel

Bethlehem Steel workers use SCBAs with compressed air from a hose (umbilical) for particular jobs that include abrasive materials.

Often, workers are required to be exposed to temperatures of ~3000 degrees, and for this they wear heat-resistant garments. They would like to see cooler and lighter weight garments for this purpose.

Ford Motor Company, General Motors Corporation

Both of these companies were reluctant to answer any questions, so very little information was received. Some automobile factory workers do wear respirators and masks, but no head protection.

Wilson Greatbatch Ltd.

This company manufactures lithium batteries as well as other products whereby they need a waterless environment (or dry room). Explosion and environmental contamination are a problem. Safety glasses and face masks are always worn. Latex gloves deter hand moisture from reacting with the lithium.

Dupont Diagnostic Imaging Co.

This is the branch of DuPont that produces film for X-rays. This is done in an ultra-clean room called the coating laboratory. Polyester film base is coated at very high speeds. Personnel in the coating laboratory wear disposable, lintless protective gear that includes booties, jumpsuits, helmets, etc. These garments are usually made of Tyvek or Slantara (both materials are manufactured by DuPont Chemicals). Other parts of this company have various levels of clean rooms.

Eastman Kodak

A clean room is used in the manufacturing of film. Lintless garments of Tyvek are worn.

Stone Container Corp. (a division of Champion Paper Co.)

The people who work in the paper industry (mills, etc.) are subject to a great deal of heat and dust and debris. The protective equipment used consists mostly of respirators and hardhats. An interest exists in light-weight hardhats that would still meet safety requirements. The respirators use spun aluminum bottles that are fairly lightweight for their air supply. This industry would like to see lightweight cooling garments that are not too expensive.

Newport News Shipbuilding Safety Dept.

This shipbuilding business uses a wide variety of negative-pressure respirators, including SCBAs, during welding, painting, blasting, etc., operations. Nuclear systems work, where an adequate supply of backup air is necessary, is also performed here. Much of the work involves being in confined areas on ships, and these areas are potentially hazardous. Robotics products that could reduce human exposure to hazardous materials are not presently available, or what is available is too expensive. This industry would be very interested in any product that could accomplish this at a marketable cost.

NUCLEAR PLANTS

South Texas Electrical Generating Station

In radiologically-controlled areas, reusable cotton coveralls are worn. These do not give any protection against radiation, and because of this, the time spent in these areas is regulated depending on the radiation level.

This industry would like to see a garment on the market that would allow workers to work for longer times in radiation-exposed areas.

MEDICAL GROUPS

National Center for Disease Control

A microbiologist in the maximum containment laboratory at the Center for Disease Control (CDC) spoke about the necessity of protection from blood-borne pathogens and highly virulent organisms that exist in the virology laboratories and the high-containment areas of CDC.

CDC uses what they call a "space suit" in the high- and maximum-containment labs. This is a one-piece positive-pressure vinyl suit which uses compressed air that is cooled and rehumidified. The conditioned air is kept at ~60 degrees.

The biggest problem with the space suits presently in use is the noise from the air being brought in from the compressor (e.g., there is not a good "muffler" system). The CDC also relies heavily on gloves and would be interested in robotics, if available.

Harris County Health Department

High-efficiency particulate arrestor (HEPA) respirators are used when Health Department personnel respond to requests for lead or asbestos testing or cleanup. Safety shoes, glasses, hardhats, masks, and coveralls are worn as well. This is the department that is tasked with inspections for toxic materials and emergency response for high contaminate pollution control.

General Medical

General Medical is the second largest supplier of medical equipment and supplies in the U.S. Gloves are the most widespread personal protective equipment supplied by this company. Gloves are made mostly of latex, and are manufactured for the most part abroad. Some hypoallergenic gloves are available for those medical personnel who cannot wear latex, and because most gloves are powdered with talc or cornstarch that some people are allergic to, gloves using no powder are available.

After gloves, the order of the volume amount of PPE supplied by General Medical is as follows:

2. Masks
3. Eyewear (goggles, shields)
4. Gowns (disposable - either impervious, or fluid-resistant)
5. Shoe covers and caps
6. Resuscitators, now with a one-way valve
7. Biohazardous waste containers.

Life support equipment, such as ventilators, are sold directly from the manufacturer.

ENGINEERING COMPANIES

Rust Construction Co.

This is a very large construction company that builds large projects and that includes an environmental section. The protection used by this company in the construction section consists mostly of hard hats and underground protection. The environmental section of Rust does remedial-type work throughout the U.S. This company would like to see cooler and lighter-weight suits that could be worn for a longer period of time in a contaminated area, as well as garments made of a more durable material.

MILITARY

U.S. Army Edgewood Research, Development, and Engineering Center

This is an Army research laboratory where masks for use in helicopters and tanks as well as with the Mission-Oriented Protective Posture (MOPP) gear are developed. The MOPP gear is part of the nuclear/biological/chemical uniforms that are used in time of war. MOPP gear consists of a cotton shell with carbon-impregnated polyurethane foam. The gas masks for MOPP gear use filtered air and must be integrated into the helmets.

U.S. Army Natick Research and Development Center

Several types of Army PPE and LSS are in the research and development phase at the present time. Many of the items being researched deal with heat stress, weight and bulk, and chemical protection. Market surveys are being conducted to determine whether commercially available items could replace materials now in use.

Navy Clothing and Textile Research Facility (NCTRF)

The Navy is presently involved in research and development for fire-fighting protection for Air Force rocket launch personnel, i.e., a fully encapsulated suit. This research and development effort is taking place in conjunction with the Air Force at Patrick Air Force Base.

For other situations, the Navy PPE ranges from basic utility clothing for fighting fires to reusable flame-retardant cotton and SCBAs.

MISCELLANEOUS

NASCAR

NASCAR drivers use fire-resistant clothing made out of Nomex. Several layers of Nomex are used and this makes the suits very hot. A new helmet with a small air fan that pumps fresh air into the helmet is now available and heavily used.

Liquid-cooled garments have been used by drivers, but the failure factor is high for these suits and when they fail the cool water turns into hot water, causing great discomfort.

National Hotrod Association (NHRA)

Garments used by hotrod drivers vary from one to several layers of Nomex. These garments are hot and heavy. Helmets with built-in vents or "breathers" pipe in filtered compressed air.

Because these drivers only race for approximately 1/4 mile, they are not interested in a liquid-cooled garment.

Championship Auto Racing Teams (CART)

CART is the group that regulates Indy-type auto racing. Clothing worn by race car drivers is usually Nomex or a similar material that is very heavy. Some drivers wear liquid cooling suits and headsox. A new type of driving helmet has recently been introduced for aerodynamic stability which cuts down the buffeting and energy absorption.

CART would like to see a continual development of safety products as well as an improvement in the cooling garments.

FINDINGS:

The opinions of nearly all the industries/agencies contacted were similar on the following improvements that they prefer in personal protective equipment survey that was taken by the world safety organization.

IMPROVEMENT DESIRED	PERCENTAGE
Improved gloves and suits -	92%
More comfortable, more user-friendly equipment	81%
Improvements in respirators, including noise control	38%
Reliable, durable, available cooling suits/equipment	35%
Recyclable, less expensive	28%
Super-absorbent materials	26%
Eye, ear protection, including hard hats	25%

CONCLUSION

A survey of PPE usage was conducted through the World Safety Organization's Fifth Annual Safety and Accident Prevention Congress held in Memphis, Tennessee, September 1993. The survey results indicated a need for a joint industry/government workshop to address the issues and development of PPE and LSS.

APPENDIX

CRITICAL/HAZARDOUS ENVIRONMENTS

- Environmental remedial groups
 - Lockheed Environmental Sciences and Technology Co. (LESAT)
 - Southdown Transportation Storage and Disposal Facility
 - Marsh & McLennan, Inc.
- Fire departments
 - NASA Fire Department
 - National Fire Protection Association (NFPA)
 - Safety Equipment Institute (SEI)
 - Red Adair Firefighting Services
- Chemical plants/refineries
 - E.I. DuPont deMours & Company
- Nuclear facilities
 - South Texas Electric Generating Facility
- Steel plants, mines, or other plants with smokestacks, etc.
 - Bethlehem Steel
 - Mine Safety and Health Administration
- Automotive, including painting, welding, etc.
 - Ford Motor Company
 - General Motors Corporation (GMC)

- **Medical/biological/dental, etc.**
 - National Center for Disease Control (CDC)
 - Harris County Health Department
 - General Medical Corp., a supplier of medical equipment
- **Construction**
 - Rust Construction Co.
- **Waterless environments**
 - Wilson Greatbatch Ltd.
- **Ultraclean work environments**
 - DuPont Diagnostic Imaging Co.
 - Eastman Kodak Company
- **Military**
 - Navy Clothing and Textile Research Facility (NCTRF)
 - U.S. Army Edgewood Research, Development, and Engineering Center
 - U.S. Army Natick Research and Development Center
- **Race car drivers**
 - National Association of Stock Car Drivers (NASCAR)
 - National Hotrod Association (NHRA)
 - Championship Auto Racing Teams (CART)
- **Paper mills**
 - Stone Container Corp., a division of Champion Paper Co.
- **Shipyards**
 - Newport News Shipbuilding

GMT

Session H2: HUMAN PERFORMANCE EVALUATION

Session Chair: Barbara Woolford

5/8-53

ADVANCED VIDEO ANALYSIS NEEDS FOR HUMAN PERFORMANCE EVALUATION

Paul D. Campbell, P.E.
Lockheed Engineering and Sciences Company

Mail Code C44
2400 NASA Road 1
Houston, Texas 77058

ABSTRACT

Evaluators of human task performance in space missions make use of video as a primary source of data. Extraction of relevant human performance information from video is often a labor-intensive process requiring a large amount of time on the part of the evaluator.

Based on the experiences of several human performance evaluators, needs were defined for advanced tools which could aid in the analysis of video data from space missions. Such tools should increase the efficiency with which useful information is retrieved from large quantities of raw video. They should also provide the evaluator with new analytical functions which are not present in currently used methods.

Video analysis tools based on the needs defined by this study would also have uses in U.S. industry and education. Evaluation of human performance from video data can be a valuable technique in many industrial and institutional settings where humans are involved in operational systems and processes.

INTRODUCTION

Background

The Human Factors Project Office (HFPO) at the Johnson Space Center provides human research and engineering expertise for a variety of NASA programs. In the course of this work, HFPO personnel often use recorded video data in evaluating the human performance of ground test subjects and flight crew members. A recent example of this type of HFPO work is the series of extravehicular activity tests conducted on Space Shuttle missions during 1993. Space-suited crewmembers performed representative tasks outside the Shuttle, and HFPO personnel evaluated the results using recorded video data taken from Shuttle cameras.

The large quantity of data which resulted from video recording during these Shuttle tests required extensive time and effort on the part of the HFPO evaluators to extract valuable human performance information. Tools that reduce this effort would be useful to the

HFPO and other organizations which routinely use recorded video as a source of information.

Purpose

This study was performed to define and integrate the needs of HFPO personnel for computer-based tools which can be applied to human performance evaluation using recorded video. The statement of needs resulting from this study can be used as guidance for future tool development efforts at NASA and may be useful to commercial developers of performance analysis tools.

METHOD

Product Examinations

Examinations of three video analysis tools were used to create an understanding of the current state of development in this area. The tools that were examined were the Posture Video Analysis Tool (reference 1), the Timelines system (reference 2), and the Vision 3000™ system (reference 3).

The Posture Video Analysis Tool (PVAT) is an interactive system developed at the NASA Johnson Space Center and hosted on an Apple Macintosh® personal computer. PVAT allows its user to compile structured information while viewing a video tape. It is primarily designed for posture analysis, but with modifications may be useful for general task analysis as well. PVAT has been tested using Space Shuttle crewmember video, and was found to be valuable in human factors evaluation of postural issues.

The Timelines system is a Macintosh®-based tool developed by the University of Toronto for video annotation or "coding". Timelines is designed to:

"...address a problem that many researchers in human interaction experience as a result of using video tape. Video is a very rich recording medium but the process of analyzing video takes significantly longer than the viewing time of the tape. Timelines allows a skilled coder to code without stopping the tape. This drastically reduces the time to code....

The basic task in coding a video tape is to reduce the amount of information while retaining what is relevant. This usually means moving from qualitative data to quantitative data" (reference 2).

The Vision 3000™ system is a personal computer-based tool that supports analysis of human postures and force-application tasks (lifting, etc.) from recorded video data. Vision 3000™ :

"...is a software solution for performing both quantitative and qualitative biomechanical and ergonomic evaluations of job tasks and other physical activities that have been recorded on videotape" (reference 4).

Human Performance Evaluator Interviews

The members of the HFPO were asked to provide data on their prior experiences with video analysis. The questions in Table I were used as the basis for initial data collection interviews.

Table I: Video Analysis Experience Questionnaire

Describe instances when you have used recorded video in evaluating human performance.

List specific types of data you gathered from video in the above instances.

Describe the final products to which your video analysis contributed.

Describe the video evaluation activities that required the most time on the part of the evaluator.

Describe any other specific computer-based tool capabilities which might help you in future video analysis for human factors purposes.

What constraints should be placed on a video analysis tool to make it useful?

The responses to the experience questionnaire were compiled, and from these responses, potential video analysis tool functions were extracted. The list of potential tool functions was then given back to the respondents in the form of a needs survey to quantify the perceived usefulness of each function. The needs survey that was distributed to HFPO personnel is shown in Table II. This survey resulted in indications of needs that may be considered by video tool developers when defining new video-based human factors analytical systems.

Table II: Video Analysis Needs Survey Form

Rating	Potential Tool Function
	Evaluator Annotation:
_____	Structured notes (predefined text items the evaluator tags to the video)
_____	Unstructured notes (free-form evaluator text inputs tagged to the video)
_____	Graphics overlay (addition of evaluator markings on frames of video)
	Data Enhancement and Information Extraction:
_____	Detection of the presence of a person(s) in the video frame
_____	Tracking of a person or other object in the scene
_____	Measurement of the motions (linear and rotational) of a person or object
_____	Detection of human speech
_____	Transcription of speech into text
_____	Speech stress analysis (based on spectral content, etc.)
_____	Detection of the start and/or end of a particular task or event
_____	Elapsed time-keeping of a particular task or event
_____	Creation of a task taxonomy/dictionary based on tasks detected
_____	Counting of the number of occurrences of each task or event
_____	Environmental event detection (such as sunrise, sunset, etc.)
_____	User selection of a portion of the video image to be enlarged to fill frame
_____	Enhancement of speech to make it more intelligible
_____	Speech compression during fast playback (to keep voices intelligible)
	Control of Video Playback:
_____	Display of the time associated with the video
_____	Simultaneous side-by-side display of more than one video
_____	Playback speed control (fast/slow, forward/reverse, pause)
_____	Reduce playback to normal speed when speech detected, increase it to high speed when speech ends
	Contextual Data:
_____	Synchronize video with other displayed data (e.g. plots of heart rate, etc.)
	Analysis:
_____	Support statistical analyses of evaluator annotations or other data
_____	Perform plotting of data generated during the video analysis session

Functional Needs

Table III shows the average scores accorded by the respondents to the potential tool functions they were asked to rate. The following scale was used:

- 0 = No usefulness in your video analysis
- 1 = Some usefulness in your video analysis
- 2 = Substantial usefulness in your video analysis
- 3 = Essential in your video analysis.

Table III: Respondents' Scoring of Potential Functions

Function:	Respondent: A	B	C	Average Score:
Structured Notes	1	2	2	1.67
Unstructured Notes	2	2	2	2.00
Graphics Overlays	2	3	3	2.67
Detect Persons	3	3	3	3.00
Track Persons	3	3	3	3.00
Measure Motions	2	3	3	2.67
Detect Speech	2	2	3	2.33
Transcribe Speech	1	2	3	2.00
Speech Stress Analysis	2	1	3	2.00
Detect Task Start/End	3	3	3	3.00
Keep Elapsed Time	3	3	3	3.00
Create Task Taxonomy	2	2	3	2.33
Count Task Occurrences	2	2	3	2.33
Detect Environment Events	1	1	2	1.33
Enlarge Part of Image	2	2	3	2.33
Enhance Speech	2	3	3	2.67
Speech Compression	2	3	3	2.67
Display Video Time	3	3	3	3.00
2 Side-by-Side Images	2	2	2	2.00
Video Speed Control	3	3	3	3.00
Reduce Speed for Speech	2	3	3	2.67
Synchronize with Other Data	2	2	2	2.00
Statistical Analysis	2	2	2	2.00
Data Plotting	2	2	2	2.00
Avg. Score by Respondent:	2.13	2.38	2.71	2.40

Note that in two cases, "transcribe speech" and "speech stress analysis", there was a wide range of individual scores (1 to 3) from the respondents, with one respondent assigning "essential" to both functions and the other respondents assigning "some" or "substantial" usefulness to them.

Figure 1 illustrates the distribution of average scores across the range of possibilities. It is evident that the respondents generally perceived substantial usefulness in all the potential functions. All average scores were greater than 1 (some usefulness), and more than 90 percent of functions (22 out of 24) had average scores of 2 (substantial usefulness) to 3 (essential).

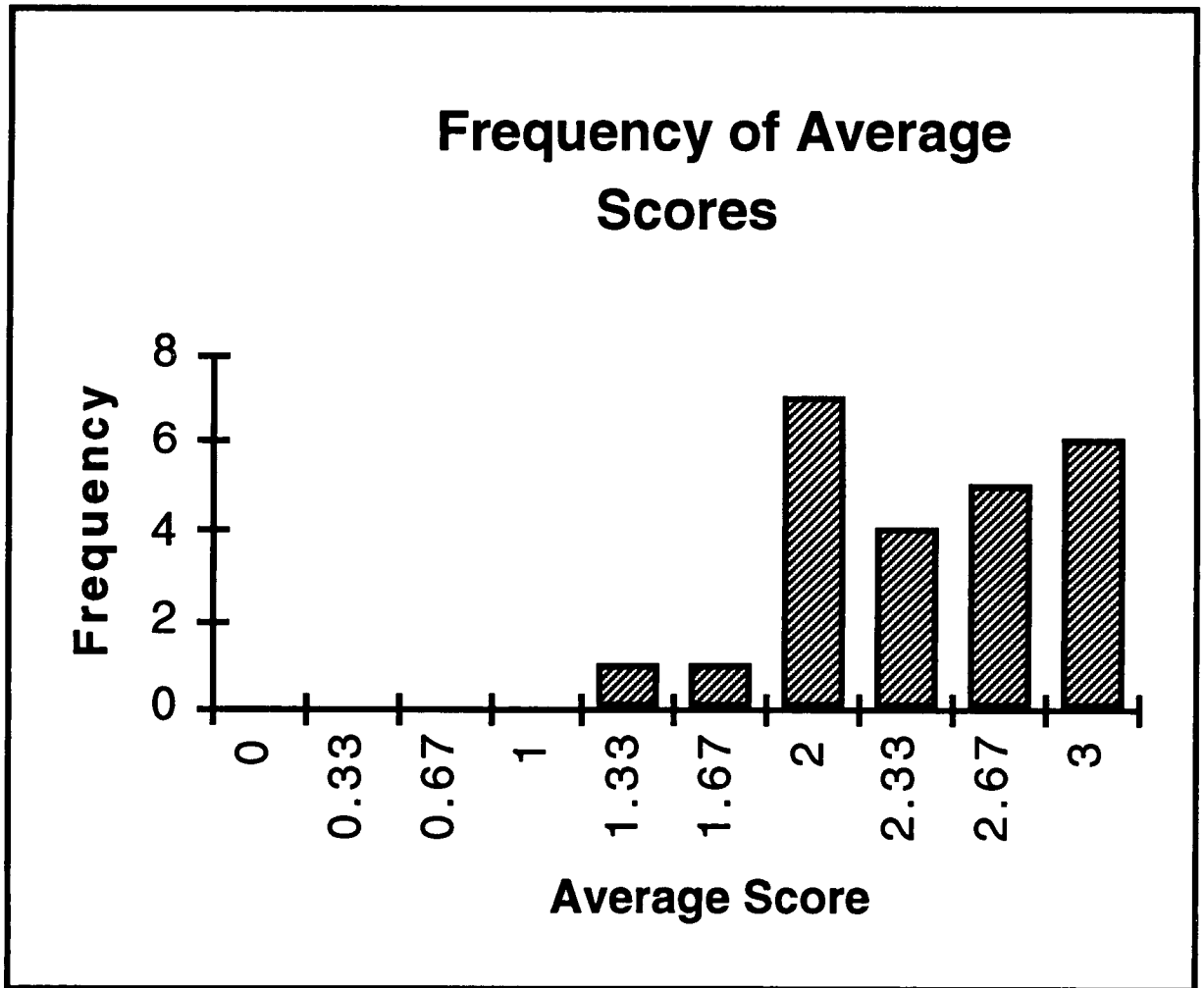


Figure 1: Frequency of Average Scores

Table IV compares the three video analysis tools that were examined to the functions scored in the needs survey. An "X" indicates that the tool supports the function. Indices of tool functionality are given at the bottom of the table, with the numbers indicating the percent of tool capability across all the advanced functions listed. The tools that were examined include only a small fraction of the advanced video analysis functions that were considered in this study.

Table IV: Tool Capabilities Compared to Functional Needs

Function:	Tool: PVAT	Timelines	Vision 3000™
Structured Notes	X	X	
Unstructured Notes		X	X
Graphics Overlays			X
Detect Persons			
Track Persons			X
Measure Motions			X
Detect Speech			
Transcribe Speech			
Speech Stress Analysis			
Detect Task Start/End			
Keep Elapsed Time		X	
Create Task Taxonomy			
Count Task Occurrences	X	X	
Detect Environment Events			
Enlarge Part of Image			X
Enhance Speech			
Speech Compression			
Display Video Time	X	X	
2 Side-by-Side Images			
Video Speed Control		X	
Reduce Speed for Speech			
Synchronize with Other Data			
Statistical Analysis			X
Data Plotting			X
Tool Functionality Index:	12.5%	25%	29%

CONCLUSIONS

Findings from this study indicate that there is a need for the development of advanced computer-based video analysis tools that will meet the needs of human performance evaluators. These tools should include many functions that are not available in currently available tools.

NASA organizations that develop video analysis tools should consider the findings of this study when developing new tools for human performance evaluation. The NASA Human Factors Project Office is a potential user of such tools.

When developing video-based analytical tools, NASA should examine their dual-use technology potential. Institutions in fields such as medical research and education as well as commercial industries could use tools having many of the functions defined in this study. Evaluation of human performance from video data can be a valuable technique in many industrial and institutional settings where humans are involved in operational systems and processes.

ACKNOWLEDGMENTS

The author thanks the members of the NASA Human Factors Project Office for their support in this study.

REFERENCES

1. Whitmore, M., Merced-Moore, D., and Adam, S. C., "PVAT-A Video Analysis Tool for Microgravity Posture Evaluation", in *Proceedings of the Human Factors Society 37th Annual Meeting*, 1993.
2. Owen, R. and Harrison, B., "Timelines: A System for Analysing Time Based Data", University of Toronto, 1993.
3. Promatek Medical Systems, Inc., "Vision 3000™, The Leading Video-Based System for Motion & Job-Task Analysis", Joliet IL, undated.
4. Promatek Medical Systems, Inc., "Promatek Vision 3000™, Presentation Summary", undated.

USE OF VIDEO ANALYSIS SYSTEM FOR WORKING POSTURE EVALUATIONS

Timothy D. McKay* and Mihriban Whitmore*
Human Factors and Ergonomics Laboratory
Flight Crew Support Division
NASA/Johnson Space Center
Houston, TX 77058

519-54

ABSTRACT

In a work environment, it is important to identify and quantify the relationship among work activities, working posture, and workplace design. Working posture may impact the physical comfort and well-being of individuals, as well as performance. The Posture Video Analysis Tool (PVAT) is an interactive menu and button driven software prototype written in Supercard™. Human Factors analysts are provided with a predefined set of options typically associated with postural assessments and human performance issues. Once options have been selected, the program is used to evaluate working posture and dynamic tasks from video footage. PVAT has been used to evaluate postures from Orbiter missions, as well as from experimental testing of prototype glove box designs. PVAT can be used for video analysis in a number of industries, with little or no modification. It can contribute to various aspects of workplace design such as training, task allocations, procedural analyses, and hardware usability evaluations. The major advantage of the video analysis approach is the ability to gather data, non-intrusively, in restricted-access environments, such as emergency and operation rooms, contaminated areas, and control rooms. Video analysis also provides the opportunity to conduct preliminary evaluations of existing work areas.

INTRODUCTION

Working posture can be affected by the interaction of the operator with the design of the workplace environment and the tasks to be performed. Some factors that contribute to an assumed working posture are:

- 1) workstation layout such as height and location of benches,
- 2) equipment and tool design characteristics such as orientation of the handles,
- 3) work methods such as lifting or translation technique used,
- 4) anthropometric characteristic of the operator, and
- 5) task requirements.

Medical and ergonomic field studies suggest that the awkward working postures may cause pains in muscle and connective tissues of tendons, joint capsules, and ligaments (Grandjean and Hunting, 1977). In many work environments, there is a need to identify and further quantify the relationship among the work activities,

*Lockheed Engineering and Sciences Company, NASA Rd. 1, MC-C81, Houston, TX 77058

working posture, and workplace design so the postural stress can be quantified and the specific causes of awkward posture identified. These results could benefit engineers in designing and/or modifying the environment or task for safety, comfort, and productivity.

In recent ergonomic literature, posture evaluations have ranged from physiological measures of workload and subjective discomfort rating scales to observational data (Fisher and Tarbutt, 1988; Keyserling, 1986a; Keyserling, 1986b; Leonard and Keyserling, 1989). Several systems have recently been developed for measuring working posture. The Posturegram system describes the position of 16 links relative to the 3 reference planes. The rater needs to estimate 48 limb angles that requires several minutes to record a single posture. The system is more applicable for static tasks with few posture changes as opposed to dynamic work (Priel, 1974). Ovako Working Posture Analysis System (OWAS) uses standard postures and a three-digit code. The first digit describes the position of the trunk, the second digit describes the position of the arms and the third digit describes the position of the legs. The OWAS requires a few seconds to record the posture, however, the available posture categories in this system may be considered too broad depending on the level of precision required to describe the working posture (Leonard and Keyserling, 1989; Karhu, Kansil and Kuorinka, 1977). Posture Targeting is another system developed to observe the worker and record the position of the head, trunk, upper arms, lower arms, upper legs and lower legs by making 10 marks on a chart. The marks represent the position of each body segment with respect to a standard reference posture. This system most applicable to static tasks (Corlett, Madeley, and Manenica, 1979). A computer-aided system was developed by Keyserling (1986a, 1986b) to observe and record the posture by selecting from a menu of standard positions of neck, trunk, shoulders, and lower extremities. Most of the systems described above require that specific reference points be identified prior to classifying the working posture and that the analyst have control of the data collection procedures. Unfortunately, strict control over the data collection is not always possible or feasible in dynamic work environments such as a nuclear power control room, spacecraft, or emergency room.

In conducting an observational analysis of existing Shuttle mission footage, the Human Factors and Ergonomics Laboratory (HFEL) at the NASA Johnson Space Center found that important information such as different postures, duration, frequency, and design issues could be extracted from the video, if a uniform methodology was available. The HFEL then developed the Posture Video Analysis Tool (PVAT) to provide the ability to quantify observational posture data.

The use of video footage as a means for collecting data non-intrusively is augmented with PVAT. This tool can be used to define critical posture categories, and dynamic or repeated body movements for use in the design of tools, workstations and

habitats. PVAT provides the structured methodology needed to extract and classify working postures, even from videos not recorded specifically for experimental analysis. PVAT helps an analyst identify problem areas that contribute to poor postures and decreased human performance and thus brings into focus those design issues requiring more rigorous evaluations. It should be noted that PVAT does not eliminate the need for traditional data collection and assessment techniques. The data can then be analyzed with off-the-shelf systems such as the ARIEL system, Motion Analyzer, or Vision 3000 for further quantification of the critical working postures. The paper herein provides a brief description of PVAT, a brief description of users testing the tool, and a discussion of applying PVAT to other industries.

TOOL DESCRIPTION

PVAT is an interactive Macintosh menu and button driven SuperCard™ software prototype. Prior to the video analysis process, the analyst must first define the assessment parameters. These parameters are entered using keyboard and pulldown menu selections from the PVAT "Startup" screen, see Figure 1. For reference purposes, the user may enter the video footage title, video camera location, a brief task description, and the rater's name. Although "rater" is not a required input, it is recommended for the purposes of evaluating inter- and intra-rater reliability.

Startup		
Video Title	Task Description ?	
GPWS Study (KC-135)	KC-135 Flight study looking at fine motor manipulation and object handling inside Glovebox mockup.	
Camera Location		
Front Cabin		
Rater	Subject Code	Other
Kent	F1	Flight Day 1 (March 1993)
Body Orientation	Body Movement & Rating Level	Behaviors & Activities
Neutral Body	1 Flex/Extend	A-I K-R S-Z Other
Body Part	Mild - Severe	Twist
Neck	2	
Camera View		
Left Side		
		Clear Setup

Figure 1. PVAT Startup Screen

The primary inputs are: subject code, body orientation, body part, camera view, body movement, and rating level. Body orientation describes the overall posture most often observed in microgravity, see Figure 2. Body part (e.g., Neck) may be selected as an additional descriptor variable. Having specified the orientation and body part of interest, the user then selects a body movement category (e.g., flex/extend) and its corresponding rating level (e.g., mild-severe). The midpoint represents "nominal." Sometimes, it is possible for the user to rate different movements simultaneously (e.g., supine/prone and flex/extend) given the same body part (e.g., hand). The camera view menu selection input is useful in describing the camera position at the time of video recording. There are six view options listed in the camera view pulldown menu (top, bottom, front, left side, right side, angle) as well as an "other" option. The importance of camera view entry is that certain movements are better observed and rated given different camera views. Such information could be then used to request multiple camera views in the future. The lower right hand portion of Figure 1 contains three buttons grouped alphabetically. These buttons are associated with the "behavior/ activity" input category. This category can be used as an independent measure or to provide supplementary information in documenting "behaviors" or "activities" associated with the primary input.

Body Orientation	Body Part
Free Floating Horizontal Neutral Body Reclined Sitting Standing Transverse Vertical	Ankle Knee Hip Low Back Neck Shoulder Elbow/Wrist Hand Fingers
Clear Other...	Clear Other...

Figure 2. Body Orientation and Body Part

Once, these entries are selected, the user presses the "Setup" button to initialize the program and to open the "Output" screen, as shown in Figure 3. Once the setup is completed, the user may begin analyzing the selected video footage using PVAT. Entries are made using labeled screen buttons. Each posture classification is stored and time stamped automatically. Provisions are made that allow the users to pause, tag incorrect selections, enter an "Unsure" response and "Notes." Having completed the assessment process, the analyst can summarize the data in terms of posture

classification frequency and overall duration. The data is saved as a "text" file with tab delimiters that can be imported into programs such as Microsoft EXCEL™.

Output			
Body Orientation	Body Part	Camera View	Subject Code
Neutral Body	Neck	Left Side	F1
Press Start			
1:41:17	Flex Mild		↑
1:41:28	Extend Severe , Twist		↓
1:41:37	Nominal		↓
Flex <input type="radio"/> Mild <input type="radio"/> Severe		Extend <input type="radio"/> Mild <input type="radio"/> Severe	
		<input type="button" value="Nominal"/> <input type="button" value="Unsure"/> <input type="button" value="Clear Last"/>	
Behaviors/ Activities	<input type="text" value="Twist"/>	<input type="text"/>	<input type="text"/>
	<input type="text"/>	<input type="text"/>	<input type="text"/>
Start	<input type="text" value="1:41:11"/>	End	<input type="text" value="1:42:11"/>
		<input type="button" value="Pause"/>	<input type="button" value="Notes"/>
		<input type="button" value="Clear ALL"/>	<input type="button" value="Start"/>

Figure 3. PVAT Output Screen

USER EVALUATION/CASE STUDY

A user evaluation was conducted to identify interface issues and required modifications to the software. Two raters were trained to test PVAT using video of the General Purpose Workstation (GPWS) taken on NASA's KC-135 aircraft. Many of the users' recommendations were incorporated into the screen designs previously shown in Figures 1 and 3 (e.g., button locations, automatic "save" prompt, simplified screen layout, smaller output field, and flagging mistaken entries for deletion with an "*"). Future software modification also will include the addition of a "Restraint Systems" input parameter, internal data reduction and calculation capabilities, an animated posture classification glossary, and automatic control of the video deck from within the software.

In conducting the video analysis of the GPWS, the two trained PVAT raters were instructed to classify the neck posture as: *Nominal*, if the neck appeared to deviate 0-20° from a reference plane; *Mild flex/extend*, if the deviation was between 20-

45°; and *Severe flex/extend*, if the deviation were greater than 45° (taken from Leonard and Keyserling, 1989). The neck was targeted for this evaluation since it was the body part most affected by the GPWS design. The purpose for conducting the KC-135 studies was to evaluate and document the microgravity posture for the 5th percentile female and 95th percentile male subjects.

It was evident from the analysis that there were differences in neck flexion between the 95th percentile male and the 5th percentile female subjects. Males were found to have more severe neck flexion (12%) versus females (1%). This was attributed to the design limitations of the GPWS where taller operators tend to bend and hover their upper body in order to look inside the glove box. The GPWS configuration is constrained in that the gloveports and workstation height are not adjustable. Data of this nature is needed to support design and operational recommendations from a human performance standpoint.

Application of PVAT to Other Industries

Industries such as manufacturing plants, offices, laboratories, and service industries such as postal or food services typically require workers to maintain prolonged postures. The posture may involve sitting, standing, or both sitting and standing to perform the work. A person should not be restricted to a workplace in such a way that he or she cannot change posture during the work shift. If a worker can achieve an extended reach only by leaning, stretching, stooping or crouching; these postures can all produce fatigue.

In addition, some work environments may require the use of specialized equipment, such as computers, chemical hoods, or glove boxes. This equipment may place constraints on the operator such as restricted arm movement or vision, which in turn may cause discomfort or fatigue and make the task awkward to perform.

PVAT can assist engineers in designing tasks and the equipment to be used. PVAT could also assist in evaluating existing workplace environments to determine problem areas resulting from posture demands of the task.

Many health professionals are concerned with the posture an individual may assume. Occupational and physical therapists and nursing professionals are aware that prolonged or awkward posture may fatigue and/or result in physical complications for a patient. These professionals could use PVAT to study postures that patients use to perform specified tasks for the home or work. It may also be used to assess the effectiveness of various sedentary positions for recovering patients.

Alternative postures could be identified, thereby allowing recovery patient with some postural flexibility.

Such evaluations may also result in the identification of aids to assist the individual in performing tasks or attaining sedentary comfort. Useful aids might be items such as a specially designed chair, foot or back support, or tools for performing a task, to name a few.

CONCLUSIONS

Posture is an important aspect of workload and a potential limitation on time and/or effectiveness of an operator's performance (Corlett, Madeley and Manenica, 1979). PVAT provides the initial step in identifying "limiting postures" and related workstation design concerns. Furthermore, it may provide the analyst with supporting data to specify "adequate" postures or "safe" durations for certain postures perceived as potentially dangerous.

Commercial industries can benefit from the PVAT concept, with little or no modification to the software. Potential applications would include industries: with limited resources for conducting comprehensive ergonomics studies; where video has already been used for monitoring work-in-progress; and where posture related injuries are common occurrences. It can contribute to various aspects of workplace design such as training, task allocations, procedural analysis, scheduling of timelines, and hardware usability evaluations.

Future plans are to finalize the design and to make it available for human factors engineers.

ACKNOWLEDGMENTS The authors wish to recognize the work conducted by Darlene Merced-Moore in helping to conceive, develop, and test PVAT.

REFERENCES

1. Corlett, E. N., Madeley, S. J. and Manenica, I. (1979). *Posture Targeting: A Technique for Recording Working Postures*. Ergonomics, 22(3), 357-366.
2. Fisher, W. and Tarbutt, V. (1988). *Some Issues in Collecting Data on Working Postures*. Proceedings of the Human Factors Society 32nd Annual Meeting. Santa Monica, CA: Human Factors Society.
3. Grandjean, E. and Hunting, W. (1977). *Ergonomics of Posture - Review of Various Problems of Standing and Sitting Posture*. Applied Ergonomics, 8(3), 135-140.
4. Karhu, O., Kansil, P. and Kuorinka, I. (1977). *Correcting Working Postures in Industry: A Practical Method for Analysis*. Applied Ergonomics, 8, 199-201.

5. Leonard, J. and Keyserling, W. M. (1989). *A Method to Evaluate Neck and Lower Extremity Postures Using Simulated Real Time Analysis*. In Mital, A. (Ed.), Advances in Industrial Ergonomics and Safety I. Great Britain: Taylor and Francis.
6. Priel, V. (1974). *A Numerical Definition of Posture*. Human Factors, 16, 576-584.

APPLICATION REUSE LIBRARY FOR SOFTWARE, REQUIREMENTS, AND GUIDELINES

Jane T. Malin
NASA
Intelligent Systems Branch
Johnson Space Center, ER2
Houston, TX

520-61

Carroll Thronsbery
MITRE
Johnson Space Center, ER2
Houston, TX

ABSTRACT

Better designs are needed for expert systems and other operations automation software, for more reliable, usable and effective human support. A prototype computer-aided Application Reuse Library shows feasibility of supporting concurrent development and improvement of advanced software by users, analysts, software developers, and human-computer interaction experts. Such a library expedites development of quality software, by providing working, documented examples, which support understanding, modification and reuse of requirements as well as code. It explicitly documents and implicitly embodies design guidelines, standards and conventions. The Application Reuse Library provides application modules with Demo-and-Tester elements. Developers and users can evaluate applicability of a library module and test modifications, by running it interactively. Sub-modules provide application code and displays and controls. The library supports software modification and reuse, by providing alternative versions of application and display functionality. Information about human support and display requirements is provided, so that modifications will conform to guidelines. The library supports entry of new application modules from developers throughout an organization. Example library modules include a timer, some buttons and special fonts, and a real-time data interface program. The library prototype is implemented in the object-oriented G2 environment for developing real-time expert systems.

INTRODUCTION

In current and future programs, control center users (flight controllers) will continue to require increasing involvement in the design of software applications. Flight controllers are taking responsibility for design of automation software in the Space Shuttle and Space Station programs, especially advanced displays and expert systems. Such software is being developed to make operations safer, more reliable and less costly.

Yet, other organizational elements of programs will remain responsible for a number of software development tasks, with goals such as efficient performance and sustaining maintenance, and conformance to human factors guidelines. Software development and integration processes designed to support more user involvement often seem, paradoxically, to make it more difficult to develop and integrate the advanced user software. Traditionally, these processes have relied heavily on documents, forms and committees. Users and developers have not been able to use these methods effectively. User understanding of tools for developing advanced displays and expert systems is still maturing. In this environment, more support is needed for successful development and integration of advanced user applications.

The roles of human factors personnel in software development and certification have been frustrating and expensive. Establishing guidelines and standards is only a first step. Users and designers need design support, not just enforcement and evaluation activities. After style guides and standards have been agreed to, they need implementation support, to be effectively utilized. The input of human factors experts needs to be transformed to be more closely allied to design and test. One way is to provide working examples that embed standards and guidelines in applications. In the areas of advanced displays and expert systems, such examples would be welcomed by users and developers.

The concept of reusable software can be applied to the process of software design by users, and this technology can be used to support user application development and to embed human factors guidelines. Reuse and modification of working documented examples will result in more productive user involvement in design of displays and automation software. It will result in faster prototyping and development, and more standard software. It will speed certification, by supporting code review by difference and building of applications from certified parts. It will also support continuous improvement of software requirements, design and code, especially for automation software.

Designers of processes, documents and tools for software reuse must beware, however. There are two significant pitfalls. The first pitfall to avoid is a high-cost retrofitting process that uses computer-aided software engineering (CASE) maintenance and metrics tools for reuse. Iterative development with user participation is fundamentally different from traditional software engineering, and it requires different processes and support tools [3]. The second pitfall to avoid is failing to solve critical reuse problems by redesign of processes and programming approaches alone. Experience has already shown that highly successful process redesign for reuse may significantly reduce defects, but may not significantly improve software development in a number of critical areas that concern users and human factors personnel [1]. The areas most resistant to this approach pertain to application of standards, and to misunderstandings of functions of reusable library items, sometimes called item relevance.

We need a process and tools that would support innovation and use of innovative software, and would support iterative prototyping with heavy user involvement. Given the disappointing track record of CASE-tool-based re-engineering, and the failure of reuse process changes to impact many critical reuse problems, we have taken a completely different approach. In this approach, the key innovation is interactive demonstration of displays and applications, to support evaluation of relevance of library items. Such a software reuse library is strong in the areas where traditional CASE is weakest, in requirements definition and prototyping support. The emphasis is on providing support for demonstrating and prototyping applications and displays, along with their corresponding requirements and guidelines.

Our design concept for an Application Reuse Library as an application designer's tool includes several key features:

1. Item relevance judgments are supported by a combination of descriptive information and a test-drive demonstration of the item and closely related alternatives. Users can inspect available items to improve requirements prior to coding, and can even use the library items to help develop requirements by difference. Test drives are easiest when a user interface (UI) is provided, but a clear distinction is maintained between the UI elements and the application elements of an item. Key benefits of this feature are avoidance of misunderstood item functions, disregarded item limitations, and faulty functional specifications.

2. Items embody standards and contain descriptive information to call attention to the embodied standards. The items provide examples of standards integrated into applications, and can make the standards less abstract. The items make application of standards another benefit of code reuse. Key benefits of this feature are avoidance of ignored or misapplied standards.

3. Indexing of items is flexible, to support an evolving library of items that are complex and volatile. The library needs to accommodate new items as they are extracted from new applications, and to reflect changes and variations in requirements and innovations in meeting requirements. To implement this feature, a fixed database design must be avoided. Some modern object-oriented databases have the desired flexibility.

4. The size of the library is not likely to be so massive as to require sophisticated search. The domain of real-time monitoring and fault management in control centers is limited enough to make it likely that there will be easy requirements matches. Most of the relevant intelligent systems applications of this type for Johnson Space Center ground control have been described [4].

5. Library item development support focuses on partitioning complete applications and selecting items opportunistically from complete applications. Extracting library items from current complete applications has two benefits: avoiding items that have not stood the test of integration and may not meet any specific current requirement, and avoiding bogging down library development by defining a complete set of new items to be developed. Selection of items can be based on functional importance rather than on complete coverage. Functionally important items in a domain can be recognized by their benefit (and difficulty to develop), their common use, and their obvious relevance [2].

Our current prototype emphasizes some of the features of this design concept, especially the support for judging item relevance and embodying standards.

PROTOTYPE

The Application Reuse Library prototype contains a variety of types of library items to illustrate broad usefulness, and to explore variations in the use of library features. All the items were extracted from existing intelligent system applications. The library items, or modules, contain Demo-and-Tester elements. Developers and users can evaluate applicability of a library module and can test modifications, by running the Demo-and-Tester interactively. Sub-modules provide application code and displays and controls. The library prototype is implemented in the object-oriented G2 environment for developing real-time expert systems. Current Library modules include some buttons and special fonts, a real-time data interface program, a mission header, an X-based event logger, an X-based plotter, a timer, and a time-conversion module. The Timer will be used to illustrate the library design concept in this paper.

Developer's Interface and Demo-and-Tester

When the developer inspects the Timer for possible reuse in a new application, the display shown in Figure 1 appears. The buttons at the top allow navigation to additional timer information. The text helps the user determine the timer's relevance to the new application. The Timer Demo-and-Tester is shown in Figure 2. By manipulating the inputs, the user tests the application in an interactive demonstration to help determine the relevance of this type of stopwatch timer. This Demo-and-Tester can be reused itself, to evaluate any modifications of the item in the new application. The inputs and outputs illustrate the

TIMER INFORMATION

✕

DEMO-AND-TESTER
 USER-INTERFACE-OPTIONS
 IMPROVEMENT-IDEAS

CODE
 KB-INTEGRATION

PURPOSE:
 The TIMER functions much like a stopwatch. Your software can start, stop, and resume the timer and set a timeout limit. Afterwards, it can read the timer-count (in seconds), the state (on or off) of the timer, and the status (nominal or timed-out) of the timer.

WARNINGS:
 This timer "ticks" once per second within G2. Depending on your operating system and the workload of your workstation, this may not be as accurate as comparing telemetry event time-stamps for start and stop times.

NEEDED IMPROVEMENTS:
 The timeout limit should be made an integral part (attribute) of the timer object. Currently, it is implemented as a separate variable. An example of this improvement appears in the TIMER-IMPROVEMENT-IDEAS workspace.

EXISTING IMPLEMENTATIONS:
 This timer is being used in PDRS-DESSY, both the MPM-MRL module and the End Effector module.

KEYWORDS:
 timer, stopwatch, event timer

AUTHOR:
 Sherry Land, ER2, 483-2064, land@aio.jsc.nasa.gov
 Carroll Thronesbery, PS/MITRE, 335-8533, cthrones@mitre.org
DATE OF ENTRY:
 2 July 1993

Figure 1. The top-level display panel for the Timer library item.

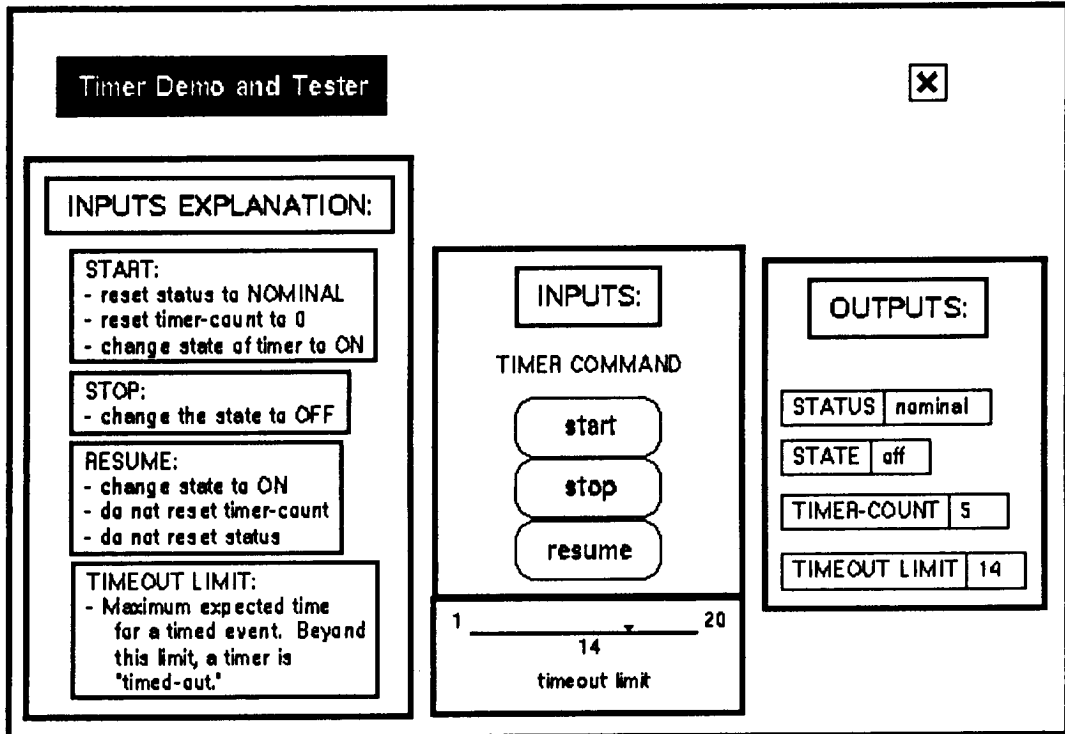


Figure 2. The Demo-and-Tester panel for the Timer library item.

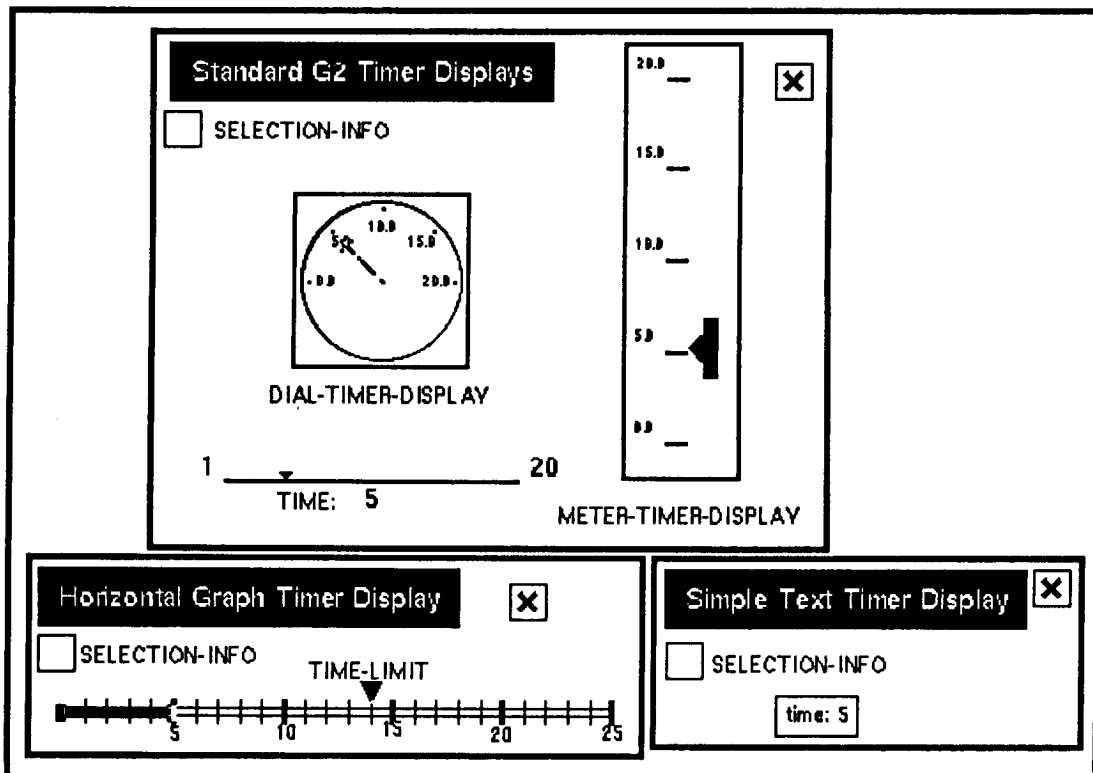


Figure 3. Alternative user interface display options for the Timer library item.

application level of the module, not the UI. Inputs would be set by the application, and outputs would be fed directly to the application or to the UI.

By selecting User-Interface-Options, the user can interactively test-drive a variety of timer display options in parallel, shown in Figure 3. All timer displays update dynamically to reflect changes caused by inputs to the Demo-and-Tester. Again, this helps determine the relevance of the displays and of this type of timer. Providing a set of options that can be run in parallel benefits the evaluation process, as patterns of common and different features can be better perceived in the simultaneous test-drive. This helps the user understand both functionality and limitations, and the distinctions that should be made in functional requirements.

On the basis of this parallel test-drive, the user can refine requirements. The clear distinction between the application and the UI helps the user sort out both types of requirements simultaneously. This is a great benefit, since both types of requirements interact, and can be handled better together.

Application of Standards

For each display option, there is descriptive display selection information, shown for the Horizontal Graph Timer display in Figure 4. Descriptions of display advantages, disadvantages, and characteristics compliant with guidelines can be provided by human factors engineers. Each display that illustrates guidelines helps bridge the communication gap between guidelines writers and software developers. Explicit information about compliance with guidelines helps developers stay in compliance when the reused code is altered for integration into the the new application.

As participants in Library development process, human factors engineers can also write in descriptions of improvement ideas. They can also modify library item sub-modules to illustrate improved application of standards, and submit these items to the Library. Improved items would be alternatives rather than replacements for old items, until the new item is integrated into the old application. Submission of entirely new items to the Library would be a less desirable human factors activity, since such items have not stood the test of integration, and might not actually fit into any real application.

Importing, Integrating, and Altering Code for New Applications

When a user has determined that a library item can be reused in a new application, the code can be copied into the developer's working files. The code can include several types of sub-modules: the application code, the UI code, example integration formats, and the Demo-and-Tester. Integration examples must be altered for the new application. The requirements of the new application are likely to be slightly different, and the application or UI sub-modules will need to be altered. Testing of the changed code is made easy, since the Demo-and-Tester can be altered for this purpose.

When an item is reused, even in altered form, certification and documentation can be made much easier. Code review and certification can concentrate on the differences, since the Library item has already been reviewed and certified.

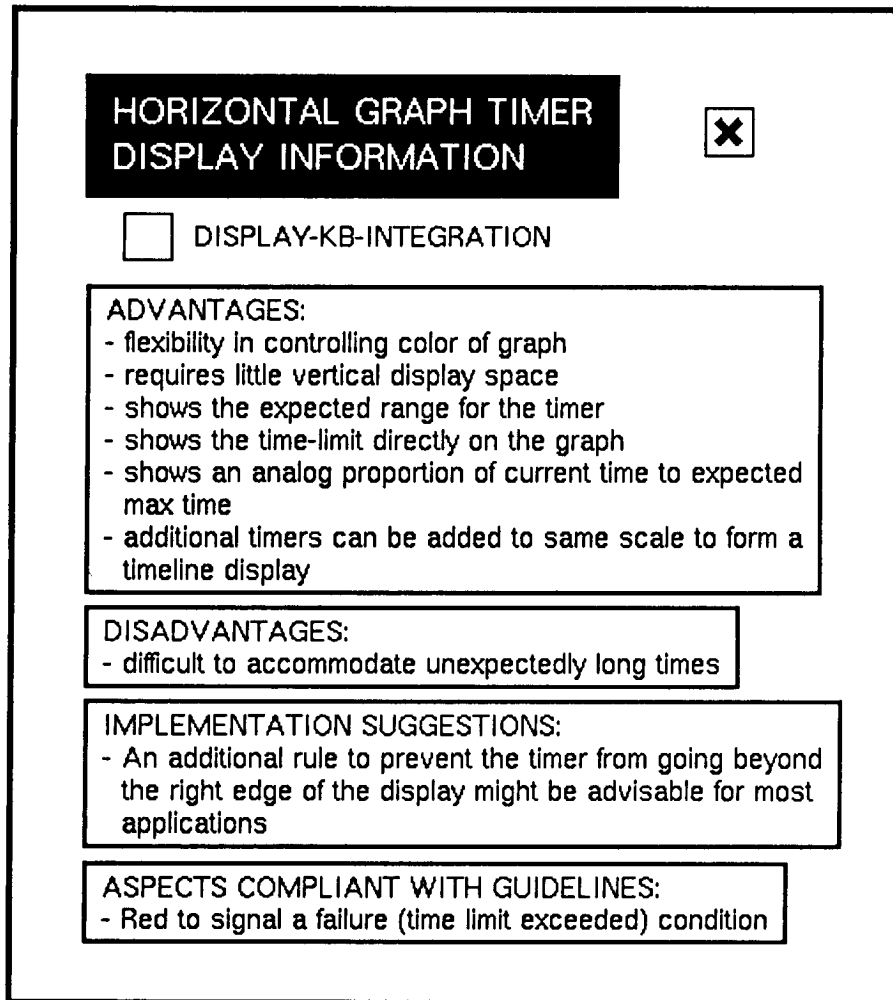


Figure 4. Selection information for the Horizontal Graph display option for the Timer item.

Developing New and Altered Library Items

All the current Library modules have been extracted from intelligent system applications. A process for developing really new items need not exist, beyond the normal process of developing applications. Because development so far has been limited and exploratory, we have not yet systematically partitioned an entire application. Items that seem powerful and commonly reusable have been entered as opportunities have arisen.

The Library provides instructions and forms to make it easy to add descriptive documentation. There are not yet any indexing capabilities beyond space in the description for providing keywords. There is not yet any configuration management support.

The process of entering code for an item can be easy if the code is modular and has undergone some unit testing. The unit test can be the basis for developing the Demo-and-Tester. If the item is innovative, developing a set of alternatives can be easy, since the innovative items often improve on simpler functions. The alternative timer displays illustrate this concept: three are standard displays that are part of the G2 commercial software, one is a simple extension, and one is a complex extension.

CONCLUSIONS

This reusable software library concept for human factors engineering and requirements analysis is unique. Nevertheless, it should be highly transferable to software development in industry. It should be especially useful in organizations where continuous improvement of working software is needed in a large project, and where more productive human factors and user involvement is sought.

Plans for future work include a Design Assistant and Library (DAL) to support, automate, and reduce costs of development of automation applications for ground monitoring and control. We plan to further support and automate the process with an intelligent design assistant (IDA) for application development, integration, test, porting and documentation. The objective is to use the DAL to assist users, including control center users and science Principal Investigators (PIs), in developing, testing and integrating application software and its functional specifications, by providing resources and support that would otherwise require a multi-disciplinary development team.

The DAL concept is based on the successful Application Software Factory developed by Celite Corporation [1]. Our goal is to reproduce the benefits experienced by Celite, and to drastically cut costs to develop certified applications. Benefits observed in the Celite Corporation case include reduction of development team personnel by 75%, and significant reductions in application development time, rework, software defects and training time.

ACKNOWLEDGEMENTS

We want to thank the many individuals who have contributed code and concepts to our Library prototype, especially Sherry Land. Other code contributors include Jane Falgout, Art Rasmussen and Mark Gnabasik.

REFERENCES

1. Burton, S., Swanson, K., and Leonard, L., "Quality and Knowledge in Software Engineering," AI MAGAZINE, AAAI, Vol. 14, No. 4, Winter, 1993, pp. 43-50.
2. Davis, R. and Resnick, P., "Multiple Dimensions of Generalization in Model-Based Troubleshooting," PROC. 11TH NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE, AAAI Press, Menlo Park, CA, 1993, pp. 160-167.
3. Malin, J., Land, S., and Thronesbery, C., "End Effector Monitoring System: An Illustrated Case of Operational Prototyping," PROC. 7TH ANNUAL WORKSHOP ON SPACE OPERATIONS APPLICATIONS AND RESEARCH (SOAR '93), Houston, TX, August, 1993.
4. Malin, J., Schreckenghost, D., et al., "Making Intelligent Systems Team Players." Vols. 1-2, 3, 4, NASA TECHNICAL MEMORANDA 104738, 104754, 104786, Houston, TX, 1991-4.

omit

Session M1: MATERIALS AND STRUCTURES I

Session Chair: David Hamilton

NASA's Man-Systems Integration Standards:
A Human Factors Engineering Standard for Everyone in the Nineties

Mr. Cletis R. Booher
NASA Johnson Space Center
Mail Code - SP33, Houston, TX 77058

321-54

Dr. Betty S. Goldsberry
Lockheed Engineering and Sciences Company
Mail Code - C44, Houston, TX 77058

100 /

Introduction

During the second half of the 1980s, a document was created by the National Aeronautics and Space Administration (NASA) to aid in the application of good human factors engineering and human interface practices to the design and development of hardware and systems for use in all United States manned space flight programs. This comprehensive document, known as NASA-STD-3000, the Man-Systems Integration Standards (MSIS), attempts to address, from a human factors engineering / human interface standpoint, all of the various types of equipment with which manned space flight crew members must deal.

Basically, all of the human interface situations addressed in the MSIS are present in terrestrially based systems also. The premise of this paper is that, starting with this already created standard, comprehensive documents addressing human factors engineering and human interface concerns could be developed to aid in the design of almost any type of equipment or system which humans interface with in any terrestrial environment. Utilizing the systems and processes currently in place in the MSIS Development Facility at the Johnson Space Center in Houston, TX, any number of MSIS volumes addressing the human factors / human interface needs of any terrestrially based (or, for that matter, airborne) system could be created.

Background

An obvious need has existed from the earliest days of the U.S. manned space flight programs for a comprehensive and cohesive standard to facilitate the implementation of human factors and crew interface guidelines into these programs. Until early 1987, no such document existed. Some of the NASA centers, the Johnson Space Center (JSC) in Houston, TX, and the Marshall Space Flight Center (MSFC) in Huntsville, AL, specifically, had, through the years, developed "guidelines" and "specification" documents, which individually covered relatively limited portions of the overall human interface arenas. Prior to 1987, however, there had never been anything even close to a comprehensive set of guidelines to which a space system designer could go for help in the determination of what special considerations had to be

accorded to the design of any given space-related system or piece of hardware with which the crew member would be interfacing.

This situation changed dramatically, however, in March of 1987 with the release and distribution of NASA-STD-3000, the Man-Systems Integration Standards. Put together by the space human factors department of Boeing Aerospace Co., with the very able and comprehensive assistance of human factors experts with space-related experience from across NASA, the aerospace industry, and academia, this document embodied, at the time of its publication, the most complete known body of data which included the requirements for the design of space-related hardware with which a crew-member must interact. Since publication of this initial version of the MSIS, an on-going effort has been sustained, as funding has allowed, to keep this standard updated and relevant to proposed NASA manned space programs. Revision A of the standard was released in October of 1989, and, plans call for a B Revision to be released some time in 1994.

Document's Uniqueness/Value

We believe that Volume I of NASA-STD-3000 is the most complete compilation of human factors requirements related specifically to the design of equipment and systems, albeit space-related systems, in existence today. This document has been created specifically for the designer, rather than for the human factors or ergonomics expert. In addition to addressing numerous specific system requirements, the MSIS also includes comprehensive information on anthropometry and biomechanics, human performance characteristics, and natural and induced environments, to give designers sufficient insight into the effects which these additional factors will exert on the human interfacing with the system. These additional influencing factors must then be taken into account by the system designer in order for him to create the most effective and efficient system designs.

There is no intent in the offering of MSIS-based documents to the human interface design community to try to replicate or replace the information contained in any of the current, well known and well-written human factors documents, such as Boff and Lincoln's 1988 Engineering Data Compendium addressing Human Perception and Performance, or Woodson's 1981 Human Factors Design Handbook. Indeed, a lot of the information contained in the MSIS was derived from these types of documents. A lot of the MSIS data was also derived from military publications related to Human Factors Engineering concerns, such as MIL-STD-1472, Human Engineering Design Criteria for Military Systems, and MIL-HDBK-759, Human Factors Engineering Design for Army Material. There does exist a definite degree of similiarity between these publications and the MSIS. As in the case of the Military Standards, the MSIS documents are continuously being

updated and improved. Even more unique, as with the "space program specific" volumes of the MSIS, the system (or individual industry) specific volumes of the standard would be created and maintained individually for each area of interest, with no "extraneous baggage" of non-needed information or requirements. Specific examples of the modes in which information is currently presented in the MSIS are shown in Figures 1., 2., and 3.

Methodology

Utilizing computer-based databases and publishing systems, the entire standards updating process has been automated, from the initial input of suggested changes, to the creation of the new pages of text and/or graphics. This includes the documenting of all comments and/or suggestions which are made by reviewers and other interested parties relative to a given suggested change, and careful, specific wording of the final version of the change which is to be inserted into the document. This documenting process provides a useful record of the inputs made by reviewers of the changes, which may be referred to at a later time, when questions concerning the rationale behind the insertion of any particular change into the document is raised.

Another singularly important aspect of the document updating process relates to the careful documentation of the sources of all changes which are proposed for inclusion in the standard. As with all standards types of documents, it is extremely important to provide a pedigree, in one way or the other, for all information which exists in the document.

Appropriate data regarding all recipients of the standards, including names, affiliations, full mailing addresses, and phone numbers, is also kept in a database. This allows easy forwarding of MSIS updates and revisions to all current holders of the documents. Additionally, during 1992, a newsletter was initiated and sent to all standards recipients.

In addition to the basic "generic" version of the MSIS initially created by Boeing, as new manned space programs are conceived and developed, additional volumes of the MSIS are created and maintained which specifically address the human factors and crew interface needs for that program. The MSIS volume which specifically addressed these needs for Space Station Freedom (SSF) was actually approved and released in January of 1987, two months prior to the release of the basic volume, and a volume which addressed these needs for the Assured Crew Return Vehicle, the SSF "lifeboat", was released in June of 1992. As specialized volumes of this type are updated and revised, the information gathered for them is also evaluated for possible inclusion in the basic volume.

Proposed Applications

Basically, all of the human interface requirements developed for application to space-related systems are also directly applicable to terrestrial systems. Not all of the MSIS requirements would be applicable to any individual, earth-based system; but taken as a whole, most of the requirements would find use in a terrestrial system somewhere. The only specific exceptions would be those requirements applicable to, or which are derived as a result of interfacing with, the micro-gravity (or zero-gravity) environment. Even the requirements for many of the Extravehicular Activity systems will be applicable to the development of equipment for the handling of hazardous materials or for use in hazardous environments. It is also probably appropriate to note at this point, that most of the requirements included in the MSIS were, indeed, derived from data extracted from testing in the terrestrial environment, with appropriate interpretation thereof.

Systems which could potentially benefit from the custom development of MSIS documents are many and varied. The design work necessary for anything from automobiles and recreational vehicles to pocket radios and video cassette recorders could benefit from this type of document. The design processes of all home appliances and home and commercial electronics hardware are candidates for the development of a document of this type. The design techniques used in the creation of all furniture products can potentially be improved through the use of this type of standard, as can the design of homes and places of business. This is not to say that there are not currently thousands of skilled designers and architects in all of these fields producing products with excellent ease of use characteristics; there are. But with a document of the type of the MSIS, this design process could be made considerably more efficient, consistent, and cost-effective in its output, while resulting in much more user-friendly end-products. Additionally, MSIS volumes, either generic or specific, could be developed to address the concerns of specific terrestrial environments, such as: ocean surface, subsea, mountainous, desert, subterranean, etc.

The first step in the development of NASA-STD-3000 types of documents for terrestrial application is to secure the proper legal clearances to enable the production of this type of document for other than federal government entities, since the standard is, at present, federally-owned, and thus available free of cost to any other organization or individual in its current configuration. Appropriate procedures will then be established whereby any professional or industry organization, company, or individual can come to NASA JSC and contract to have a version of the MSIS produced to fulfill their specific needs. These terrestrially-oriented volumes of the document would be created utilizing the equipment and processes developed for creation of the program-unique volumes discussed previously. See Appendix A for a generic outline of the work process for the development of a new MSIS volume.

Figure 4 lists all chapters of the MSIS, along with a few words relative to how the contents of each of these chapters can be utilized to incorporate human factors principles into the design of terrestrially-based systems.

Conclusion

As the vehicles, appliances, electronic equipment, and all other items which we buy or come into contact with, become more sophisticated and complex, the jokes about not being able to program our own Video Cassette Recorders or not understanding all of the buttons and controls in our new vehicles become more and more true. These situations can not only become frustrating and embarrassing, but, in the cases of certain systems, down-right dangerous. The insertion of some good, rational, human factors engineering and logic into the design and development phases of items of the types listed above, not to mention simpler units like furniture, ladders, and yard equipment, would go a long way to making all of our time, both at work and during our leisure hours, much more efficient and enjoyable.

APPENDIX A
New MSIS Volume Development Process

Following identification of a specific product, organization, or industry for which an MSIS Volume is to be developed, a process generally adhering to the following steps would be developed:

- I. Identification of a cadre of experts who would be available to identify the specific sections in the current generic MSIS volume which would be applicable to the new "specialized" volume. This might well include the identification of areas not currently covered in the generic volume. This same group of experts would also be expected to critique sections of these new, specialized volumes as they are created and/or extracted from the existing generic MSIS volume.
- II. Call a meeting of these experts to develop a Table of Contents (TOC) and/or outline of the new specialized volume.
- III. MSIS contractor personnel would create an appropriate TOC and/or outline as directed by the experts and return the draft TOC and/or outline to same. Meetings of the experts would be held at intervals, as felt appropriate by the sponsoring entity and/or the experts, as funds allowed.
- IV. Continue an iterative process of evaluation, comment, and inputting by the experts of new information, concepts, and/or data relating to the specific MSIS volume being created; and creation and updating of appropriate sections of this new MSIS volume by the MSIS contractor, until all parties agree that the new, specialized volume is as complete as appropriate for its intended use.
- V. Completion by the MSIS contractor of the final draft of the new, specialized volume for perusal and approval by the experts.
- VI. Following final approval by the experts, publication and distribution of the final version of the specialized MSIS volume by the MSIS contractor. The list of recipients of this distribution must be approved by the sponsoring entity of the new volume.

c. Exposed edges 0.5 to 3.0 mm (0.02 to 0.12 in.) thick shall be rounded to a full radius. See Figure 6.3.3-3.

d. The edges of thin sheets less than 0.5 mm (0.02 in.) thick shall be rolled or curled. See Figure 6.3.3-4.

6.3.3.2 Exposed Corner Requirements for Facilities and Mounted Hardware

a. Exposed corners of materials less than 25 mm (1.0 in.) thick shall be rounded to a minimum radius of 13 mm (0.5 in.). See Figure 6.3.3-5.

b. Exposed corners of materials which exceed 25 mm (1.0 in.) in thickness shall be rounded to 13 mm (0.5 in.) spherical radius. See Figure 6.3.3-6.

6.3.3.3 Protective Covers

Equipment requiring pointed, sharp, or dangerous edges shall be covered or shielded when not in use.

6.3.3.4 Holes

Holes that are uncovered and are round or slotted in the range of 10.0 to 25.0 mm (0.4 to 1.0 in.) shall be avoided.

6.3.3.5 Latches

Latches or similar devices which can pinch fingers shall not be used. A protective guard or cover shall be used where suitable substitutes cannot be found.

6.3.3.6 Screws and Bolts

Screws or bolts with more than two exposed threads shall be capped to protect against the sharp threads.

6.3.3.7 Securing Pins

Securing pins in hand rails shall be designed to prevent their inadvertently backing out above the handhold surface.

6.3.3.8 Levers, Cranks, Hooks, and Controls

Levers, cranks, hooks, and controls shall not be located where they can pinch, snag, or cut the crewmembers or their clothing.

h. Rapid Closing - Hatches used to isolate interior areas of Space Station Freedom shall be designed to allow rapid closing.

8.10.3.4 Operating Forces Design Requirements

Hatch and door cover operating forces shall meet the following requirements:

- a. Emergency Operation - Forces for emergency backup operation or breakaway of jammed internal doors shall not exceed 445 Newtons (100 lbf).
- b. Latch Operations - The force required to operate the door and hatch latches shall not exceed the strength of the fifth percentile design population as defined in Paragraph 4.9.3.
- c. Open/Close Force - The opening and closing forces for internal hatches and doors shall not exceed 22 Newtons (5 lbf) with zero delta-pressure through the opening.
- d. Restraints - Restraints shall be provided as necessary to counteract body movement when opening or closing hatches and doors.

8.10.3.5 Minimum Size Design Requirements

The minimum size of personnel hatch and door openings shall accommodate passage of the largest IVA ORU or crewmember (whichever is larger) intended to pass through the opening.

8.10.3.6 Operations Interface Requirements

The location and operation of crew interfaces (gauges, levers, valves, handles, etc.) for hatches in all pressurized elements shall be visually and functionally identical. This shall include the procedures and protocols for opening, securing, closing, stusing and performing maintenance.

8.10.3.7 Shape

The hatch should be shaped such that it can pass through the opening that it is designed to seal to allow for removal, maintenance, repair, relocation, etc.

8.11 WINDOWS INTEGRATION

8.11.1 INTRODUCTION - N/A

8.11.2 WINDOWS INTEGRATION DESIGN CONSIDERATIONS - N/A

9.3.3.4.1.1 Layout

a. **Alphanumerics** - The basic alphanumeric character arrangement for standard keyboards shall conform to USA Standard Typewriter Pairing of the American Standard Code for Information Interchange (ASCII). See Figure 9.3.3.4.1.1-1.

b. **Reserved**

c. **Number Keypad** - When appropriate, a number keypad shall be added to the keyboard. This shall be to the right-hand side of the main keyboard, if workstation layout permits. The arrangement of the numeric keypad shall conform to Figure 9.3.3.4.1.1-2.

d. **Function Keys** - The use of function keys will depend on the specific system that the keyboard is a part of.

1. **Keying Process** - Function keys shall be used to make the keying process faster and to minimize keying errors where fast response is required (e.g., contingencies).

2. **Location of Function Keys** - Certain functions that occur most frequently or that tend to occur together should be placed in the same area.

3. Function Key Types

(a) **Fixed Function Keys** - Fixed function keys shall be provided for those functions that are widely and frequently used. Examples of commonly used fixed function keys are RESET, BREAK, TRANSMIT, CONTROL, and a means of cursor control.

(b) **Cursor Movement Keys** - Cursor movement keys shall be arranged in a spatial configuration reflecting the direction of actual cursor movement see Figure 9.3.3.4.1.1-3).

(c) **Variable Function Keys** - Variable function keys (user programmable) shall be provided whenever it is thought that the system will at present or in the future require the flexibility of these keys.

4. **Minimization of Errors** - The keyboard layout shall minimize the effect of likely errors, especially those that are critical. For instance, the delete key shall never be located next to the send key or other frequently used keys.

5. **Non-ASCII Key Locations** - The locations of keys which are not defined by the ASCII USA Standard Typewriter Pairing shall be located using the following guidelines.

(a) **Frequently Used Keys** - Frequently used keys shall be placed in the locations in which they are most convenient to use.

**Applicability of Space-Related MSIS Chapters to Terrestrially
(and Airborne) Related Designs**

CHAPT.#	CHAPTER TITLE	TERRESTRIAL RELATIONSHIP
1	Introduction	Document Development & Use Information
2	General Requirements	General Commonality Statements
3	Anthropometry and Biomechanics	Anthropometric and Biomechanical Data as Needed
4	Human Performance Capabilities	Utilize as Applicable to Specific Systems
5	Natural and Induced Environments	Utilize as Applicable to Specific System Environments
6	Crew Safety	Utilize Safety Reqmts. as Applicable to Specific Systems
7	Health Management	Partially Applicable to Health-Related Systems
8	Architecture	Utilize as Applicable to Specific Terrestrial Systems
9	Work Stations	Human Interface, Control, Display, & Computer System Interfaces as Applicable
10	Activity Centers	Living, Work, & Recreation Areas Descriptions as Required
11	Hardware and Equipment	Small Equipment and Hardware Descriptions as Required
12	Maintainability	Ease of Maintenance Requirements as Required
13	Facility Management	How to Keep Track of Parts and Components as Required
14	Extravehicular Activity	Partially Applicable to Systems to be Utilized in Hostile Environments

21 -

**Session H3: SYSTEMS/PROCESSES IN
HUMAN SUPPORT TECHNOLOGY**

Session Chair: John Schuessler



SOLAR PV POWERED HEAT PUMP	CREW AND THERMAL SYSTEMS DIVISION	
	Michael Ewert	Feb. 2, 1994

SOLAR PHOTOVOLTAIC POWERED HEAT PUMP

Michael Ewert
Jeff Dominick

DUAL-USE SPACE TECHNOLOGY TRANSFER CONFERENCE

NASA JOHNSON SPACE CENTER



SOLAR PV POWERED HEAT PUMP	CREW AND THERMAL SYSTEMS DIVISION	
	Michael Ewert	Feb. 2, 1994

HEAT PUMPS

- A heat pump is a device which moves or "pumps" heat from a cooler to a warmer region
- Your home air conditioner and refrigerator are heat pumps
- What is commonly referred to as a heat pump is basically an air conditioner which can also be run in reverse to pump heat into the house in winter
- Heat pumps require a considerable amount of energy to operate
 - This energy is used to compress vapor refrigerant in a traditional vapor compression heat pump



SOLAR PV POWERED HEAT PUMP	<i>CREW AND THERMAL SYSTEMS DIVISION</i>	
	Michael Ewert	Feb. 2, 1994

SOLAR PHOTOVOLTAIC POWER

- Solar photovoltaic (PV) panels produce direct current (DC) electricity from sunlight due to the properties of two different semi-conductor layers
- PV panels have been commercialized for small scale power production since the first practical panels flew on NASA satellites in the 1950s
- The cost is still high, but great advances have been made
- In recent years, electric utilities have showed great interest in using PV for peak power production
 - Several demonstration plants have been built
 - Southern US utilities are particularly interested since their system peak demand is on hot sunny days when the most sun is available



SOLAR PV POWERED HEAT PUMP	<i>CREW AND THERMAL SYSTEMS DIVISION</i>	
	Michael Ewert	Feb. 2, 1994

NASA'S INTEREST IN HEAT PUMPS AND SOLAR PV POWER

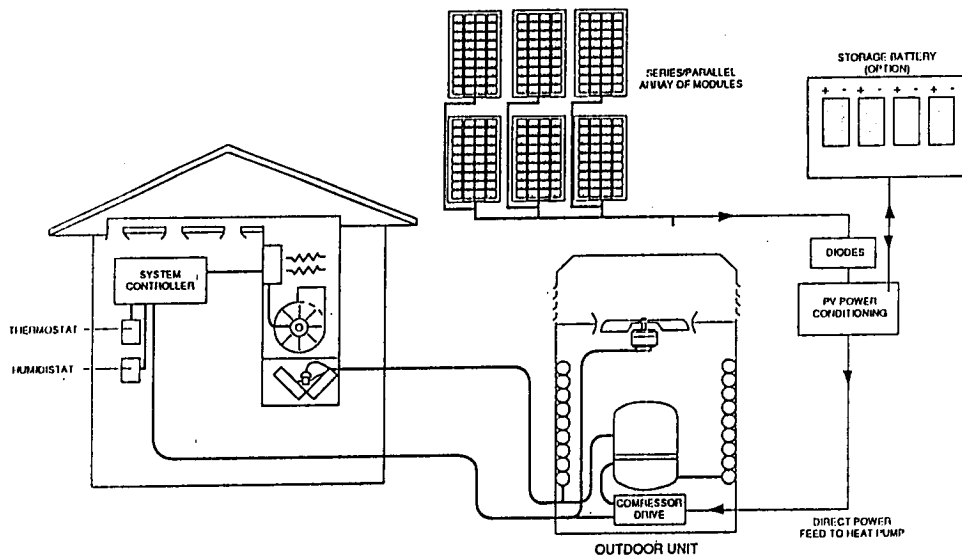
- Some future NASA missions, such as a lunar base, will require a heat pump to reject heat from the crew habitat to the hot surrounding environment
 - NASA JSC has an on-going "high-lift, high-efficiency heat pump" development effort
- Most aerospace missions derive power from solar PV arrays
- Especially on the lunar surface, solar PV power is the logical choice for a heat pump since power is not required at night and the most power is required at mid-day



SOLAR PV POWERED HEAT PUMP	<i>CREW AND THERMAL SYSTEMS DIVISION</i>	
	Michael Ewert	Feb. 2, 1994

EPRI SOLAR HEAT PUMP PROJECT

- The Electric Power Research Institute (EPRI) is a private non-profit research organization funded by electric utilities
- EPRI began a 2 year test program of solar PV powered heat pumps in 1993 in conjunction with 2 member utilities, Foster Miller (an engineering firm), Trane (a heat pump manufacturer) and Solarex (a solar array manufacturer)
- NASA JSC Crew and Thermal Systems Division learned of the project and discussed common goals with EPRI
- JSC was invited to join the project as an additional test site in a cost sharing relationship with EPRI and Houston Lighting and Power Co., an EPRI member utility
- The EPRI project has now grown to 7 test sites nation-wide



General Layout of the PV-Assisted Heat Pump



SOLAR PV POWERED HEAT PUMP	<i>CREW AND THERMAL SYSTEMS DIVISION</i>	
	Michael Ewert	Feb. 2, 1994

JSC TEST SITE DESCRIPTION

- A 3 kW Solarex PV array will be installed on the roof of building 241 at JSC
- A 5 ton cooling capacity commercial Trane heat pump unit will be provide part of the heating and cooling for the building, which houses advanced life support laboratories
- The solar array will be directly connected to the heat pump compressor, which has a variable speed brushless DC motor for high efficiency
- The heat pump will automatically use grid power to make up the difference when solar array output is not sufficient
- During the cooling season, the solar array output profile is expected to follow the heat pump power demand profile nicely



SOLAR PV POWERED HEAT PUMP	<i>CREW AND THERMAL SYSTEMS DIVISION</i>	
	Michael Ewert	Feb. 2, 1994

JSC TEST SITE SCHEDULE

- The project is currently in the procurement phase
- Installation of the system at JSC is expected in April 1994
- Test data will be collected for a period of 2 years
- Following the test period, NASA will integrate the results into on-going heat pump development efforts

**SOLAR PV POWERED
HEAT PUMP***CREW AND THERMAL SYSTEMS DIVISION***Michael Ewert****Feb. 2, 1994**

CONCLUSION

- **NASA JSC is entering into a partnership with industry to develop solar PV powered heat pump technology**
- **NASA, HL&P, EPRI, Foster Miller, Trane and Solarex will conduct a two year test program at JSC with commercially available hardware**
- **The benefits of the project will include:**
 - **Early hands-on experience with hardware for all participants**
 - **Investigation of various system configurations**
 - **Exchange of technical information between government and private industry**
 - **Creation of local interest in future energy saving technologies**

54

KINETIC STUDY OF METHYL ACETATE OXIDATION IN A Pt/Al₂O₃ FIXED-BED REACTOR

K. Y. Li*, Jeffrey S. Li, S. M. Chen, C. L. Yaws, H. W. Chu, and W. E. Simon
Lamar University, P.O. Box 10053, Beaumont, TX 77710

Michael Hoy

NASA Johnson Space Center, 2101 NASA Rd. 1, Houston, TX 77058, MAIL CODE EC3

ABSTRACT: To support technology development for future long-term missions, a metabolic simulator will be used in a closed chamber to test the functions of a Controlled Ecological Life Support System (CELSS). Methyl acetate (MA) was selected as the fuel because its metabolic respiratory quotient is near that of humans. A kinetic study of the catalytic oxidation of MA over Pt/Al₂O₃ was then conducted to support the design and operation of the simulator. Kinetic data were obtained as a conversion percentage of MA versus retention time. The reaction was studied at one atmosphere and temperatures from 220 to 340°C. The inlet MA concentration was varied from 100 to 2000 ppm with retention times from 0.01 to 10 sec. A first-order rate law and a Langmuir-Hinshelwood rate equation were tested by nonlinear regression of the kinetic data to estimate rate constants in the rate law. Regression results of the L-H equation explain the kinetic data better than the results of the first-order rate law. A Taguchi experimental design was used to study the effects of temperature, retention time, and concentrations of MA, CO₂, and O₂ on the conversion of MA. Results indicate that temperature has greatest effect, followed by retention time, and finally MA concentration. It was further determined that the effects of CO₂ and O₂ concentrations, and the cross effects, are negligible.

INTRODUCTION

One of the most challenging problems associated with long-duration manned space flight is found in the development of Controlled Ecological Life Support Systems (CELSS). This includes the technologies of air revitalization, water recovery, waste processing and food production, and the integration of these systems into an optimal closed life support system for future space missions [Elikan, 1966; Schwartzkopf, 1992]. While early studies addressed some of these technology areas on an individual basis, little effort was put forth to develop an integrated, closed-cycle ecological space life support system. At the present time, JSC engineers from the Crew and Thermal Systems Division (CTSD) are in the process of converting the 20' man-rated vacuum chamber in Building 7A to a full-up Systems Integration Research Facility (SIRF), with the intent of eventually conducting a fully integrated, reduced-pressure (10.2 psia), test with four humans and all required closed-cycle life support equipment to sustain the crew for one year in an isolated condition.

To achieve this goal, subsystem development work has begun with the construction of experimental waste-processing and plant growth facilities at JSC, as well as contractor development of various subsystems. To facilitate efficient development and testing of these systems, expensive man-in-the-loop testing of early development subsystems must be avoided. To accomplish this, a human metabolic simulator is desired to simulate the presence of

* correspondence concerning this paper should be addressed to K. Y. Li

humans in the closed environment. Such a metabolic simulator has been conceptually designed by JSC engineers and their support contractors. This conceptual design, in summary, is based on the catalytic oxidation of methyl acetate (MA). This simulator is designed to provide proper carbon dioxide, humidity, and metabolic heat load to a test-bed environment [Lin, 1982; Henninger, 1993]. However, experimental kinetic data associated with the methyl acetate oxidation were very limited [Maurel, 1982; Lang, 1991 & 1992] and specific data at design temperature were unavailable. Therefore, in this project a subscale kinetic investigation was initiated to obtain the needed kinetic data for the catalytic oxidation of methyl acetate. The information obtained from this study will be beneficial for the design and operation of both the metabolic simulator and industrial catalytic oxidation units [Kesselring, 1986; Becker, 1989].

The Taguchi method for the design of experiments was used in this study to determine both the effects and the cross-effects of the controlled variables, i.e., temperature, retention time and concentrations of methyl acetate, oxygen and carbon dioxide. This experimental design method proved to be helpful in analyzing a system with several uncertainties, since it offered maximum amount of information with minimum sets of experiments [Ross, 1988].

EXPERIMENT

Experimental Configuration

The experimental setup is shown in Figure 1. Air was supplied by an air cylinder, with one air stream bubbled through a methyl acetate storage tank submerged in an isothermal bath for temperature regulation. This air stream containing methyl acetate vapor was then mixed with another air stream from the air cylinder through a static mixer. The concentration of methyl acetate in the air was controlled by varying the flow rates of these two air streams while the total air flow rate was kept constant. The mixed air stream containing methyl acetate vapor passed through a preheater and then entered the catalytic reactor, which was controlled at a predetermined reaction temperature.

The catalytic reactor was fabricated from one-inch diameter by one-foot stainless steel tubing. The reactor was packed with 10 grams of 1% platinum-on-alumina (Pt/Al₂O₃) non-porous 1/8" catalyst pellets. The methyl acetate, oxygen, and carbon dioxide contents of the inlet and outlet streams of the catalytic reactor were analyzed by a gas chromatograph (GC) (Varian 3400 with Alltech CTR column) through an automatic sampling valve system. A Thermal Conductivity Detector (TCD) and a Flame Ionization Detector (FID) were used in the GC for the chemical analysis. Details of the preheater, the reactor, and the gas sampling system are shown in Figure 2.

Pre-Test Considerations

Prior to the reaction exposure, the catalyst was pretreated by calcination in a flowing air stream. The calcination process was implemented by first drying the catalyst at 125°C in air flowing at 1.0 liter/min for 2 hours. The catalyst temperature was then raised gradually to 425°C and maintained for 16 hours. After calcination, the catalyst temperature was decreased to 250°C and hydrogen was then passed through the catalyst to reduce any oxidized surface.

Before beginning the experiment, the deactivation of the catalyst was tested to see if it would present a problem during the time period of the experiment. The catalyst was continuously exposed to a dry air stream containing 3000 ppm methyl acetate at 258°C for 70 hours. During this time period, the conversion of methyl acetate was then monitored and recorded continuously. No significant change in methyl acetate conversion could be detected during a period of 70 hours. This suggested that the catalyst did not lose its activity for at least this period of time. Consequently, the catalyst used to obtain the kinetic data was exposed to the oxidation process for less than 70 hours. Once 70 hours were reached, the catalyst was replaced or regenerated with hydrogen.

A blank run was initially conducted to determine the conversion percentage of methyl acetate without catalyst prior to beginning the catalytic kinetic measurement. For this blank run, glass beads (approximately 6.9g) were placed in the reactor as a substitute for the catalyst. Methyl acetate was then fed through the reactor at the lowest flow rate (0.6 ml/min) and the highest temperature (380°C) planned for the test with the catalyst. An average 2.75% conversion of the methyl acetate was obtained for different concentrations of MA. This suggested that the noncatalytic conversion due to thermal decomposition could be neglected as compared to the catalytic conversion.

In an experiment of this type, it is important to obtain true kinetic data without a distortion of interphase mass transfer resistance. Otherwise, the reaction rate obtained from the experiment may not be a true reaction rate [Satterfield, 1970]. In this study, Mears criterion was applied as a check to see if the reaction was external-mass-transfer limited [Mears, 1971]. The Mears criterion is described below:

$$\frac{r \rho_b R^n}{k_c C_A} < 0.15 \quad (1)$$

where r = reaction rate, mole/kg_{catalyst} sec
 ρ_b = bulk density of catalyst bed, kg/m³
 R = catalyst particle radius, m
 n = reaction order
 C_A = bulk concentration, mol/m³
 k_c = mass transfer coefficient, m/sec

The value of the mass transfer coefficient, k_c , in a fixed-bed reactor is calculated from a semi-empirical equation [Dwivedi and Upadhyay, 1977]. The left-hand side of equation 1 is a ratio of reaction rate to mass transfer rate. If this ratio is less than 0.15, this means that the external diffusion limitation is insignificant, i.e., it is not a problem.

For the nonporous catalyst used in this study, it was found that the external-mass-transfer resistance could be neglected during all of the experiments of kinetic measurement [Li, 1993]. Under this condition, the concentration gradient between the bulk fluid and the external surface of the catalyst can be neglected. Therefore, the measured bulk concentration is equivalent to the concentration on the catalyst surface, and the measured reaction rate can be used for kinetic analysis.

EXPERIMENTAL RESULTS AND DISCUSSION

Taguchi Experimental Design

The Taguchi method of experimental design is advantageous in analyzing a system with several uncertainties because it offers maximum information with minimum sets of experiments. A Taguchi orthogonal array, L_{27} , (shown in Table I) was constructed based on 3-level factors with interaction, i.e., reaction temperature (A), methyl acetate concentration (B), gas flow rate (C), and 3-level factors with no-interaction, i.e., pressure (D), carbon dioxide concentration (E), and oxygen concentration (F). The three levels of each factor are also indicated in Table I. The pressure factor was eliminated due to the difficulty of conducting a reduced-pressure test at Lamar University. Experimental results, namely, conversion % of methyl acetate of the 27 experimental sets, were then obtained and these are shown in Table I.

Based on these experimental design results, an analysis of variance was then performed. The results indicated that the random error and the error fluctuations from E, F and BxC factors are very small [Li, 1993]. It is therefore convenient to lump these fluctuations into one random error term, and a simplified analysis of variance can be computed as shown in Table II [Gunst, 1980; Roy, 1990]. The F values were obtained from the ratio of the mean square of the factor to that of the lumped random error. Results indicate that the reaction temperature (A-factor) is the most significant factor affecting the conversion of MA. The next significant factor is the gas flow rate, followed by methyl acetate concentration. The cross effect of either AxB or AxC is insignificant. The effect of MA concentration on the conversion of MA implies that the adsorption step may be critical in determining the rate equation. The determination of the rate equation will be discussed in the following section.

Selection of Rate Equation

For a fixed-bed integral reactor, the combination of rate equation and design equation leads to [Froment, 1979]:

$$\frac{W}{F_{MA}^0} = \int_0^x \frac{dx}{r} \quad (2)$$

- where W = catalyst weight, g
 F_{MA}^0 = initial flow rate of reactant A, mole/sec
 x = conversion of reactant A
 r = rate equation

For a fixed amount of catalyst, W , the conversion of methyl acetate, x , is a function of the flow rate, F . Therefore, a set of experimental data was acquired with different flow rates (retention time). The rate equation could be a simple first-order rate law or a complicated rate law considering chemical reaction, adsorption, and desorption steps.

Two assumptions were made prior to the data analysis. First, the concentration of oxygen was assumed to be constant because the consumption of oxygen (maximum 0.75 %

based on inlet MA=2000 ppm) was not significant compared with the oxygen concentration percent in air. Second, the small amount of water produced (0.75% maximum) due to the reaction did not affect the reaction rate. The rate equation could then be set up by considering methyl acetate concentration only. Two rate equations were proposed, a first-order power law and a Langmuir-Hinshelwood model, as listed in Table III. After the rate equation was selected, Equation 2 was rearranged and integrated. The integrated equations are also shown in Table III. Then a program of nonlinear regression using Marquardt's method was used to fit the kinetic data at the same temperature to estimate parameters such as k (and k_a for the Langmuir-Hinshelwood model).

A first-order power law was taken as the simplest approach in formulating the rate equations. This approach ignores the adsorption and desorption processes occurring on the catalyst surface and provides less information about how the reaction occurs than a more elaborate formula. However, when the objective is reactor design, the first-order power law is advantageous because it has fewer adjustable parameters. Additionally, such an equation can frequently correlate the experimental data accurately for industrial reactor design. Therefore, it was first selected to fit the experimental data and the result was shown in Figure 3. As can be seen from this figure, the conversion of methyl acetate decreases as its concentration increases at a reaction temperature of 240°C. However, the first-order rate law fails to predict the effect of methyl acetate concentration. The rate constant, k , for modeling purposes was treated as a combination of a surface reaction rate constant and an adsorption equilibrium constant.

Another model used to correlate the experimental data, the Langmuir-Hinshelwood model, is also included in Table III. The rate constant, k , in the numerator was also a combination of a reaction rate constant and an adsorption equilibrium constant. Moreover, the term $k_a C_{MA}$ in the denominator took into account the inhibition due to an adsorption of reactant on the active site of the catalyst surface, where k_a is the adsorption equilibrium constant. Hence, with the increase of inlet concentration of methyl acetate, the reaction rate decreased due to a decreased active site area available on the catalyst surface. Therefore, the conversion of methyl acetate also decreases. Although the correlation at high MA concentration deviates due to a lack of experimental data points, the Langmuir-Hinshelwood model explained the methyl acetate concentration effect better than the first-order power law. This model should be used in the design of the catalytic reactor of methyl acetate oxidation.

CONCLUSIONS

The following conclusions may be drawn from this study:

1. The 0.5%Pt/Al₂O₃ catalyst was determined to be capable of converting methyl acetate to carbon dioxide and water at a temperature as low as 350°C.
2. Results from application of the Taguchi experimental design method indicate that the major factors affecting the conversion of methyl acetate are (in descending order of influence): temperature, flow rate, and methyl acetate concentration. Both concentrations of carbon dioxide and oxygen have insignificant effect on the conversion of methyl acetate.

3. The rate law of the Langmuir-Hinshelwood model was found to be superior to the first-order rate law in predicting the conversion of methyl acetate. The L-H model is successful in predicting the effect of methyl acetate concentration on its conversion.

ACKNOWLEDGMENTS

The funding of this project by NASA-JSC is gratefully appreciated. Technical discussions with NASA personnel were also beneficial in guiding the study.

LITERATURE CITED

- Becker, E. R., and Pereira, C.J., Catalyst and Reactor Design for Emission Control. AICHE Advanced Seminar Series, (1989).
- Dwidevi, P. N., and Upadhyay, S. N., "Processes Development for Catalysis," Industrial and Engineering Chemistry. Process Research and Development, 16,157-165, (1977).
- Elikan, L., "An Approach to Water Management for Long Duration Manned Space Flights," Aerospace Life Support volume of Chemical Engineering Progress Symposium Series Number 63, Volume 62, American Institute of Chemical Engineers, New York, NY (1966).
- Froment, G. F., and Bischoff, K. B., Chemical Reactor Analysis and Design. New York: John Wiley & Sons, Inc. (1979).
- Gunst, R. F., and Mason, R. L., Regression Analysis and Its Application. New York: Marcel Dekker, Inc. (1980).
- Henninger, D., "Controlled Ecological Life Support Systems (CELSS) Research and Technology Development at the Johnson Space Center," CELSS Conference of 1993, Alexandria, Virginia, March 1-3, 1993.
- Kesselring, J. P., "Catalytic Combustion", Advanced Combustion Method. London: Academic Press, 237-275, (1986).
- Lange, K. E., "Summarized Work on the Analysis of a Proposed Metabolic Simulator", NASA/JSC Document, (1991).
- Lange, K. E., "Requirements for the Test Bed Metabolic Simulator," Report to NASA-JSC, (1992).
- Li, Jeffrey S., "The Kinetic Study of Methyl Acetate Oxidation over Platinum Alumina Catalyst", Masters Thesis, Lamar University, Beaumont, Texas (1993).
- Lin, C. H., "Space Station Environmental Control and Life Support System Preliminary Conceptual Design," NASA/JSC Doc. No. CSD-SS-059, (1982).
- Maurel, J.-L., and J.-L. Vernet, "Catalyse. - Etude de la Combustion Catalytique du Formlate et de l'Acetate de Methyle," C. R. Seances Acad. Sci. Ser. 2, 294(1) 25-28 (1982).
- Mears, D. E., "Test for Transport Limitations in Experimental Catalytic Reactors," Industrial and Engineering Chemistry Process Design and Development. 10, 541-548, (1971).
- Ross, P. T., "Taguchi Techniques for Quality Engineering", McGraw Hill, (1988).
- Roy, R. "A Primer on the Taguchi Method," Van Nostrand Reinhold, (1990)
- Satterfield, C. N., Mass Transfer in Heterogeneous Catalysis, Robert E. Krieger Publishing Co,(1970).
- Schwartzkopf, S. H., "Design of a Controlled Ecological Life Support System," BioScience Vol 42, No. 7, 526, (1992).

Table I Taguchi's Experimental Design Orthogonal Array

Set	A	B	C	E	F	MA Conv.
1	1	1	1	1	1	0.09
2	1	1	2	2	2	0.03
3	1	1	3	3	3	0.05
4	1	2	1	2	3	0.088
5	1	2	2	3	1	0.09
6	1	2	3	1	2	0.05
7	1	3	1	3	2	0.04
8	1	3	2	1	3	0.06
9	1	3	3	2	1	0.04
10	2	1	1	3	2	0.57
11	2	1	2	1	3	0.33
12	2	1	3	2	1	0.24
13	2	2	1	1	1	0.47
14	2	2	2	2	2	0.3
15	2	2	3	3	3	0.27
16	2	3	1	2	3	0.25
17	2	3	2	3	1	0.25
18	2	3	3	1	2	0.15
19	3	1	1	2	3	0.93
20	3	1	2	3	1	0.83
21	3	1	3	1	2	0.81
22	3	2	1	3	2	0.89
23	3	2	2	1	3	0.84
24	3	2	3	2	1	0.76
25	3	3	1	1	1	0.87
26	3	3	2	2	2	0.75
27	3	3	3	3	3	0.61

level*	A:Temp.	B:MA con	C:flowrate	E:CO2 con.	F:O2 con.
1	250°C	100 ppm	1.5 l/min	0%	21.0%
2	320°C	500 ppm	3.0 l/min	3.0%	28.5%
3	420°C	2000 ppm	4.5 l/min	8.5%	40.4%

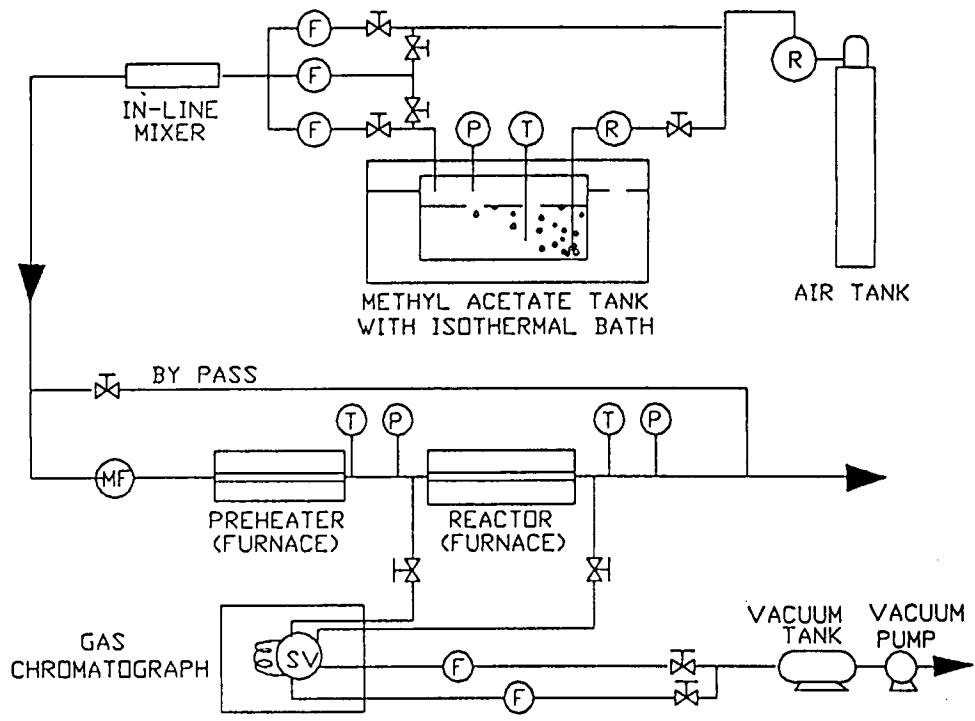
Table II Analysis of Variance of A Simplified Result of Taguchi's Experimental Design

Factor	Degree of Freedom	Sum of Square	Mean Square	F*
A	2	26197.91	13098.95	533.78
B	2	481.16	240.58	9.80
C	2	832.99	416.45	16.97
AxB	4	189.74	47.44	1.93
AxC	4	292.71	73.18	2.98
e+E+F+BxC	12	294.46	24.54	
Total	26	28288.97		

* Confidence Level = 95% 99%
 F(2,12) = 3.88 6.93
 F(4,12) = 3.26 5.41

Table III
 Equations for Regression Analysis with Different Rate Laws

Rate Laws	Equations for Regression Analysis
first-order power law $rate = k C_{MA}$	$\frac{W}{F_{MA^0}} = \frac{-1}{k C_{MA^0}} \ln(1-x)$
Langmuir-Hinshelwood $rate = \frac{k C_{MA}}{1 + k_a C_{MA}}$	$\frac{W}{F_{MA^0}} = \frac{-1}{k C_{MA^0}} \ln(1-x) + \frac{k_a}{k} x$



- Note:
- (T) Thermalcouple
 - (P) Pressure Gauge
 - (MF) Mass Flow Meter
 - (R) Gas Pressure Regulator
 - (F) Float Type Flowmeter

Figure 1

Flow Diagram of Experimental Equipment

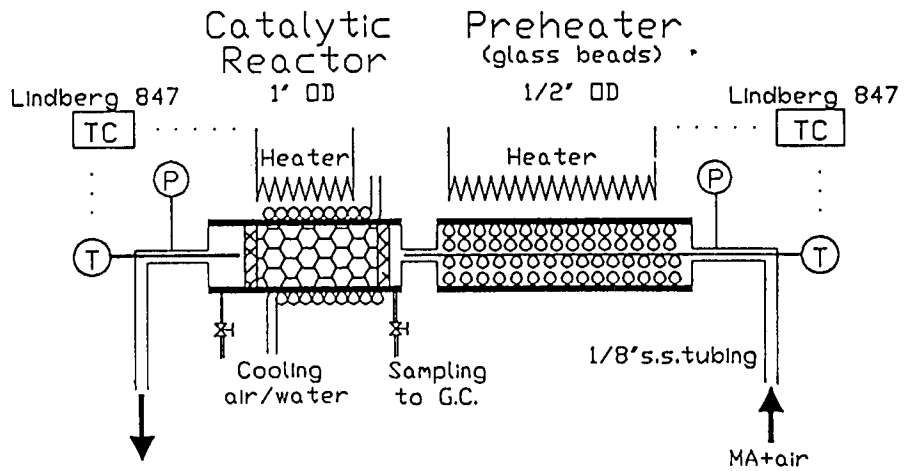


Figure 2

Structure of the Preheater and Reactor

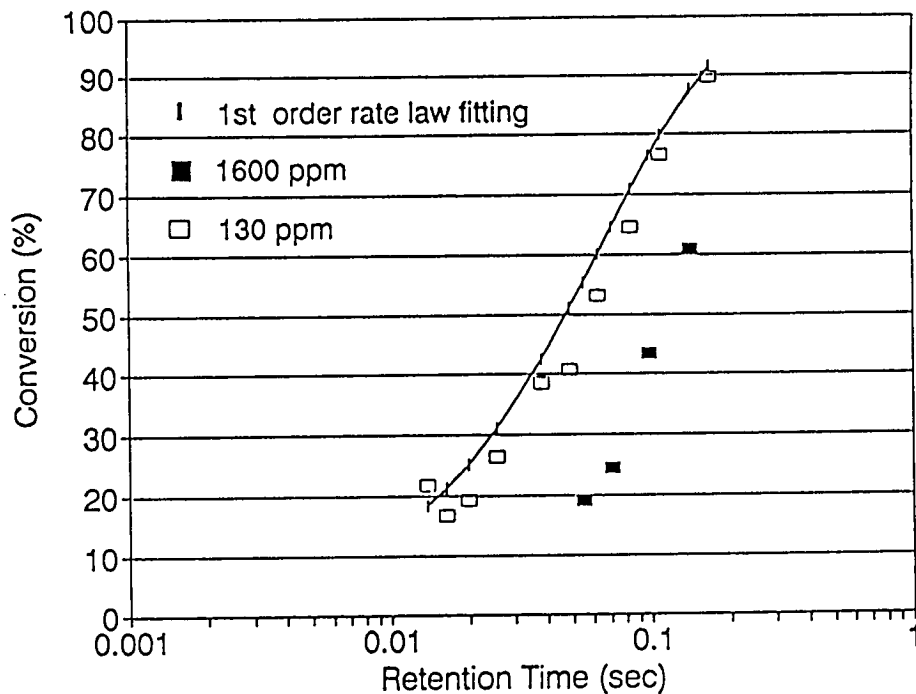


Figure 3

Curve fitting result from the first-order rate law for the experiment at $T=240^{\circ}\text{C}$ and inlet MA=130 and 1600 ppm.

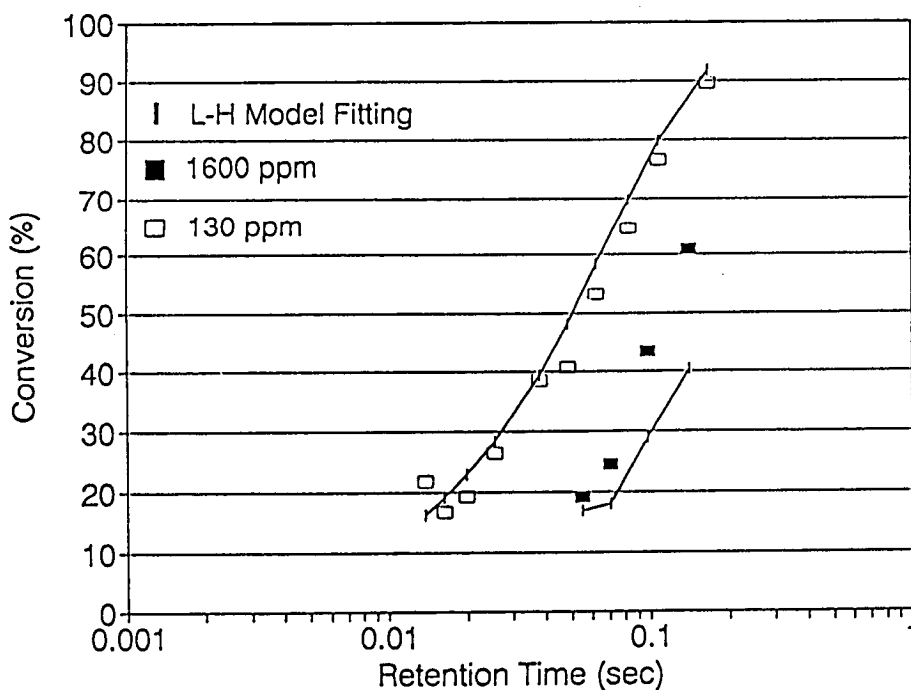


Figure 4

Curve fitting result from the Langmuir-Hinshelwood model for the experiment at $T=240^{\circ}\text{C}$ and inlet MA=130 and 1600 ppm.

SOLID STATE OXYGEN SENSOR DEVELOPMENT

Jeffery T. Cheung Ph. D.

Scott R. Johnson

Rockwell International

ABSTRACT

To anticipate future long-duration mission needs for life support sensors, we explored the feasibility of using thin-film metal-oxide semiconductors. The objective of this task was to develop gas sensors for life support applications which would be suitable for long-duration missions. Metal oxides, such as ZnO, SnO₂, and TiO₂ have been shown to react with oxygen molecules. Oxygen lowers the metal oxide's electrical resistance. Critical to the performance is the application of the oxide in a thin film on an inert substrate: the thinner the film, the more readily the oxygen penetration and hence the more rapid and sensitive the sensor. Metal oxides are not limited to oxygen detection, rather, oxides offer detection and quantification applications to the complete range of gases of interest, not only for life support systems, but for propellants as well.

After a preliminary assessment of various metal oxides, we chose ZnO as the active sensory element for oxygen - the most critical near-term need for life support sensors. Our results to date are that we:

- Verified that metal oxides respond linearly to oxygen concentration
- Experimented with deposition techniques for a thin film of ZnO

- Determined the appropriate operational parameters by which ZnO could be used to measure oxygen in concentrations from 10% to 30%
- Selected a conceptual design suitable for proof of concept
- Developed techniques to fabricate the prototype sensor compatible with integrated circuit and microelectronic packaging

Subsequent work will optimize the ZnO properties, refine prototype design and fabrication, and install and test the sensors in the KSC Plant Growth Chamber.

INTRODUCTION

There is a need for a reliable, light weight oxygen sensor for advanced life support systems. Current oxygen sensors, such as those in the Space Shuttle Orbiter, are potentiometric cells with a solid electrolyte. These cells exhibit insufficient lifespan for missions in excess of 30 days. Furthermore, electrochemical cells degrade over time, requiring frequent re-calibration. These electrochemical cells are currently changed out every few days, and require refurbishment on the ground. Additional drawbacks of current systems include: bulkiness and a continuous

drift due to the saturation in the electrolyte and the change in reference electrode potential.

The use of metal-oxide semiconductor thin films as oxygen detection devices offers an alternative approach. In this report, we present results on material development and characterization, the oxygen sensing characteristics, and the design of a miniature solid state heater in order to realize an integrate solid state oxygen gas sensor. Future directions are also outlined based on these results.

It has been well known since 1962 that these oxides, such as ZnO, SnO₂, and TiO₂, can react with oxygen molecules which are chemisorbed on the surface by electron transfer. The depletion of mobile electrons from the lattice lowers the material's electrical resistance which can be used to monitor the oxygen (or other gases) concentration. The chemisorption process is temperature sensitive. Within these oxides, there are different mechanisms occurring at various operating temperatures.

In order to satisfy the requirements for space mission applications, the oxide material of choice must have a low to medium operating temperature range, to minimize power use and maximize longevity, and the entire device be patterned into a miniaturized integrated solid state package. Moreover, the response is preferably linear in the oxygen concentration range of interest: 10% to 30%. Other issues such as selectivity and long term stability are also important.

Based on these considerations, and after a suitable literature search, we choose ZnO thin film to be the active sensing material. The ideal operating

temperature for ZnO is between 250°C to 300°C when oxygen gas and ZnO interacts according to the so called "ionosorption" mechanism. During this process, electrons near the ZnO surface are depleted by chemisorbed oxygen molecules to form O₂ - the electrical resistance of the ZnO increases linearly with the oxygen concentration. At high temperatures (above 300°C), oxygen vacancy diffusion inside the lattice dominates: the diffusion process is very slow and results vary. At temperatures lower than 250°C, ZnO has insufficient thermally activated electrons and therefore the ionosorption process is too slow to be practical. Thus, we must heat the ZnO to between 250°C - 300°C and maintain this temperature for proper sensor function.

In the "ionosorption" regime, the resistance change (i.e. sensitivity) can be optimized by the doping density of the film, its thickness, and exposed surface area. For large resistance change, films with small thickness, large exposed area, and low electron density are most desirable. These material parameters are easily adjustable by using the Pulsed Laser Deposition technique to grow ZnO thin films. Finally, the temperature range of interest is sufficiently low that it can be achieved by micro-filament heaters.

MATERIAL DEVELOPMENT

The key issue in growing high quality ZnO films is the ability to incorporate sufficient oxygen into the lattice to maintain its stoichiometry. While traditional thin film growth techniques such as Chemical Vapor Deposition and Sputtering are commonly used, we find that Pulsed Laser Deposition offers the most advantage and flexibility in fulfilling this requirement. This

novel technique employs a pulsed high energy UV laser as an external power source for evaporation and deposition.

The growth using Pulsed Laser Deposition was carried out in 0.02 torr of oxygen background pressure to provide a reactive environment. Typical growth rate was 1 micron/hr. Sapphire, silicon, quartz plate, and silicon wafers with metal film for electrical contact were used as substrates. Under a laser fluence of $2\text{J}/\text{cm}^2$ to ablate an undoped ZnO target, we were able to grow ZnO films at a temperature as low as 25°C . ZnO films grown at room temperature are transparent with 82% transmission and an absorption edge at 3840\AA wavelength. Structurally, it is preferentially c-axis oriented as shown in the X-ray diffraction pattern in Figure 1.

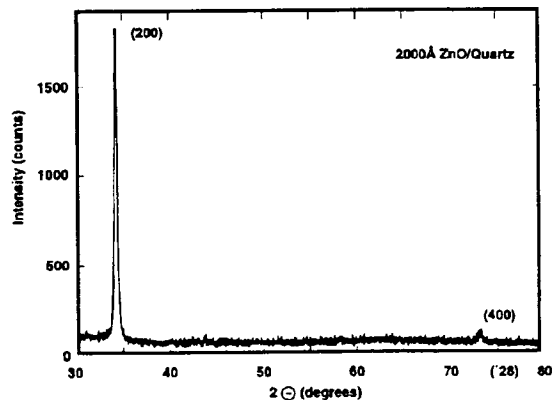


FIGURE 1. X-ray diffraction spectrum of a ZnO film grown on quartz substrate at room temperature by Pulsed Laser Deposition

On the contrary, ZnO films deposited at room temperature by other techniques (such as sputtering) are inferior. These films often exhibit a metallic appearance due to a comparatively severe oxygen deficiency. High quality films of this type are only

possible if they are grown at elevated temperature or subject to post annealing at very high temperature.

For comparative purposes, the oxygen sensing properties of films grown at different temperatures were measured and compared. In addition, we were also able to control the electron density in the ZnO film by introducing nitrogen into the lattice substitutionally. Undoped ZnO films are *n-type* due to the presence of oxygen vacancies. The electron carrier density is between 10^{19} - $10^{20}/\text{cm}^3$. In nitrogen doped ZnO films, electron density is reduced by compensation. This is significant as the resistivity of the doped films is **3 to 6 orders of magnitude higher** than the undoped samples. Clearly, this has important implications for sensor sensitivity.

SENSOR PROPERTIES MEASUREMENT

After a film is grown, we patterned a pair of Au/Ti interdigitated electrodes by photolithographic lift-off technique. The ratio of the total length of the electrode to the spacing distance between the two electrode is about 100:1. Electrodes are surrounded by a guard ring biased at ground potential to reduce spurious noise. Thin gold wires are attached to the electrode pads by silver paste for electrical measurement. The sensor is mounted on an inconel holder whose temperature can be regulated up to 700°C with an accuracy of $\pm 0.5^\circ\text{C}$.

The resistance of the ZnO thin film was measured by applying a constant voltage across the electrodes while measuring the current flow. The test chamber has a capacity of only 3 liters to facilitate rapid gas purging. Oxygen and nitrogen gases are introduced through

mass flow meters and well mixed in a T-shaped tube filled with glass grit. During the measurement, the test vessel was wrapped tightly with aluminum foil to eliminate any photoconductive effect.

Figure 2 shows the response of a 1.2 micron thick ZnO film at 275°C to an injection of oxygen into nitrogen to achieve a 20% Oxygen/80% Nitrogen mixture. The film was grown at 175°C. The response time, defined as the time required to achieve 95% of the final constant value, is about 270 seconds (4 ½ minutes). It does not vary considerably within the studied temperature range from 250°C to 300°C. The rate limiting step is believed to be the diffusion of oxygen along the grain boundary to be chemisorbed at the active surface site. This suggests that the response time can be shortened by reducing the thin film thickness.

In order to speed up response time, we tested a film grown under identical conditions but with a thickness of only 5000Å, it shows a shorter response time of only 155 seconds (2.58 minutes). The response time can also be shortened by nitrogen-doping into the ZnO film. Figure 2 compares the response time between a nitrogen-doped ZnO film and an undoped film. The nitrogen-doped film, although twice the thickness as the undoped film, shows a faster response: 92 sec (1.53 minutes). A plausible explanation is that the introduction of low concentration levels of nitrogen atoms to substitutionally replace the oxygen atoms in the lattice. This replacement can cause internal stress because of the different atomic radii of the two species. The stress enhances the rate of diffusion of oxygen along the grain boundary. Further investigation out to optimize the film thickness and

nitrogen doping level with respect to improve the response time is needed.

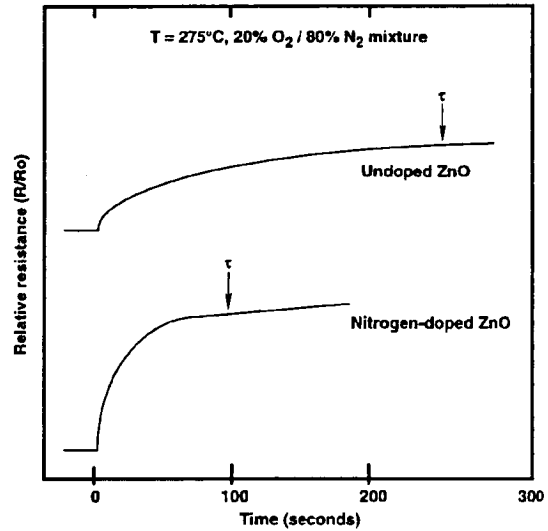


FIGURE 2. Response of the electrical resistivity of various ZnO films to oxygen in a 20% oxygen/80% nitrogen mixture

SENSITIVITY

Figure 3 shows the resistance of ZnO thin film as a function of oxygen concentration in an oxygen/nitrogen mixture at 275°C. For undoped ZnO, the sensitivity increases with the decrease in thickness. Since the thickness of the electron depleted layer occupies to a larger fraction of the total thickness, sensitivity also increases with the decrease in electron carriers since the number of depleted electrons will account for a larger portion of the total electrons. This is reflected in the results showing that thin undoped ZnO film has a higher sensitivity than thick films while nitrogen doped ZnO film exhibits the highest sensitivity. A systematic study will be carried out to optimize the sensitivity by adjusting the nitrogen doping level.

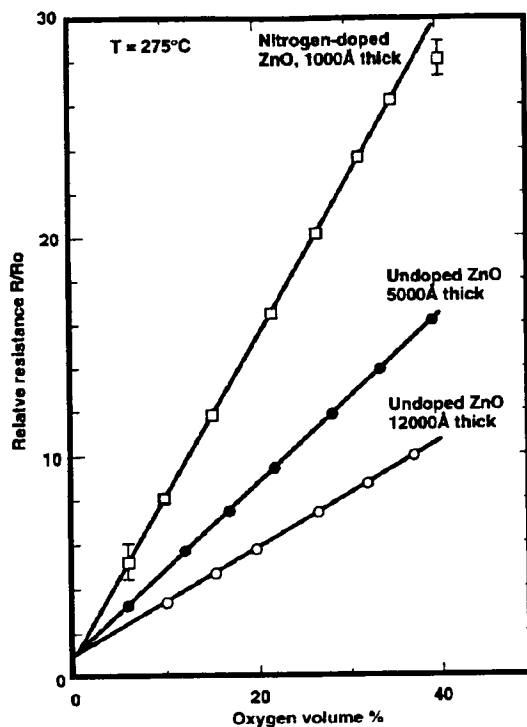


FIGURE 3. Relative electrical resistivity R/R_0 of various ZnO films vs. oxygen concentration in oxygen nitrogen mixture. (R_0 is the resistance of the film in the absence of oxygen)

In addition to basic performance characteristics, we also found that common gases such as CO_2 and NH_3 have no interference effect. Preliminary investigation on reliability shows that during the first 60 hours of operation, the sensors show some degradation. However, after the initial "burn in" period, they show excellent long term stability. The stability seems to be dependent on the crystallinity of the ZnO film.

MICRO-ELECTRONIC PACKAGING

For operational use, the entire sensor assembly will be miniaturized to a chip-sized solid state device with an integrated micro-heater and a built in temperature sensor. The initial sensor characterization was carried out in a environmentally controlled chamber with a

bulk heater. Of course, this assembly is impractical and not representative of life support applications. To enable the assembly and realistic test of a miniaturized prototype, a photolithographic set of five masks has been designed. Fig. 4 shows the cross sectional view of this device.

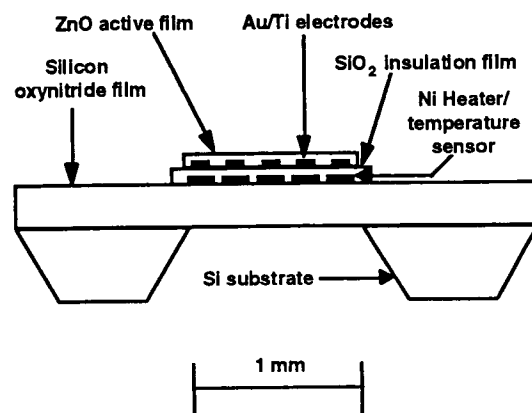


FIGURE 4. Cross sectional schematic view of a miniature solid state oxygen sensor

The assembly of this prototype is anticipated for December 1993. The active element is at the center of a low stress and low thermal conductivity silicon oxynitride free standing film on a silicon substrate for thermal isolation. A meandering thin film nickel filament micro-heater and a resistive type temperature sensor will be deposited and patterned. The heater will then be covered by a SiO_2 layer for electrical insulation. This will be followed by the deposition and patterning of a pair of interdigitated Au electrodes aligned just on top of the heating element. Finally, ZnO film will be deposited on the interdigitated electrode. Each wafer is 8 mm x 8 mm in size and will be mounted on a 20 mm x 20 mm 20-pin header. Each wafer contains three identical gas sensor devices. Integral, triple-redundant cells will serve as a cross check to assure accuracy and inherent backup in case a

single unit failure. The active area of each device is only 0.5 mm x 0.5 mm. An estimate power consumption to achieve 300°C in such a small localized area is approximately 1.3 watts.

CONCLUSIONS

We have demonstrated the merits of using nitrogen-doped, thin-film ZnO for oxygen detection and measurement. This technique offers great promise not only for oxygen detection, but other gases. The response of thin-film ZnO is stable over time, a significant advantage over current oxygen sensors. Because of the relatively small size of our prototype sensors, distributed networks of low-weight, low power sensors can more accurately assess local environmental conditions than can current sensors. Low power use and low weight are obvious advantages for space systems. All of these benefits offer a significant improvement in current life support systems control.

FUTURE WORK

1. We will continue to optimize the ZnO film properties, especially thickness and nitrogen doping level to improve sensitivity and shorten response time.
2. We will gain a better understanding the degradation mechanism during the initial "burn in" period. This has been speculated to be related to the crystalline quality.
3. Integrated solid state oxygen gas sensors will be fabricated and tested.

REFERENCES

- U. Dibbern; A Substrate for Thin-film Gas Sensors in Microelectronic Technology, *Sensors and Actuators B*, 2 (1990), pg 63 - 70.
- H. Nanto, T Minami, S. Takata; Zinc-oxide Ammonia Gas Sensors with High Sensitivity and Excellent Selectivity, *J. Applied Physics*, 60 (2), 1986, Pg 482.
- H. Nanto, H. Sokooshi, T Usuda; Smell Sensor Using Aluminum-doped Zinc Oxide Thin Film Prepared by Sputtering Technique, *Sensors and Actuators B*, 10 (1993), pg 79 - 83.
- Uwe Lampe, Jörg Müller; Thin-Film Oxygen Sensors Made of Reactively Sputtered ZnO, *Sensors and Actuators*, 18 (1989), pg 269 - 284.
- A. R. Raju, C.N.R. Rao; Gas-sensing Characteristics of ZnO and Copper-impregnated ZnO, *Sensors and Actuators B*, 3 (1991), pg 305 - 310.
- G. Sberveglieri, P. Nelli, S. Groppelli; Oxygen Gas Sensing Characteristics at Ambient Pressure of Undoped and Lithium-doped ZnO-sputtered Films, *Materials Science and Engineering*, B7 (1990), pg 63 - 68.

CHIT TO
P

**Performance Enhancement of
Heat and Mass Transfer Devices
With Electrohydrodynamics (EHD)**

**Dr. J. Seyed-Yagoobi
Department of Mechanical Engineering
Texas A&M University
College Station, Texas**

**K. Miller
Crew and Thermal Systems Division
NASA Johnson Space Center
Houston, Texas**

**Presentation at the Dual-Use Technology Conference
NASA Johnson Space Center
February 2, 1994**

NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

Outline

- Background on EHD J. Seyed-Yagoobi
- Applications of EHD Technology
- Advantages of EHD Technology
- Overview of the Texas A&M Program
- Overview of the NASA-JSC Program K. Miller
- Dual-Use Potential
- References

Background

Enhancement Techniques

- Passive (no direct application of external power)
- Active (need external power)

There are currently over 20 enhancement techniques in use; 11 are passive and 9 are active.

Background (Continued)

Electrohydrodynamic (EHD) pumping is produced by the interaction of electric fields and free charges in a dielectric fluid medium (Coulomb Force).

Background (Continued)

Electrohydrodynamics: interaction of electric fields and fluids

$$F_e = \rho_e E - 1/2 E^2 \nabla \epsilon + 1/2 \nabla \left[E^2 \rho \left(\frac{\partial \epsilon}{\partial \rho} \right) \right]$$

F_e	body force
E	electric field
ϵ	permittivity
ρ	mass density
ρ_e	charge density

Background (Continued)

F_e : External Force

Coulomb force requires free charges. This force can be significant in single and two phase flows.

Electrophoretic force requires electrical permittivity gradient. This force is generally negligible in single phase fluids (gas or liquid). However, it can be significant in two-phase flows.

Electrostriction force requires electric field gradient. In order for this force to be significant the change in permittivity with density also has to be significant.

Background (Continued)

Free Charges:

- a. Direct injection of free charges by means of a corona source (Ion-Drag Pumping)
- b. Induction charging based upon the establishment of an electrical conductivity gradient perpendicular to the desired direction of fluid motion (Induction Pumping). For instance, the electrical conductivity gradient is established by the temperature gradient present in the channel or pipe.

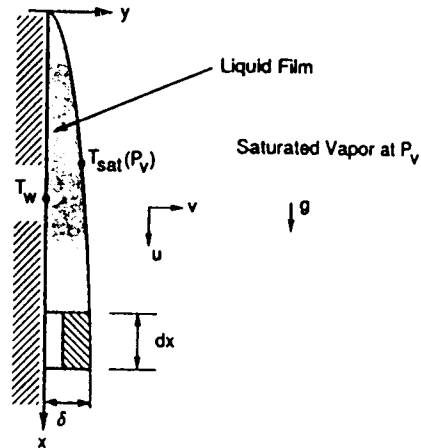
Background (Continued)

Requirements:

- a. The dielectric fluid being pumped must contain free charges.
- b. An electric field must be present to interact with the free charges in the fluid.

Background (Continued)

- Condensation on Vertical Surface



- Restricted by Liquid Film
 - > reduce film thickness (electrophoretic force)

Possible Applications of EHD Technology

(Outer Space and Terrestrial)

- a. Enhancement of Heat and Mass Transfer (Single-Phase and Two-Phase Flows)
- b. Pumping of Fluids Such as Cryogenics, Hydrocarbons and Radioactive Fluids.
- c. Mixing of Fluids
- d. Others

Advantages of EHD Technology

- Non-mechanical (no rotating parts, no vibration)
- Low Power Input
- No Need of External Pressure for Operation; Uniform Low Pressure Generation Along Pumping Section
- Inexpensive and Simple Design
- Lightweight
- Rapid Control of Heat and Mass Transfer Coefficients
- Localized Cooling of Complex Curved Passages
- Negligible Heat Generation

Disadvantages of EHD Technology

- High Voltage Required for Operation
- Dielectric Fluid Required
- Long Term Operation Has Not Been Addressed

Overview of the Texas A&M Program

(Video Presentation)

- Pumping of Hydrocarbons and Refrigerants (ion-drag mode and induction mode)
- Enhancement of Heat Transfer in Single Phase Fluid Flow
- Enhancement of Heat Pipe Performance
- Enhancement of Condensation Heat Transfer

Overview of the NASA JSC Program

- The program is a joint effort with NASA JSC, McDonnell Douglas Aerospace (MDA), and Texas A&M University.
 - NASA JSC - facilities, test hardware, test and technical support; funding provided by Center Director's Discretionary Funding and other internal resources
 - MDA - full-time principal investigator, Mr. James Bryan; funding provided by the MDA Independent Research and Development Program
 - Texas A&M University - part-time principal investigator, Dr. Jamal Seyed-Yagoobi and graduate student; support for heat pipe modelling and investigation of dual-use potential; funding provided by NASA JSC Grant
- Overall Program Objective - to develop advanced EHD-augmented heat transfer devices with significantly increased performance over existing heat rejection, acquisition and transport devices

Overview of the NASA JSC Program (cont'd)

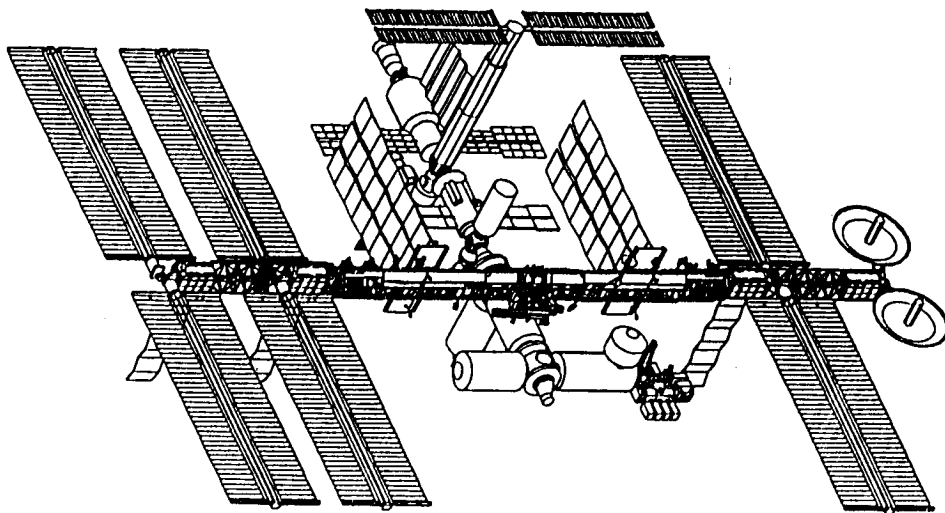
- An example of the need for advancements in space thermal control systems is the Space Station Alpha program.
 - extremely large radiator area required, 780 sq m (8400 sq ft)
 - extremely massive system due to large radiator size, 6820 kg (15000 lbs)
- The NASA JSC Program consists of two parts:
 - EHD Development Loop, and
 - EHD Heat Pipe Loop Studies.

NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

Space Station Alpha



NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

Overview of the NASA JSC Program (cont'd)

EHD Development Loop

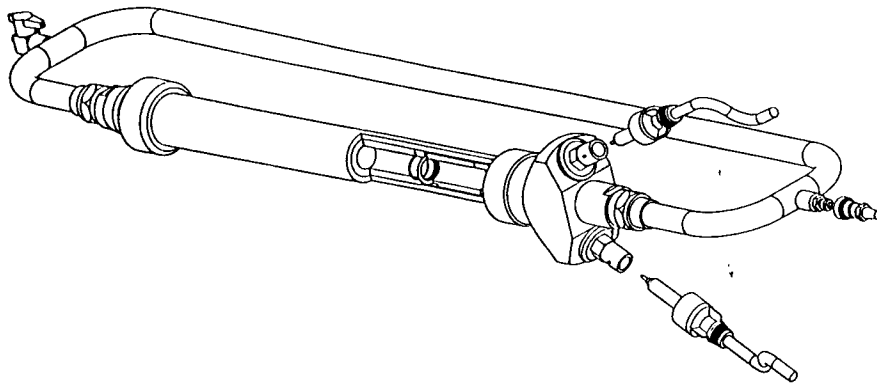
- **Description**
 - The testbed is located in the Building 7 High Bay, Room 2006.
 - The major components include: the test loop, test bench, Laser Doppler Velocimetry (LDV) system, and data acquisition system.
 - The primary objectives of this testbed are:
 - to characterize the flow local to the electrodes and elsewhere in the flow stream, and
 - to support design optimization; specifically investigating electrode design and spacing.

NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

EHD Development Loop



NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

Overview of the NASA JSC Program (cont'd)

EHD Heat Pipe Loop Studies

- **Description**

- The test bed will be located in the Building 7 Highbay, Room 2006.
- The heat pipe test article is designed to be modular to allow testing of various evaporator, condenser, and EHD pumping sections.
- The major components include the modular heat pipe, test bench, and data acquisition system.

- **The overall project objectives include:**

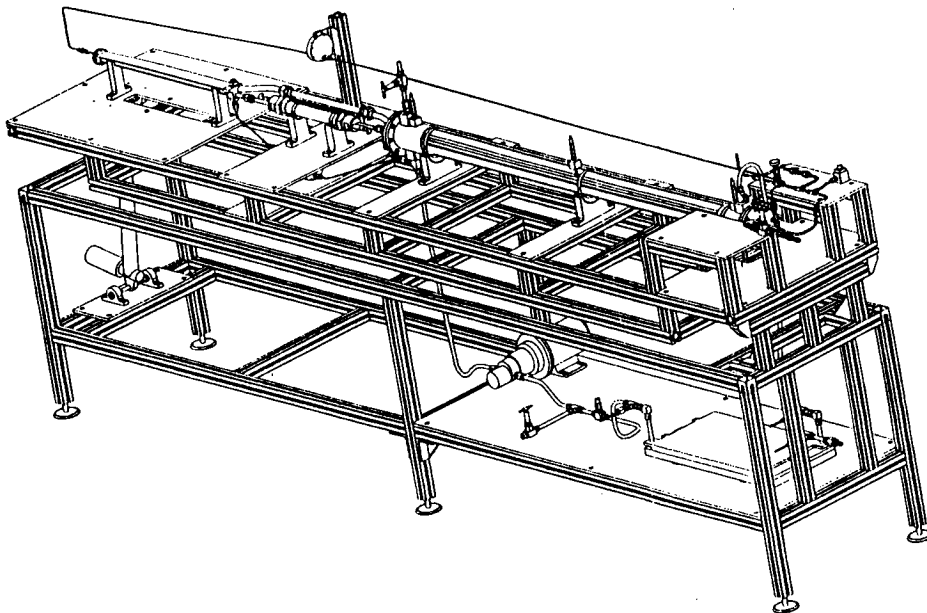
- to study/assess EHD and heat pipe capabilities,
- to develop an EHD/heat pipe database,
- to develop/validate EHD/heat pipe math models,
- to demonstrate/develop EHD applications for commercial use, and
- to develop a conceptual space flight experiment design.

NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

EHD Heat Pipe Test Bench



NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

Overview of the NASA JSC Program (cont'd)

- **Status**
 - Buildup of the test bench is complete.
 - MDA and NASA personnel have completed training on the LDV system.
 - Functional testing is underway.
- **Future Work**
 - Extended testing is planned with a variety of fluids and electrode configurations.

NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

Overview of the NASA JSC Program (cont'd)

- **Status**
 - The design PDR and CDR have been completed.
 - Fabrication/buildup are underway.
- **Future Work**
 - Buildup should be complete by the end of February 1994.
 - Extended testing is planned for this year.
 - FY95 efforts will focus on validation and enhancement of the models and testbed, and on development of the conceptual flight experiment design.

NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

Dual-Use Potential

- **Potential Dual-Use Applications include:**
 - **Heat Exchangers**
 - **Applications:** plate fin, shell-and-tube, tube-and-tube
 - **Markets:** space - station, orbiting platforms
industry - refrigeration, process, automotive
 - **Pumped Loops**
 - **Applications:** heat pipes, active one- or two-phase thermal control systems, electrical transformers
 - **Markets:** space - orbiting platforms, satellites, attached payloads
industry - refrigeration, utilities

NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

Dual-Use Potential

- **Static Mixing Technology**
 - **Applications:** liquid-liquid, liquid-gas
 - **Markets:** space - EHD pumping augmentation
industry - processing
- **Electronics Cooling**
 - **Applications:** electronics boxes, chips and circuits
 - **Markets:** space - satellites, attached payloads, STS payloads
industry - computer, test and measurement
- **Several companies (Allied Signal Aerospace, Carrier Air Conditioning, and Koch Engineering) are already looking at commercial use of this technology.**

NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

Dual-Use Potential

- For more information, please contact:
 - MDA, Mr. James Bryan, (713) 283-1853
 - Texas A&M University, Dr. Jamal Seyed-Yagoobi, (409) 845-4063
 - NASA JSC, Katy Miller, (713) 483-4546

NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

References

Babin, B.R., Peterson, G.P., and Seyed-Yagoobi, J., "Experimental Investigation of an Ion-Drag Pump-Assisted Capillary Loop," *Journal of Thermophysics and Heat Transfer*, Vol. 7, No. 2, April-June 1993.

Castaneda, J. A., and Seyed-Yagoobi, J., "Electrohydrodynamic Pumping of Refrigerant 11," *Proceedings of the IEEE-IAS Annual Meeting, Dearborn, Michigan, 1991.*

Seyed-Yagoobi, J., "A Theoretical Study of Induction Electrohydrodynamic Pumping in Outer Space," *Transactions of the ASME, Journal of Heat Transfer*, Vol. 112, November, 1990.

Yabe, A., "Active Heat Transfer Enhancement by Applying Electric Fields," *ASME/JSME Thermal Engineering Proceedings*, Vol. 3, 1991.

Yamashita, K., et al., "Heat Transfer Characteristics on an EHD Condenser," *ASME/JSME Thermal Engineering Proceedings*, Vol. 3, 1991.

NASA Johnson Space Center

McDonnell Douglas Aerospace

Texas A&M University

CMIT

Session L1: HUMAN FACTORS AND HABITATION

Session Chair: R. Bond

HUMAN-COMPUTER INTERACTION
WITH MEDICAL DECISION SUPPORT SYSTEMS

524-52
7

Jurine A. Adolf*
Kritina L. Holden*

NASA/Johnson Space Center
Human Factors and Ergonomics Laboratory (HFEL)
Frances E. Mount, Manager

ABSTRACT

Decision Support Systems (DSSs) have been available to medical diagnosticians for some time, yet their acceptance and use have not increased with advances in technology and availability of DSS tools. Medical DSSs will be necessary on future long duration space missions, because access to medical resources and personnel will be limited. Human-Computer Interaction (HCI) experts at NASA's Human Factors and Ergonomics Laboratory (HFEL) have been working toward understanding how humans use DSSs, with the goal of being able to identify and solve the problems associated with these systems.

Work to date consists of identification of HCI research areas, development of a decision making model, and completion of two experiments dealing with "anchoring". Anchoring is a phenomenon in which the decision maker latches on to a starting point and does not make sufficient adjustments when new data are presented. HFEL personnel have replicated a well-known anchoring experiment and have investigated the effects of user level of knowledge. Future work includes further experimentation on level of knowledge, confidence in the source of information and sequential decision making.

INTRODUCTION

The Human Factors and Ergonomics Laboratory (HFEL) at Johnson Space Center has been investigating human-computer interaction (HCI) with Decision Support Systems (DSSs). The impetus for this line of research was the need to identify problems with decision support systems, which have resulted in their lack of use within the medical community. While DSSs are not unique to the medical domain, the focus on medical DSSs was selected because of NASA's interest in technologies which will support future long-duration space missions. Access to medical resources

and personnel will be extremely limited on long duration flights, thus the need for a DSS. This paper provides an overview of the approach used, a description of the work in progress, and future directions for a line of research on DSSs.

Medical DSSs are defined here as computer-based tools which provide diagnoses and associated rationale, probabilities, and strategies for further investigation based on data input by the user (e.g., medical professional) and available in existing databases. These

* Lockheed Engineering and Sciences Company, 2400 NASA Rd. 1, MC-C81, Houston, TX 77058

systems do not replace the medical expert, but act as collaborator in the decision making process. Thus, the human remains "in the loop". Rouse (1988) describes this concept as "adaptive aiding", whereby the system provides aiding only at the points in time when assistance is needed. Advantages in using a DSS are increased speed and accuracy of diagnosis, and the avoidance of systematic biases and errors characteristic of the human decision making process. Even so, the acceptance and use of DSSs has been minimal.

The literature indicates that the lack of use of DSSs may be largely due to poor interface design. Kassirer, et. al. (1987) stated that while DSSs have made decision analysis effective and useful, they are "hard to use". Others have cited human-computer interface design as a major factor in the success of DSSs [Sandblad, Lind and Schnieder (1986); Shultz and Brown (1989); Luedtke and Luedtke (1984); Miller and Masarie (1990)]. Thus, the present line of research is aimed at understanding decision making with DSSs and providing interface design solutions. Although still in the early phases, accomplishments thus far include the identification of HCI research areas, the development of a decision making model and the completion of two laboratory experiments.

IDENTIFICATION OF HCI RESEARCH AREAS

After reviewing existing literature on DSSs, eight major areas of HCI research were identified. These include (1) Decision making errors, (2) Communication variables, (3)

Explanation facility issues, (4) Decision making issues under emergency conditions, (5) Trust/acceptance variables, (6) Time-consuming data input problems, (7) Medical terminology diversity and (8) Temporal-based reasoning issues. While all of the areas are candidates for HCI-based solutions, the first one in the list, "Decision making errors", was selected for the initial effort.

DEVELOPMENT OF A HUMAN DECISION MAKING MODEL

After reviewing existing literature concerning human decision making and decision making errors, a conceptual model of human decision making (see Figure 1) with the aid of a DSS was developed. The model provided a framework for discussion, and a visual representation of where in the process decision making errors occur. Figure 1 shows that while computer-aided decision making involves a larger information base, and eliminates many human decision making errors, some of these errors may actually be introduced or increased because of the way information is presented to the user. This is the problem of interest in the current line of research.

SELECTION OF DECISION MAKING ERRORS TO INVESTIGATE

Due to existing anecdotal information in the literature and interest within the medical community, the human decision making error known as "anchoring" was selected for initial investigation. Anchoring is a phenomenon in which the decision maker latches on to a starting point and does not make sufficient adjustments when new data are presented. For example, a physician makes an initial

diagnosis based on the data at hand. When new independent data (e.g. lab tests) are presented that may contradict the initial data, there is a tendency for the physician to persevere in believing the initial data (i.e., anchoring), and ignore the updated data. The result is a

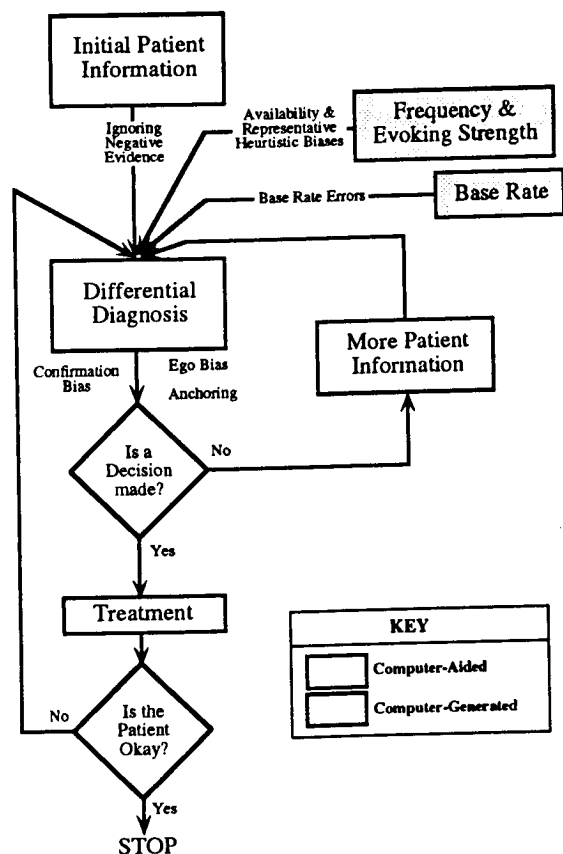


Figure 1. Medical Decision Making Model

faulty diagnosis, or decision making error. Anecdotal evidence for anchoring was found in a study by the Naval Submarine Medical Research Laboratory (D. M. Stetson, personal communication, 1990). The study was designed to investigate display design issues, not anchoring per se. Medical corpsmen were presented with rank-ordered likely diagnoses on a computer screen, accompanied by numerical probability

values. Researchers found that corpsmen were “paying too much attention” to the numbers, and ultimately had to remove them, leaving a rank-ordered list.

Tversky and Kahneman (1974) coined the term “anchoring” in their landmark demonstration, but the same phenomenon has been found in other studies as well (Roby, 1967; Peterson & DuCharme, 1967; Perrin, et al., 1993). Although the goal of the present research is to investigate anchoring within a medical framework, obtaining subjects with medical expertise is difficult and costly. Instead, a building block approach was adopted, such that initial experimentation to answer basic questions is accomplished outside of the medical domain, with experiments increasing in fidelity as the experimental methodology matures. Components of a real world task are incrementally added as experimentation progresses. The first step in investigating anchoring is to demonstrate its existence and identify the factors which affect it. Only then can potential solutions such as training, warnings or improved interface design be evaluated.

TVERSKY AND KAHNEMAN'S (1974) ORIGINAL DEMONSTRATION

Tversky and Kahneman (1974) asked subjects to estimate various quantities in percentages (e.g. percentage of African countries in the United Nations). A wheel of fortune was spun in the subject's presence and the subject was asked to indicate whether the number was higher or lower than the value of the quantity estimated by the subject. The subject was then asked to give an estimate of the quantity. Different groups of subjects were given different

starting values (from the wheel of fortune). The researchers found that the final estimates tended to vary with the starting value given. In other words, subjects anchored on the random number. Subjects receiving a starting value of 10, had a mean estimate of 25, and subjects having a starting value of 65, had a mean estimate of 45. This finding is very interesting, considering that the random number was totally unrelated to the question asked of the subjects.

EXPERIMENT 1 - REPLICATION OF TVERSKY AND KAHNEMAN (1974)

The first step in the investigation of anchoring was an attempt to replicate the Tversky and Kahneman (1974) findings. In addition, it was desirable to begin to try and understand why anchoring occurs. It was hypothesized that perhaps subjects anchored on the wheel of fortune value, simply because they were asked to use that value in the higher/lower judgment. In other words, they had to consider the wheel of fortune number in order to make their response. Experiment 1 was an opportunity to test this theory by NOT asking one group of the subjects to make any direct comparisons with the wheel of fortune number.

METHOD

Subjects consisted of 123 undergraduate psychology students at a local university. They participated in the experiment for course credit. Subjects were randomly assigned to one of two experimental groups. Materials for Group 1 represented a near replication of the Tversky and Kahneman (1974) wheel of fortune demonstration. These materials consisted of a sheet of paper with an introductory statement about a

wheel of fortune and two questions (see Figure 2). The wheel of fortune statement and the two questions were the same as the ones used in the Tversky and Kahneman (1974) original demonstration. The introductory statement provided a random number between 0 and 100, inclusive. This was a slight variation on the Tversky and Kahneman (1974) demonstration because they gave groups of subjects only one of two numbers (i.e., 10 or 65); whereas, the present materials provided a new random number for each subject. Materials for Group 2 were the same as for Group 1, except that Question 1 was omitted. In other words, these subjects were only asked to estimate the percentage, not to make a higher/lower judgment.

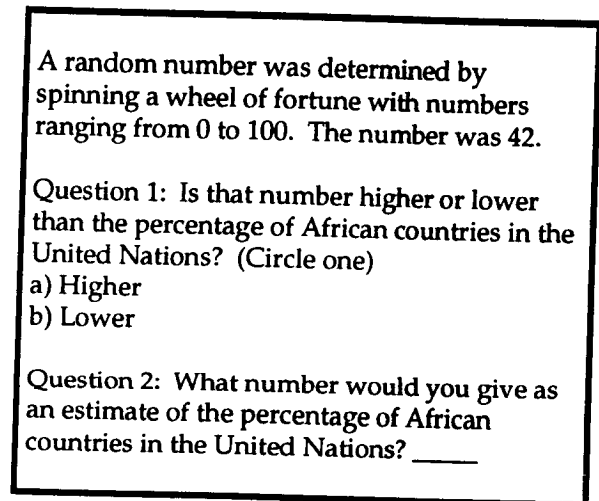


Figure 2. Materials for Experiment 1

RESULTS

Figure 3 shows a graph of the data for Groups 1 and 2, and includes an approximation of Tversky and Kahneman's (1974) original findings (slope estimated from the four data points discussed in their paper). Subjects' estimates are plotted against

the random numbers given. Note that perfect anchoring (i.e., selecting the wheel of fortune value as the estimate) would be demonstrated by a slope of 1. Complete absence of anchoring would be demonstrated by a slope of 0. The similarity in slopes between the original demonstration and Group 1 (0.27 vs. 0.25) indicates that the replication was successful; subjects do anchor on initial information, even when it is unrelated to the question posed. The slope for Group 2 is also very similar (0.31).

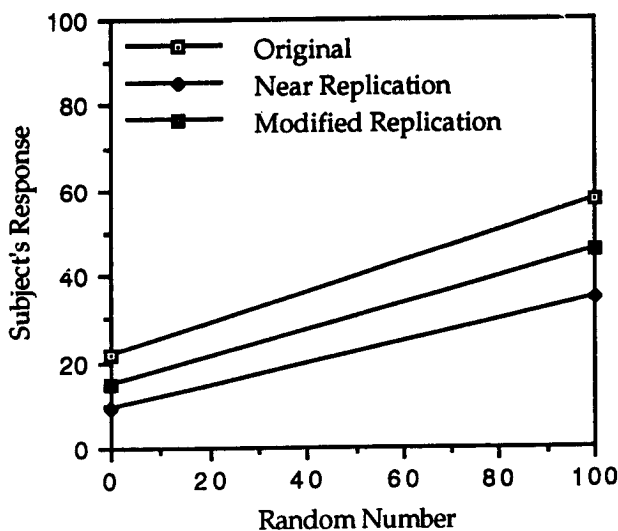


Figure 3. Tversky and Kahneman's (1974) Original Results and Experiment 1 Replications

The two measures of interest in this analysis are (1) the effect of Random Number and (2) the Random Number X Group interaction. A significant effect of Random Number would be direct evidence of anchoring. If no anchoring occurred, there should be no relationship between the random number and the subjects' estimate. A

significant Random Number X Group interaction would indicate that having subjects make a higher/lower comparison against the random number affects the amount of anchoring that occurs. An analysis of the data indicates a significant effect of Random Number ($p < 0.05$), but no significant Random Number X Group interaction. Thus, anchoring did occur, but there were no differences in anchoring for the two groups. The amount of variance accounted for in the analysis was 21.9%. While this is a low figure, the pattern of results are very similar to those found by Tversky and Kahneman (1974); thus the objective of Experiment 1 was met -- anchoring was demonstrated.

Having demonstrated that anchoring does occur, the next step was to identify and investigate variables that might affect anchoring. Given that the ultimate domain of interest in these investigations is the medical domain, the issue of trust of the decision support system became a topic of discussion. Two variables were identified as relevant to trust of the system: (1) level of knowledge of the user and (2) trustworthiness of the source of information. It was hypothesized that someone with a high level of knowledge (e.g., a medical doctor) will have less uncertainty and be less susceptible to anchoring on initial information provided by a DSS. Thus, the primary variable of interest in Experiment 2 was Level of Knowledge of the User. In addition to the effect of knowledge level, more anchoring would be expected when the source of information is very trustworthy. The source of information in Experiment 1 was a wheel of fortune, which would rate low on trustworthiness. In contrast,

Experiment 2 used a highly trustworthy source of information.

EXPERIMENT 2 - EFFECT OF USER LEVEL OF KNOWLEDGE ON ANCHORING

Since Level of Knowledge was the variable of interest, it was necessary to have a subject pool consisting of various levels of knowledge in some domain. Given the difficulty of finding a large pool of medical personnel willing to participate without compensation, a knowledge domain appropriate for a university student population was used for this initial investigation -- conservation.

METHOD

Subjects consisted of 98 undergraduate psychology students at a local university. They participated in the experiment for course credit. The number of levels of knowledge was set at three (e.g., High, Medium, and Low). The experimental materials were the same for all three groups. The materials consisted of instructions, a conservation pretest and an experiment sheet. The pretest consisted of multiple choice and True/False questions regarding conservation. It also included some rating scales (5 point Likert) used for obtaining demographic information. One of these questions asked subjects to rate their knowledge about conservation. The experiment sheet was similar to that used in Experiment 1 (see Figure 2). In contrast to the wheel of fortune statement, subjects were led to believe that an EPA representative estimated the percentage of landfill space taken up by paper. Subjects then responded with their estimate of the actual percentage. The higher/lower question (Question 1) used in

Experiment 1 was not repeated in this experiment because it was found to have no effect on anchoring.

RESULTS

The first task in compiling results was to compute pretest scores for all subjects. Scores ranged from 15% correct to 100% correct with a mean of 54%. A correlation was computed between the pretest scores and the self-rating score. The correlation ($r=0.07$) was not significant. Because the pretest score represented a number of data points, versus the one data point in the self-rating, the pretest score was selected as the criterion for subject assignment to Level of Knowledge groups. Pretest scores were plotted and subjects divided into three groups [i.e., High (N= 23), Medium (N= 52) and Low (N= 23) Level of Knowledge] based on an evaluation of the distribution of scores. Data for each Level of Knowledge group were plotted (see Figure 4).

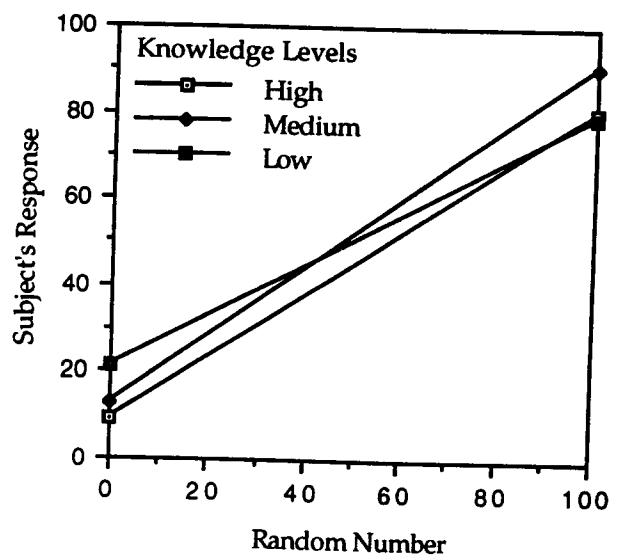


Figure 4. Anchoring Effect for Three Knowledge Levels

The slopes for all three groups are: High=0.64, Medium=0.85, Low=0.66. As in Experiment 1, a significant effect of Random Number would indicate anchoring, while a significant Random Number X Level of Knowledge interaction would indicate differences in the amounts of anchoring for each group. An analysis of the data reveals a significant effect of Random Number

The lack of a Level of Knowledge effect could indicate that: (1) there is truly no anchoring differences among the Level of Knowledge groups, or (2) the classification into Level of Knowledge groups has no validity. Explanation 2 may have some merit given the low correlation between the pretest scores and the self-rating. It could very well be that the subjects in Experiment 2 do not represent three levels of knowledge on conservation, but rather one large group with approximately the same amount of knowledge. Thus, the manipulation of Level of Knowledge is questionable, and follow-up research is required to either verify or nullify the results.

Trustworthiness of the Source of Information (Data from Experiments 1 and 2)

Figure 5 shows a graph of the data for Experiment 1 (N=123), accompanied by a graph of the data from Experiment 2 (N=98). This comparison was done to begin investigating the Trustworthiness of Source of Information variable. It is recognized that the two experiments used different experimental materials, but given that the procedures were identical, it is deemed worthwhile to compare the results of the two experiments.

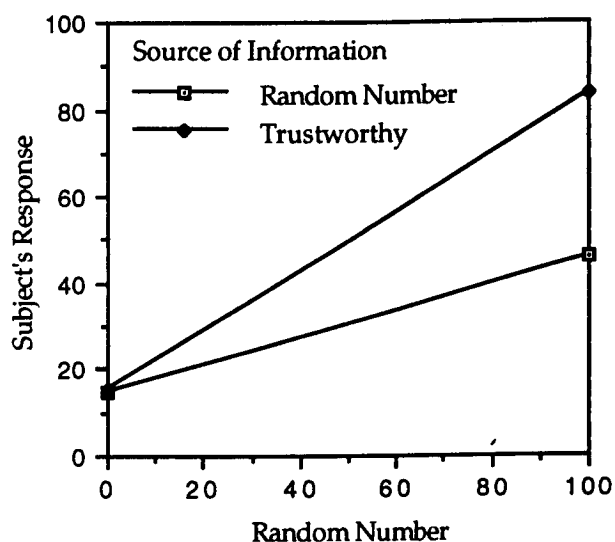


Figure 5. Trustworthiness of Source of Information (Data from Experiments 1 and 2)

Note that the Experiment 2 data have a much higher slope than the Experiment 1 data (0.75 vs. 0.25 respectively). A formal analysis reveals that this difference in slopes is significant ($p < 0.05$). These results would be consistent with the original hypothesis that a highly trustworthy source of information leads to a greater degree of anchoring. It is recognized that this was not a planned comparison, and the difference in experimental materials could have had some effect; however, the most likely explanation is that the trustworthiness of the source of information (i.e. wheel of fortune vs. a well-respected EPA representative) is responsible for the difference in slopes.

CONCLUSIONS

The experiments described empirically demonstrate that anchoring is a real phenomenon that can ultimately be the cause of decision making errors when subjects are working with decision support systems. In other words, even

when subjects were presented with random numbers that should have no bearing on their estimates (as in Experiment 1), anchoring occurred. The effect was much stronger when the initial information provided was from a highly trustworthy source (as in Experiment 2). Since there is some question about the validity of the Level of Knowledge manipulation in Experiment 2, a follow-up experiment is being planned to replicate Experiment 2, using a different method to identify and classify subjects by Level of Knowledge. A domain will be selected that will make it possible to identify true experts and novices. Future anchoring experiments will investigate trustworthiness of source of information and sequential decision making, and include the medical domain. The end product will be a set of guidelines for the design of DSS HCIs. Guidelines of this type have the potential for positively impacting not only medical DSSs, but DSSs in a number of domains. As we look toward long duration space flights, where astronauts will depend even more heavily upon technology, it is the early research, such as that described, that can make DSSs usable, trusted and accepted.

REFERENCES

- Kassirer, J.P., Moskowitz, A.J., & Lau, J. "Decision Analysis: A progress report," *ANNALS OF INTERNAL MEDICINE*, 106, 1987, pp. 275-291.
- Luedtke, P. & Luedtke, R. "Ease of Use," *COMPUTERS FOR MEDICAL OFFICES*, Harcourt, Brace, Jovanovich, NY, 1984, pp. 90-96.
- Miller, R.A. & Masarie, F.E. "Editorial: the demise of the "Greek oracle" model for medical diagnostic systems," *METH INFORM MED*, 29, 1990, pp. 1-2.
- Perrin, B.M, Barnett, B.J., & Walrath, L.C. "Decision making bias in complex task environments," *PROCEEDINGS OF THE HUMAN FACTORS AND ERGONOMICS SOCIETY 37TH ANNUAL MEETING*, 1993, pp. 1117-1121.
- Peterson, C.R. & DuCharme, W.M. "A primacy effect in subjective probability revision," *JOURNAL OF EXPERIMENTAL PSYCHOLOGY*, 73, 1, 1967, pp. 61-65.
- Roby, T.B. "Belief states and sequential evidence," *JOURNAL OF EXPERIMENTAL PSYCHOLOGY*, 75, 2, 1967, pp. 236-245.
- Rouse, W.B. "Adaptive Aiding for Human/Computer Control," *HUMAN FACTORS*, 30, 4, 1988, pp. 431-443.
- Sandblad, B., Lind, M. & Schrieder, W. "Requirements for human computer interfaces in health care," *HUMAN-COMPUTER COMMUNICATIONS IN HEALTH CARE*, North-Holland: Elsevier Science Publishers, North Holland, 1986.
- Shultz, E.K. & Brown, R. "Graphmaker: A Step in the Design of a Universal Interface for Hospital Information Systems," *PROCEEDINGS OF SCAMC*, 1989.
- Tversky, A. & Kahneman, D. "Judgment under uncertainty: Heuristics and biases," *SCIENCE*, 185, 1974, pp. 1124-1131.

DEVELOPMENT OF SHELF STABLE TORTILLAS FOR SPACE MISSIONS

Charles T. Bourland^{1*}, Vickie L. Kloeris¹, Michael F. Fohey²,
and Kimberly D. Glaus-Läte³

¹NASA Johnson Space Center, Houston TX 77058

²Lockheed Engineering and Sciences Company, Houston TX 77058

³Formerly Lockheed Engineering and Sciences Company, Houston TX 77058

ABSTRACT

Flour tortillas are a favorite bread item for the Shuttle astronauts and have been used on most Shuttle missions since 1985. Spoilage problems were encountered with commercial tortillas on missions longer than 7 days. A shelf stable tortilla with a shelf life of 6 months was developed by modifying the formulation to reduce the water activity (a_w) below 0.90 and packaging them in a reduced oxygen atmosphere. The water activity was reduced by substituting glycerin for some of the water in the basic tortilla formula. Reduction of the oxygen content was accomplished by packaging in a high-barrier container with a nitrogen atmosphere and including an oxygen scavenger in the package. Additional chemicals were added to the formula to lower the pH and further inhibit mold growth. The shelf life was verified by storage studies at 22° C. The shelf stable tortillas have been well accepted by astronauts and have been used on eight Shuttle missions with durations beyond 7 days.

INTRODUCTION

Tortillas are a flat, unleavened bread product traditionally used by Central and South American Indian populations and made using corn or wheat flour, water and lard. Wheat flour tortillas are the preferred bread item on Shuttle missions and have been used on most missions since 1985. Tortillas are an acceptable bread substitute due to ease of handling and reduced crumb generation in microgravity. Meat salad spreads, cheese spread, frankfurters, and peanut butter & jelly are some of the foods spread or folded into the tortillas to facilitate consumption. The tortillas are also used as a bread accompaniment to many of the entrees, i.e., beef tips in gravy and ham slices.

The Shuttle galley does not have refrigeration for food storage, hence all foods are stowed in locker trays at ambient temperature. Commercial

tortillas are purchased and delivered to the Kennedy Space Center and stowed in the Shuttle fresh food locker within 24 hours of the launch. The commercially produced tortillas have an ambient shelf life of only 5 to 7 days, after which time, yeast and mold growth appears. As mission length has increased, the need to develop a tortilla that would be shelf stable at ambient temperature became essential.

In order to develop an extended shelf life tortilla, it was necessary to limit the possibility of yeast, mold, and pathogenic bacteria growth. This required the development of a tortilla with a reduced water activity (a_w) and pH, packaged in an oxygen free environment. The extended shelf life bread developed by the U.S. Army Natick Research, Development and Engineering Center as part of the Meal, Ready-to-Eat (MRE) military ration was used as a model¹. The MRE bread has an a_w of less than 0.90 and is packaged in a three-ply foil laminate pouch with an oxygen absorber. The effectiveness of oxygen absorbers in inhibiting mold growth on pouched bread was demonstrated by Powers and Berkowitz². The MRE bread has a shelf life of three years at ambient temperature. The extended shelf life tortilla was developed by modifying the basic formula for tortillas to reduce the a_w and lower the pH to prevent bacterial growth, and packaged in an oxygen free environment to prevent mold growth.

MATERIALS AND METHODS

Several combinations of glycerin were evaluated by mixing various levels into the basic tortilla formula, evaluating the sensory properties and measuring the a_w of the finished product. Additional chemicals were added to lower the pH and inhibit mold growth. All ingredients used in the formulation are approved by the U.S. Food and Drug Administration. A_w was measured using a Decagon Devices CX-1 Water Activity System. Moisture was determined by the vacuum oven method. Residual headspace oxygen in the sealed packages was measured using a Mocon-Toray LC 700F Oxygen Analyzer. Sensory scores were determined by a laboratory taste panel. In the sensory test, tortillas were rated on appearance, color, odor, flavor, texture, and overall using a 9 point hedonic scale where 9 equals "like extremely" and 1 equals "dislike extremely." Utilizing these techniques, a successful formulation was developed. The formulation and the processing procedure are included in Appendix A. A year long storage study was conducted to determine shelf life. The packaged tortillas were stored at 22° C and tested at monthly intervals for sensory quality, pH, a_w , headspace oxygen, and microbiological quality.

RESULTS AND DISCUSSION

Water Activity (a_w) is defined as the ratio of the vapor pressure (P) of water in the food to the vapor pressure of pure water (P_0). This relationship is expressed by the equation: $a_w = P/P_0$. A level of 4.02 % glycerin was found to be the optimum for obtaining a a_w below 0.90 and retaining acceptable sensory scores. Reducing the pH further inhibits the potential for growth of anaerobic pathogens⁴.

Results of the initial tests performed on the tortillas produced from the new formulation were: Oxygen headspace 0%, pH 5.1, a_w 0.89, and moisture 26.3%. Microbiological counts were zero per gram (0/g) for total aerobic, yeast and mold count, and negative for coliforms, coagulase positive staphylococcus, salmonella and *Clostridium perfringens*. There were no significant changes in these values during the 1 year storage study. Sensory scores diminished with storage (Table 1). As a general rule, foods that receive an overall score of 6 or less are not qualified as candidate Shuttle foods. Sensory tests were terminated at the end of 9 months when it was obvious that the overall score was below 6.0.

TABLE I – OVERALL SENSORY SCORES FOR STORED TORTILLAS

<u>Months of Storage</u>	<u>Score</u>	<u>Standard Deviation</u>
0	7.00	0.62
1	7.18	0.60
2	6.67	0.65
3	7.31	0.37
4	6.94	0.30
5	5.89	0.78
6	6.23	0.75
7	5.70	1.14
8	5.72	0.67
9	5.68	1.25

The sensory data indicates that the formulation for the extended shelf life tortilla is acceptable through six months. After that, a bitter off-flavor developed which lowered the acceptability. This off-flavor was more objectionable to some panelists as noted by the increase in the standard deviation at 7 months. The bitter off-flavor is believed to be related to the lipid phase of the tortillas, but this has not been investigated.

The shelf stable tortillas were analyzed for nutrient content and compared to the reference tortilla in Nutritionist III⁵ software. These comparisons are shown in Table 2.

TABLE II – TORTILLA NUTRIENT CONTENT*

<u>Nutrient</u>	<u>Shelf Stable</u>	<u>Reference Tortilla</u>
Kcal	240	238
Protein	7.9 g	6.3 g
Carbohydrate	45 g	43 g
Fat	4.7 g	4.5 g
Dietary Fiber	3.4 g	1.9 g

* Per 2 tortillas (75g)

The shelf stable tortilla was higher in protein and dietary fiber than the reference tortilla. Other nutrients were similar.

The shelf stable tortillas have been used on eight Shuttle missions where mission duration exceeded the shelf life of fresh tortillas. Although the shelf stable tortillas have a slight sweet taste, they are a popular food item for the Shuttle astronauts. Fresh tortillas are still used on the early portion of long duration flights, with the shelf stable being used for mission days beyond seven.

CONCLUSIONS

A shelf stable tortilla was developed using reduced a_w to control bacterial growth and reduced oxygen to eliminate mold growth. The shelf stable tortillas have an effective shelf life of 6 months when stored at room temperatures. The tortillas have been used on eight Shuttle missions and are well accepted by astronauts. There is a high probability that the shelf life could be extended by modifying the fats used in the formulation. Shelf stable tortillas have a potential use in many areas including military, emergency foods, and remote operations.

REFERENCES

1. Military Specification: Bread, Shelf Stable, for Meal, Ready-to-Eat. MIL-B-4436-0A, March 1993.
2. Powers, E.M. and D. Berkowitz, "Efficacy of an Oxygen Scavenger to Modify the Atmosphere and Prevent Mold Growth on Meal, Ready-to-Eat Pouched Bread", J of Food Protection, vol 53, 1990, pp767-771.
3. Desrosier, N.W. and J.N. Desrosier, "Principles of Semi-moist Foods", The Technology of Food Preservation, Avi Publishing, Westport CT, 1977, pp278-299.
4. Baird-Parker, A. C. and B. Freame, "Combined Effect of Water Activity, pH and Temperature on the Growth of *Clostridium botulinum* from Spore and Vegetative Cell Inocula," J Applied Bacteriology, vol 30, pp420-429.
5. Nutritionist III. N-Squared Computing, Salem OR.

APPENDIX A: TORTILLA PROCESSING PROCEDURE

TORTILLA FORMULATION

<u>Ingredient</u>	<u>% By Weight</u>
Wheat flour	61.79
Water	26.58
Glycerin*	4.02
Shortening**	3.71
Mono/Diglycerides†	1.24
Salt	0.99
Baking powder	0.87
Dough conditioner††	0.31
Fumaric acid	0.19
Potassium sorbate	0.15
Carboxymethyl cellulose	0.12
Calcium propionate	<u>0.03</u>
Total	100.00%

* Optim 99.7%, Dow Chemical, Midland MI

** Fluid Flex, Van Den Burgh Foods Co., Joilet IL

† BFP 65, American Ingredients Co., Kansas City MO

†† Reduce 150, American Ingredients Co., Kansas City MO

Dry ingredients were combined in a Kitchen Aid mixer using the wire beater attachment on the "stir" setting for 1 minute. Shortening and mono/diglycerides were then added and blended to cornmeal consistency, about 3 - 5 minutes using the wire beater attachment on speed 2. Fumaric acid and potassium sorbate were weighed separately into 100 ml water and set aside. Glycerin and the remainder of water were combined and added to the mix and the wire beater was replaced with the dough hook. The fumaric acid and potassium sorbate solutions were added as the dough was mixed on speed 2. Mixing continued for approximately 10 minutes. After mixing, the dough was allowed to rest 5 min and then divided into 32 equal portions using a dough divider.

The divided dough was rounded by hand and placed into muffin pans and covered with plastic wrap. Dough balls were then placed in a 35.5° C proofing chamber for 1 - 2 hrs. Each dough ball was lightly dusted with flour and then formed in a Mini Wedge Flour Tortilla Press (BE & S Co, San Antonio TX). The pressed tortilla was then placed in a heated frying pan (190 to 204° C). When the uncooked surface began to bubble, the tortilla was flipped to cook the other side. After both sides were baked, the

tortillas were moved to a cool surface lined with waxed paper and allowed to cool. After a few minutes, the tortillas were turned to prevent condensation from forming between the waxed paper and the tortillas.

After cooling to room temperature, two tortillas were folded in half separately and placed in a three-ply foil laminate pouch, o.d. 6 1/2" x 8 1/8" (Reynolds Metals, Plano TX). An oxygen absorber (Multiform Desiccants, Buffalo NY) was inserted into each pouch immediately prior to beginning the sealing operation. The filled pouch was placed in a vacuum seal chamber where it was back-flushed with nitrogen 3 times before sealing at 10 in. Hg vacuum.

LARGE-SCALE NUMERICAL SIMULATIONS OF HUMAN MOTION

Frank C. Anderson, James M. Ziegler, and Marcus G. Pandy
*Department of Kinesiology, Department of Mechanical Engineering, and Biomedical
Engineering Program, University of Texas at Austin, Austin, Texas 78712*

Robert T. Whalen
NASA-Ames Research Center, Life Sciences Division, Moffett Field, California 94305

ABSTRACT

This paper examines the feasibility of using massively-parallel and vector-processing supercomputers to solve large-scale optimal control problems for human movement. Specifically, we compare the computational expense of determining the optimal controls for the single support phase of walking using a conventional serial machine (a Silicon Graphics Personal Iris 4D25 workstation), a MIMD parallel machine (an Intel iPSC/860 comprising 128 processors), and a parallel-vector-processing machine (a Cray Y-MP 8/864). With the human body modeled as a 14 degree-of-freedom linkage actuated by 46 musculotendinous units, computation of the optimal controls for walking could take up to 3 months of CPU time on the Iris. Both the Cray Y-MP and the Intel iPSC/860 are able to reduce this time to practical levels. The optimal control solution for walking can be found with about 77 hours of CPU time on the Cray, and with about 88 hours of CPU time on the Intel. Although the overall speeds of the Cray and the Intel were found to be similar, the unique capabilities of each machine are best suited to different parts of the optimal control algorithm used. The Intel performed best in the calculation of the derivatives of the performance criterion and the constraints. In contrast, the Cray performed best during parameter optimization of the controls. These results suggest that the ideal computer architecture for solving very large-scale optimal control problems is a hybrid system in which a vector-processing machine is integrated into the communication network of a MIMD parallel machine.

INTRODUCTION

Determination of the musculoskeletal forces generated during human movement is one of the most important, unsolved problems in biomechanics. The difficulty arises because the ligaments and number of muscles spanning a joint form a structurally indeterminate system and straightforward methods of applying force and moment equilibrium cannot be used. To circumvent this difficulty, we have been combining optimal control theory with detailed musculoskeletal models of the human body to determine the time history of all muscle forces generated during movement. A major advantage of using an optimal control approach is that it incorporates musculotendon dynamics and muscle excitation-contraction dynamics into the formulation of the problem. Once an optimal control solution is found, not only is there a wealth of information available to compare with experimental data (i.e., limb motions, ground reaction forces, and muscle activation patterns), but, more significantly, the solution can be analyzed to understand various aspects of multi-joint coordination and muscle function during movement (Bobbert and van Ingen Schenau, 1988; van Soest et al., 1993).

The major difficulty in solving optimal control problems for human movement is the computational expense of accurately modeling the musculoskeletal system. Recently, a four-segment, eight-muscle model of the human body was developed and used to compute the optimal control solution for vertical jumping (Pandy et al., 1990). This study demonstrated the utility of combining musculoskeletal modeling with optimal control theory. However, it also

emphasized the need for more efficient computational algorithms and alternative computer architectures. With conventional serial machines, which have processing speeds of about 2 Mflops, computation of the optimal controls for jumping is manageable, requiring approximately 3 days of dedicated CPU time. In contrast, for an activity such as walking, which requires a much more detailed representation of the human body, computation of the optimal controls on a serial machine may require several months of dedicated CPU time. Fortunately, with the emergence of high-speed, vector- and parallel-architected supercomputers, it is now possible to use very high-dimensional dynamical models of the human body to accurately simulate human movement and determine musculoskeletal loading histories during daily physical activity.

The major goals of this paper are (i) to introduce efficient computational algorithms for solving very large-scale optimal control problems for human movement, and (ii) to assess the feasibility of computing optimal control solutions using massively-parallel and vector-processing supercomputers. To establish the size and complexity of the optimal control problems studied here, we begin by presenting the details of the musculoskeletal model used to simulate the single support phase of walking. We then describe two distinct computational algorithms for optimal control: one developed for execution on both serial and vector-processing computers, and the other designed specifically for MIMD-architected parallel computers. Finally, we evaluate the computational expense of determining the optimal controls for the single support phase of walking using a conventional serial machine (Silicon Graphics Iris 4D25), a parallel-vector-processing machine (Cray Y-MP 8/864 housed at the Center for High Performance Computing at The University of Texas), and a MIMD-architected parallel machine (an Intel iPSC/860 maintained at NASA-Ames Research Center).

METHODS

Musculoskeletal models

The skeleton was modeled as an 8-segment, 14 degree-of-freedom linkage, joined together by frictionless revolute (Fig. 1). The metatarsal joint of the stance leg, the ankle joints, and the knee joints were each modeled as a single degree-of-freedom revolute. The hip joints were each modeled as a 3 degree-of-freedom ball-and-socket joint. To separate the three-dimensional rotations of the pelvis from those of the HAT segment (head, arms, and torso), a 3 degree-of-freedom, ball-and-socket joint was placed at the L3 level of the spine (Tashman, 1992). To model the interaction between the foot and the ground, a highly-damped, stiff, nonlinear, torsional spring was placed at the toes. The skeleton was actuated by 46 musculotendinous units. The mechanical behavior of each unit was described by a Hill-type contractile element which modeled the muscle's force-length-velocity properties, a series-elastic element which modeled its short-range stiffness, and a parallel-elastic element which modeled its passive response. The elastic properties of tendon were modeled by a linear stress-strain curve. Driving each musculotendinous unit was a first-order model for muscle excitation-contraction dynamics which computes muscle activation given any value of the input neural excitation between zero (no excitation) and one (full excitation).

Parameters defining nominal muscle properties for each musculotendinous unit spanning the hip, knee, and ankle were estimated from Delp (1990). The linear stress-strain curve for tendon was specified using values of elastic moduli obtained from Alexander and Vernon (1975), Woo et al. (1982), and Butler et al. (1984), while cross-sectional areas were chosen to give a reasonable strain in tendon at peak isometric force (i.e., in the range 2-6%). Since no experimental data exist for tendon rest length, we adjusted this parameter for each actuator in the model until the total isometric active torque about each joint peaked at a joint angle corresponding to *in vivo* measurements of joint torque. Musculotendon origin and insertion sites for each actuator in the model were defined on the basis of data reported by Delp (1990) and Yamaguchi et al. (1990). Finally, body-segmental parameters (i.e., segment mass and length, moment of inertia, and location of the center of mass of each segment) were scaled

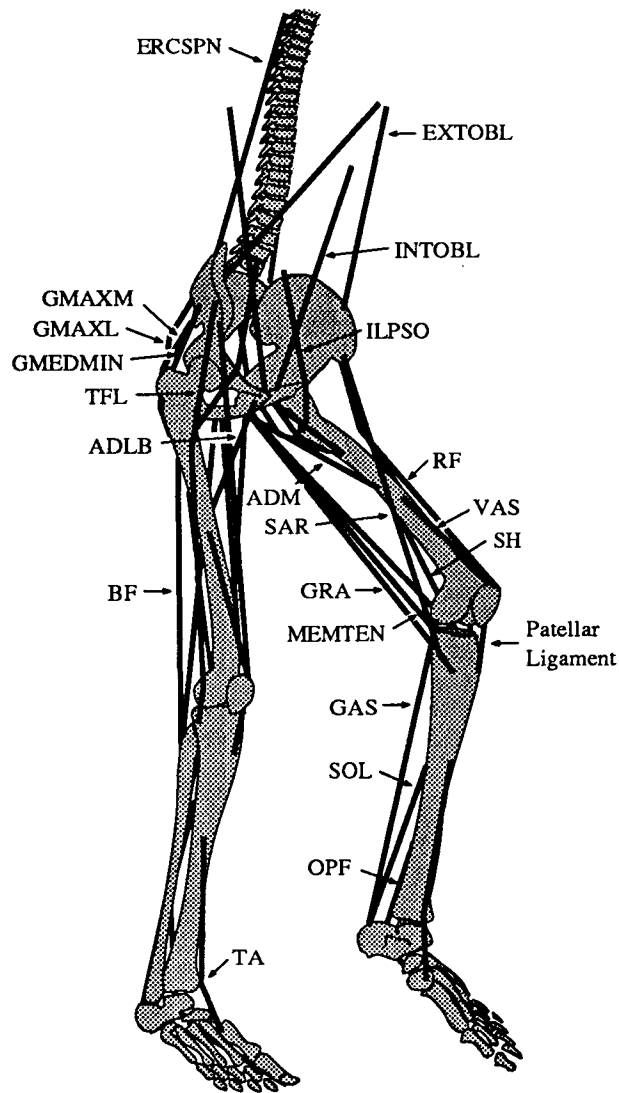


Fig. 1: Schematic representation of the musculoskeletal model used to simulate the single support phase of walking. The skeleton was modeled as an 8-segment, 14 degree-of-freedom linkage held together by frictionless joints and actuated by 46 muscles. Symbols appearing in the diagram are: tibialis anterior (TA), soleus (SOL), other plantarflexors (OPF), gastrocnemius (GAS), vasti (VAS), rectus femoris (RF), short head of biceps femoris (SH), semitendinosus and semimembranosus (MEMTEN), the long head of biceps femoris (BF), adductor magnus (ADM), iliopsoas (ILPSO), gluteus medius and gluteus minimus lumped together (GMEDMIN), medial portion of gluteus maximus (GMAXM), and lateral portion of gluteus maximus (GMAXL), tensor fasciae latae (TFL), adductors longus/brevis (ADLB), gracilis (GRA), sartorius (SAR), the erector spinae muscles (ERCSPN), the internal obliques (INTOBL), and the external obliques (EXTOBL). In addition, the model for walking included piriformis (not shown) and pectineus (not shown).

according to a 185 cm, 70 kg adult male using nominal data reported by Winter (1979).

Numerical simulation

A direct dynamics approach was used to simulate the single support phase of walking. Given the initial states of the model, body-segmental motions were produced by specifying the input neural excitation to each muscle. At each time step, the muscle activations, muscle forces, and body-segmental accelerations were found using the following differential equations:

$$\dot{a}_i = \frac{1}{\tau_{\text{rise}}} \{u_i^2 - u_i a_i\} + \frac{1}{\tau_{\text{fall}}} \{u_i - a_i\} \quad (1)$$

$$\dot{F}_i^{\text{MT}} = f(F_i^{\text{MT}}, \ell_i^{\text{MT}}, v_i^{\text{MT}}, a_i) \quad (2)$$

$$\ddot{\theta} = A(\theta)^{-1} \left\{ B(\theta) \dot{\theta}^2 + C(\theta) + M(\theta) \underline{F}^{\text{MT}} + \underline{T}(\theta, \dot{\theta}) \right\} \quad (3)$$

where i is the muscle number ($i=1,46$); u_i is the neural excitation sent to the i th muscle; a_i is the activation level of the i th muscle; τ_{rise} and τ_{fall} are rise and decay time constants for muscle activation; $\underline{F}^{\text{MT}}$ is the vector of musculotendon forces (46×1); ℓ_i^{MT} and v_i^{MT} are the musculotendon length and velocity of the i th actuator; $\theta, \dot{\theta}, \ddot{\theta}$ are 14×1 vectors of joint angular displacements, velocities, and accelerations; $M(\theta)$ is the moment-arm matrix (14×46); $A(\theta)$ is the system mass matrix (14×14); $C(\theta)$ is a vector containing only gravitational terms (14×1); $B(\theta) \dot{\theta}^2$ is a vector describing both Coriolis and centrifugal effects (14×1); and $\underline{T}(\theta, \dot{\theta})$ is a vector containing the external torque applied by the torsional spring to the foot segment when the heel is on the ground. The details of our models for muscle excitation-contraction dynamics (Equation (1)) and musculotendon dynamics (Equation (2)) are given in Pandy et al. (1992) and Pandy et al. (1990), respectively. Analytic expressions for the dynamical equations of motion for the skeleton (Equation (3)) were found using a commercial software package called Autolev (Schaechter et al., 1991).

The central problem in the direct dynamics approach is finding a set of neural excitations, \underline{u} , which produce muscle forces and limb motions similar to those present during actual human movement. Optimal control theory offers a methodology for solving this problem (Chow and Jacobson, 1971; Hatze, 1976; Davy and Audu, 1987; Pandy et al., 1990).

Optimal control problem

For walking, we minimized the time derivative of muscle force squared, summed over all the muscles and integrated over the duration of the activity,

$$J = \int_0^{t_f} \sum_{i=1}^{46} \left[\frac{\dot{F}_i^{\text{MT}}}{F_{oi}^{\text{M}}} \right]^2 dt. \quad (4)$$

Here, \dot{F}_i^{MT} is the time derivative of force in the i th musculotendon actuator and F_{oi}^{M} is the peak isometric force of the i th muscle. We chose this performance criterion because it appears to be well suited to modeling nonballistic movement (Pandy et al., 1993).

The constraints which define the optimal control problem for walking were the dynamical equations of motion (Equations (1)-(3)), a set of inequality constraints placed on the neural excitations,

$$0 \leq u_i \leq 1; \quad i = 1, 46, \quad (5)$$

and a set of terminal equality constraints which specify the joint angular displacements at time t_f , the instant of opposite heel strike,

$$\theta_j \Big|_{t_f} = \theta_{jf}; \quad j = 1, 14. \quad (6)$$

Note that the optimal control problem for walking is a *free* final-time problem.

Since a major goal of this paper is to evaluate computational performance in solving very large-scale optimal control problems for human movement, the precise forms of the performance criterion and constraints used in the simulation of walking have little bearing on the results reported below. Equations (4)-(6) serve only to illustrate the overall structure of the optimal control problem studied here.

Computational algorithms

We have developed two distinct computational algorithms for solving general optimal control problems for human movement: a serial algorithm which can be executed on both serial- and vector-processing computers, and a parallel algorithm designed specifically for MIMD-architected machines.

Each algorithm is based on the premise that any optimal control problem can be converted into a parameter optimization problem (see Pandy et al. (1992) for details). Converting the optimal control problem is accomplished by approximating the continuous muscle excitation histories with a set of discrete nodal points which become the independent variables in the resulting parameter optimization problem. By linearly interpolating between the control nodes, provided that the spacing is sufficiently small, the continuous excitation histories for each of the muscles can be reconstructed. To solve the parameter optimization problem, we implemented a sequential quadratic programming routine (Powell, 1978). This parameter optimization routine requires an evaluation of the performance criterion and the constraints, as well as an evaluation of the first derivatives of the performance criterion and constraints with respect to each of the controls.

Serial computational algorithm

A single iteration of the serial algorithm is comprised of three parts: a nominal forward integration to evaluate the performance criterion and the constraints, computation of the first derivatives of the performance criterion and constraints with respect to each of the controls, and execution of the parameter optimization routine (Fig. 2).

Evaluation of the performance criterion and constraints consists of a numerical integration of the state vector forward in time. This is accomplished using Equations (1)-(3) to compute the time derivative of the state vector at each time step. Once the values of the states at the final time have been determined, the performance criterion and constraints can be evaluated using Equations (4)-(6). All numerical integrations were performed using a Runge-Kutta-Feldberg 5-6 variable-step integrator which makes six evaluations of the time derivative of the state vector at each time step. Typically, the forward integrations required 600 to 1000 time steps.

Computation of the first derivatives of the performance criterion and constraints requires a series of forward integrations, each based on a slightly perturbed set of controls. The computational steps involved in this portion of the serial algorithm are (i) the *ith* control is perturbed, (ii) a forward integration is performed to determine the values of the perturbed performance criterion and the perturbed constraints, (iii) the *ith* control is returned to its original value, and (iv) the *ith* derivative is calculated based upon the values of the perturbed performance and constraints as well as on the values of the performance and constraints obtained from the nominal integration. So, for the walking problem a total of at least 737 perturbed forward integrations must be executed.

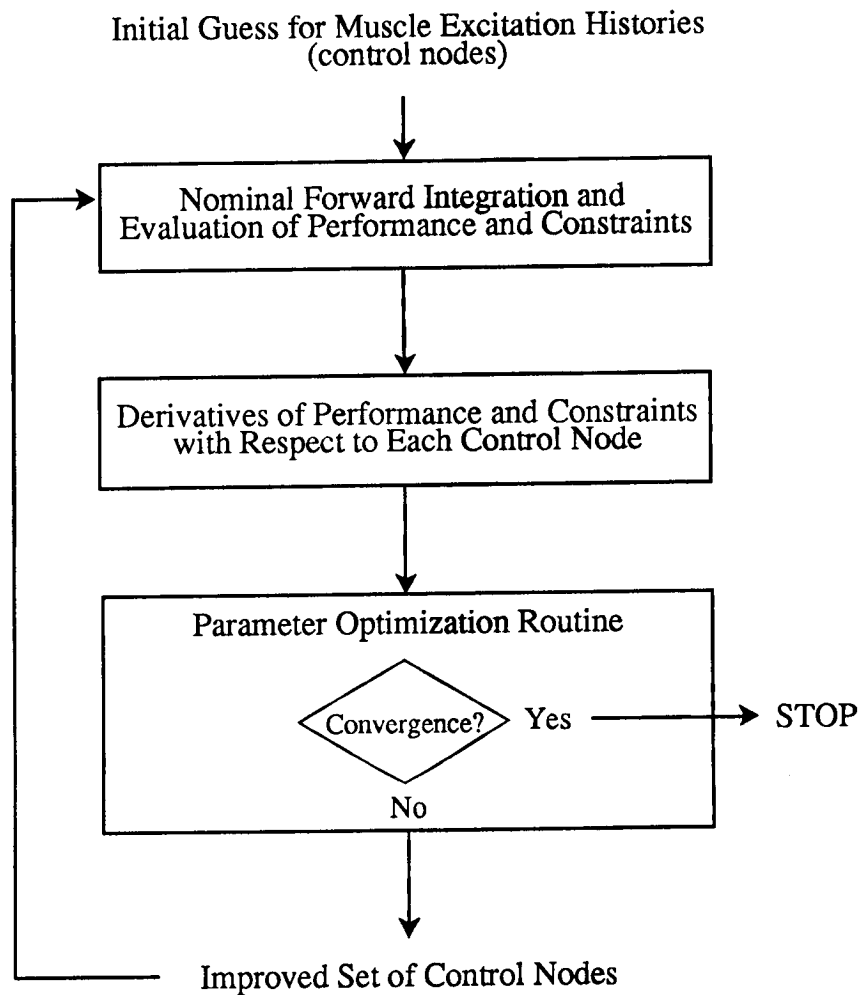


Fig. 2: Flowchart of the serial computational algorithm which can be executed on both serial and vector-processing computers. Each iteration comprises (i) an evaluation of the performance criterion and constraints, (ii) computation of the derivatives of the performance criterion and constraints with respect to each of the controls, and (iii) execution of the parameter optimization routine.

Finally, execution of the parameter optimization routine produces a new set of controls which improves the values of the performance criterion and/or the values of the constraints. The old controls are then replaced by the new controls and another iteration of the serial algorithm is begun. This procedure is repeated until the performance criterion is minimized and the constraints are satisfied (Fig. 2).

Execution of the serial algorithm on a vector-processing machine does not necessitate any change in the structure of the algorithm itself. High computational performance on the Cray Y-MP, which is composed of 8 Cray X-MP vector processors, is achieved primarily through code vectorization which means that whole variable arrays, rather than the individual elements of the arrays, are treated as the operands of floating point operations. Code vectorization is accomplished on the Cray by the cf77 compiler at the level of the innermost do-loops, and the extent of vectorization is proportional to the size of these do-loops (NAS User Guide). Therefore, the basic strategy for attaining high computational performance on the Cray Y-MP is to maximize the size of the innermost do-loops.

Parallel computational algorithm

The Intel iPSC/860 is a MIMD parallel computer which combines 128 individual i860 processors into a sophisticated communication network. Each i860 processor is rated at a peak speed of 60 Mflops double precision. As denoted by MIMD (Multiple Instruction, Multiple Data), each i860 processor of the Intel has its own memory and can work independently of the other processors. To execute code on the Intel, separate programs are loaded onto each i860 processor. The results from the various processors are then combined via communication routines furnished as extensions to C and Fortran. Therefore, to use the Intel effectively, computer code must be separated into independent parts and these parts distributed among the various i860 processors.

Fortunately, our optimal control algorithm can be distributed very efficiently among the individual processors of the Intel. The nominal integration is executed in the same way on the Intel as it is on a serial machine. The controls, \underline{u} , are passed to the first processor on which the nominal integration is performed in order to compute the values of the performance criterion and the constraints. The derivatives of the performance criterion and the constraints are then calculated in *parallel* by distributing the perturbed forward integrations among the available processors of the Intel (Fig. 3). The perturbed forward integrations can be performed in parallel because these computations are completely *independent*. That is, each forward integration requires only that the values of the controls and the initial states of the model be known.

Finally, the values of the performance criterion and constraints, as well as the first derivatives of the performance criterion and constraints, are input into the parameter optimization routine which returns an improved set of controls (see Fig. 3). Because the nominal integration and the parameter optimization routine cannot be executed in parallel on the Intel, these operations represent serial-code "bottlenecks" in our parallel computational algorithm.

RESULTS

The CPU times needed to compute the optimal controls for walking using the Iris 4D25, the Cray Y-MP 8/864, and the Intel iPSC/860 are given in Table I. One full iteration of the walking problem required 21.36 hours of CPU time on the Iris (Table I Iris, Total for walking). Assuming that convergence to the optimal control solution for walking requires 100 iterations, computation of the optimal controls would take about 3 months of dedicated CPU time on the Iris.

Most of the CPU time on the Iris is spent calculating the derivatives of the performance criterion and the constraints (Table I Iris, Numerical Derivatives). In particular, the Iris took 67,200 seconds (18.67 hours) to compute derivatives, which represents about 87% of the total

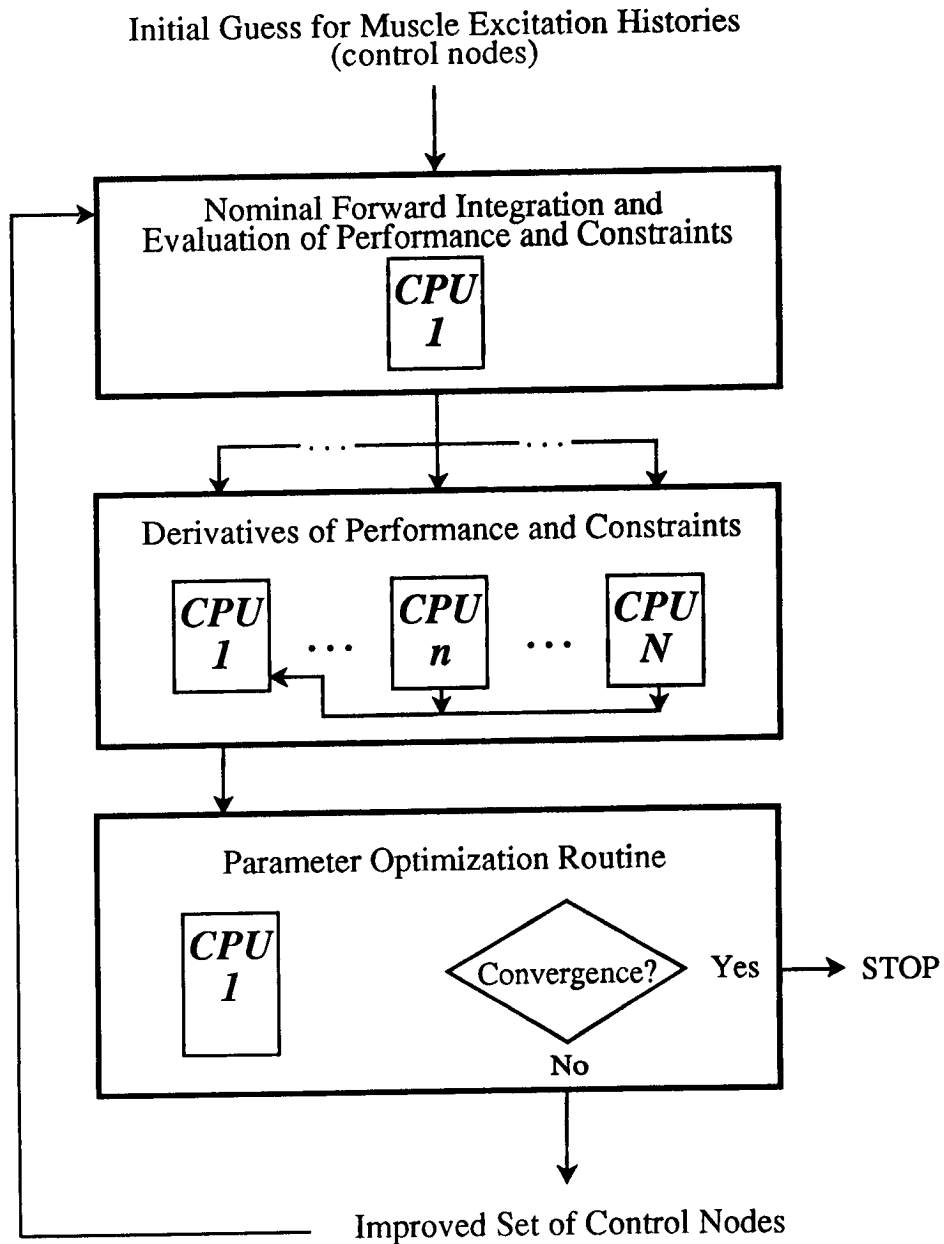


Fig. 3: Flowchart for the parallel computational algorithm. The derivatives of the performance criterion and the constraints are calculated in parallel by distributing the necessary forward integrations among the individual processors of the Intel iPSC/860 (CPU 1 through CPU N, where N can be as large as 128). Note that the nominal forward integration and the parameter optimization routine are not executed in parallel.

Table I: Comparison of CPU times for a single iteration of the optimal control algorithm (i.e., a nominal integration, numerical derivatives of the performance criterion and constraints, and optimization of the controls) for the single support phase of walking on an Iris 4D25, a Cray Y-MP 8/864, and an Intel iPSC/860. Derivatives of the performance criterion and constraints with respect to the controls were calculated using second-order numerical derivatives (central differences).

Computer	Nominal Integration (seconds)	Numerical Derivatives (seconds)	Optimization of the controls (seconds)	Total (hours)
Iris	150	67,200	9,540	21.36
Cray	5.5	2,500	260	0.77
Intel				
1 cpu	31	14,165†	2,990*	4.77
2 cpu	31	7,092†	2,990*	2.80
4 cpu	31	3,556†	2,990*	1.83
8 cpu	31	1,787	2,990*	1.34
16 cpu	31	915	2,990*	1.09
32 cpu	31	472	2,990*	0.97
64 cpu	31	247	2,990*	0.91
128 cpu	31	133	2,990*	0.88

† The CPU times for computation of the derivatives on the Intel iPSC/860 for 1, 2, and 4 processors are not actual profiles but instead are estimates based on the time required to execute a single forward integration. It was necessary to estimate these numbers because of the large amount of time required to calculate the derivatives with fewer than 8 processors.

* The parameter optimization routine could not be run on the Intel iPSC/860 because the individual i860 processors did not have sufficient memory. The numbers reported are the profiles for optimization of the controls on a supporting Silicon Graphics 4D480 located at NASA-Ames Research Center. The SG 4D480 is rated at 12 mflops in single processor mode.

CPU time needed per iteration. Since the time needed to converge to a solution for walking using conventional serial machines is prohibitive, alternative computing methods must be sought which focus on reducing the time taken to calculate derivatives.

Computation of derivatives

Although both the Cray and the Intel showed only modest improvements over the Iris in executing a nominal integration (Table I, Nominal Integration), significant improvements were made in computing the derivatives of the performance criterion and the constraints. The Cray reduced the time taken to calculate derivatives by a factor of 27 over the Iris (Table I Cray, Numerical Derivatives for walking). The Intel showed a similar improvement when 8 processors were used, and when all 128 processors were used, the Intel computed derivatives 505 times faster than the Iris (Table I Intel, Numerical Derivatives for walking).

The vast improvement in computational performance on the Intel is *not* due to the fact that each i860 processor is faster than the Iris. (For our code, each i860 processor ran up to 5 times faster than the Iris 4D25.) Rather, the large increase in computational speed is due to the fact that computation of the derivatives is ideally suited to the application of a MIMD-architected machine. Fig. 4 shows the speed-up of the Intel as the number of processors was increased from 1 to 128. The performance of our algorithm approximates ideal performance and scales almost *linearly* with the number of processors.

Parameter optimization of the controls

Computational performance of the Cray was clearly superior to that of the Iris and the Intel when performing parameter optimization of the controls (Table I). The Cray ran the parameter optimization routine 37 times faster than the Iris. The Intel iPSC/860 unfortunately did not have enough memory to execute the parameter optimization routine. Assuming that the Intel did have enough memory, we estimate that the Cray would have executed the parameter optimization routine about 15 times faster than the Intel. This estimate is based on a comparison of smaller parameter optimization problems which could run on both the Cray and the Intel. The Intel performs poorly relative to the Cray because the parameter optimization routine is not currently formulated to run in parallel on a MIMD architecture.

DISCUSSION

Although optimal control theory offers a potentially powerful methodology for determining musculoskeletal forces during movement (Hatze, 1976; Davy and Audu, 1987; Pandy et al., 1990), computing the optimal controls for large-scale musculoskeletal systems incurs great computational expense. In fact, for activities such as walking, which requires a very detailed, high-dimensioned model of the human body, computation of the optimal controls is currently impractical on conventional serial machines. Because optimization remains the only methodology available for determining muscle forces non-invasively (Seireg and Arvikar, 1973; Hardt, 1978; Patriarco et al., 1981; Yamaguchi and Zajac, 1990), we examined the feasibility of using parallel and vector-processing supercomputers to solve very large-scale optimal control problems for human movement.

Both the Cray Y-MP and the Intel iPSC/860 are capable of significantly reducing the time needed to solve optimal control problems characterized by large-scale musculoskeletal systems. The optimal control solution for the single support phase of walking can be computed with about 77 hours of CPU time on the Cray, and with about 88 hours of CPU time on the Intel when all 128 processors are used. These times are dramatically less than the 2136 hours (3 months) of dedicated CPU time required by the Iris 4D25.

Close examination of the performance of the Cray and the Intel reveals that each machine is best suited to executing different parts of the optimal control algorithm used (see Fig. 5). The MIMD architecture of the Intel makes this machine ideally suited to calculating the derivatives of the performance criterion and constraints; the Intel was able to calculate the derivatives 19

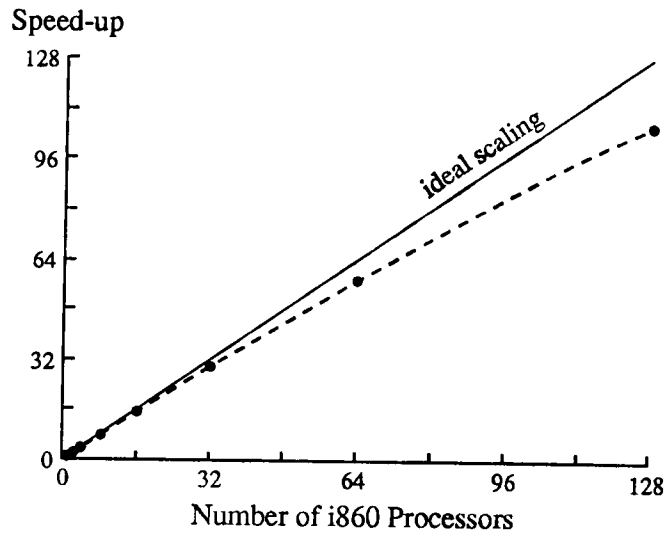


Fig. 4: Speed-up factor of the Intel iPSC/860 in the computation of the derivatives of the performance criterion and constraints for walking (dashed line) as the number of processors was increased from 1 to 128. The speed-up factor is multi-processor performance divided by single-processor performance. Ideal performance (solid line) was calculated under the assumption that as the number of processors is doubled, computational speed ought to be doubled. As indicated by the proximity of the dashed line to ideal performance, calculation of the derivatives is well suited to the MIMD architecture of the Intel.

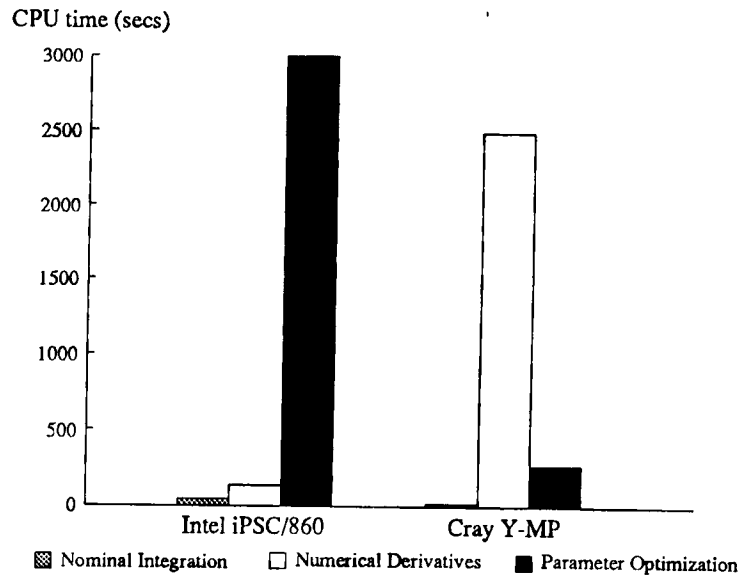


Fig. 5: Elapsed CPU time for each part of the optimal control algorithm when solving the walking problem on the Intel iPSC/860 and the Cray Y-MP. The architecture of each computer is best suited to different parts of the optimal control algorithm used. Due to its parallel-processing capabilities, the Intel was able to calculate the derivatives of the performance criterion and constraints much faster than the Cray (compare empty bars). On the other hand, the vector-processing capabilities of the Cray allowed this machine to execute the parameter optimization routine much faster than the Intel (compare black bars).

times faster than the Cray and 505 times faster than the Iris. On the other hand, because the parameter optimization routine is composed of many large do-loops, and because the Cray Y-MP derives its speed mainly from code vectorization, the Cray was able to execute the parameter optimization routine about 15 times faster than either the Intel or the Iris.

Although the demands of the parameter optimization routine currently make the Cray Y-MP more attractive in terms of overall computational performance, in the long term, MIMD-architected supercomputers such as the Intel iPSC/860 hold the greatest potential for reducing the time needed to solve very large-scale optimal control problems for human movement. Overall performance which surpasses the Cray can be attained on MIMD supercomputers by (i) formulating a parameter optimization routine which can be executed in parallel and/or by (ii) using a MIMD supercomputer which incorporates fast *vector*-processors into its communication network. Although a nonlinear programming routine tailored for execution on MIMD machines will likely yield the greatest improvement in computational speed, this option represents a major undertaking in terms of code development. Fortunately, supercomputers which integrate fast vector-processors into a MIMD architecture are already available. The Connection Machine-5, comprising 512 vector-processors, will not only allow the derivatives in our optimal control algorithm to be calculated in parallel, but it will also permit the parameter optimization routine to be executed at speeds which approach those attained on the Cray Y-MP.

ACKNOWLEDGEMENTS

We thank David Hull for helping with the development of the optimal control algorithm. Partial support for this work was provided by NASA-Ames Research Center, the Whitaker Foundation, and NASA Grant NAG5-2217. We also gratefully acknowledge the support of the Numerical Aerodynamic Simulation (NAS) Program at NASA-Ames Research Center and the Center for High Performance Computing at The University of Texas at Austin.

REFERENCES

- Alexander, R.McN. and Vernon, A., "The dimensions of knee and ankle muscles and the forces they exert," J. HUM. MVMT. STUD., 1, 1975, pp.115-123.
- Bobbert, M.F. and van Ingen Schenau, G, "Coordination in vertical jumping," J. BIOMECHANICS, 21, 1988, pp.240-262.
- Butler, D.L., Grood, E.S., Noyes, F.R., Zernicke, R.F., and Brackett, K., "Effects of structure and strain measurement technique on the material properties of young human tendons and fascia," J. BIOMECHANICS, 17, 1984, pp.239-306.
- Chow, C.K. and Jacobson, D.H., "Studies of human locomotion via optimal programming," MATH. BIOSCI., 10, 1971, pp.239-306.
- Davy, D.T. and Audu, M.L., "A dynamic optimization technique for predicting muscle forces in the swing phase of gait," J. BIOMECHANICS, 20, 1987, 187-201.
- Delp, S.L., "Surgery simulation: A computer graphics system to analyze and design musculoskeletal reconstructions of the lower limb," PhD thesis, Stanford University, Stanford, California, 1990.
- Hardt, D.E., "Determining muscle forces in the leg during human walking: An application and evaluation of optimization methods," J. BIOMECHANICAL ENGG., 100, 1978, pp.72-78.
- Hatze, H., "The complete optimization of a human motion," MATH. BIOSCIENCES, 28, 1976, pp.90-99.
- NAS USER GUIDE, Numerical Simulation Program, NASA-Ames Research Center, Moffett Field, California, 1992.

- Pandy, M.G., Garner, B.A., and Anderson, F.C., "Optimal control of non-ballistic muscular movements: A constraint-based performance criterion for rising from a chair," *J. BIOMECHANICAL ENGG.*, in press.
- Pandy, M.G., Anderson, F.C., and Hull, D.G., "A parameter optimization approach for the optimal control of large-scale musculoskeletal systems," *J. BIOMECHANICAL ENGG.*, 114, 1992, pp.450-460.
- Pandy, M.G., Zajac, F.E., Sim, E., and Levine, W.S., "An optimal control model for maximum-height human jumping," *J. BIOMECHANICS*, 23, 1990, pp.187-201.
- Patriarco, A.G., Mann, R.W., Simon, S.R., and Mansour, J., "An evaluation of the approaches of optimization models in the prediction of muscle forces during human gait," *J. BIOMECHANICS*, 14, 1981, pp.513-525.
- Powell, M.J.D., "A fast algorithm for nonlinearly constrained optimization calculations," In Matson, G.A. (ed.), *NUMERICAL ANALYSIS; LECTURE NOTES IN MATHEMATICS*, Springer-Verlag, 630, 1978, pp.144-157.
- Schaechter, D.B., Levinson, D.A., and Kane, T.R., *AUTOLEV USER'S MANUAL*, On-Line Dynamics, Inc., Sunnyvale, California, (1991).
- Seireg, A. and Arvikar, R.J., "A mathematical model for evaluation of forces in the lower extremities of the musculoskeletal system," *J. BIOMECHANICS*, 6, 1973, pp.313-326.
- Tashman, S., "Experimental analysis and computer modeling of paraplegic ambulation in a reciprocating gait orthosis: Implications for the design of hybrid FNS/Orthotic systems," PhD thesis, Stanford University, Stanford, California, 1992.
- van Soest, A.J., Schwab, A.L., Bobbert, M.F., and van Ingen Schenau, G.J., "The influence of the biarticularity of the gastrocnemius muscle on vertical jumping achievement," *J. BIOMECHANICS*, 6, 1993, pp.1-8.
- Winter, D.A., *BIOMECHANICS OF HUMAN MOVEMENT*, Wiley, New York, 1979, pp. 150-152.
- Woo, S., Gomez, M.A., Woo, Y., and Akesopn, W.H., "Mechanical properties of tendons and ligaments. II. The relationships of immobilization and exercise on tissue remodeling," *BIORHEOLOGY*, 19, 1982, pp.397-408.
- Yamaguchi, G.T., Sawa, A.G., Moran, D.W., Fessler, M.J., and Winters, J.M., "A survey of human musculotendon actuator parameters," In Winters, J.M. and Woo, S.L-Y. (eds.), *MULTIPLE MUSCLE SYSTEMS: BIOMECHANICS AND MOVEMENT ORGANIZATION*, Springer-Verlag, New York, 1990, pp.442-444.
- Yamaguchi, G.T. and Zajac, F.E., "Restoring unassisted gait to paraplegics via functional neuromuscular stimulation: A computer simulation study," *IEEE TRANS. BIOMED. ENGG.*, 37, 1990, pp.886-902.

327-54
1000

**ADVANCED LIFE SUPPORT SYSTEMS:
OPPORTUNITIES FOR TECHNOLOGY TRANSFER**

B. Fields¹, D. Henninger², D. Ming³, C. E. Verostko²

¹Lockheed Engineering & Science Company

NASA - Johnson Space Center

²Crew & Thermal Systems Division

³Solar System Exploration Division

ABSTRACT

NASA's future missions to explore the solar system will be of long-duration possibly lasting years at a time. Human life support systems will have to operate with very high reliability for these long periods with essentially no resupply from Earth. Such life support systems will make extensive use of higher plants, microorganisms, and physicochemical processes for recycling air and water, processing wastes, and producing food. Development of regenerative life support systems will be a pivotal capability for NASA's future human missions. A fully functional closed loop human life support system currently does not exist and thus represents a major technical challenge for space exploration. Technologies where all life support consumables are recycled have many potential terrestrial applications as well. Potential applications include providing human habitation in hostile environments such as the polar regions or the desert in such a way as to minimize energy expenditures and to minimize negative impacts on those often ecologically-sensitive areas. Other potential applications include production of food and ornamental crops without damaging the environment from fertilizers that contaminate water supplies; removal of trace gas contaminants from tightly sealed, energy-efficient buildings (the so-called sick building syndrome); and even the potential of gaining insight into the dynamics of the Earth's biosphere such that we can better manage our global environment.

Two specific advanced life support technologies being developed by NASA, with potential terrestrial application, are the zeoponic plant growth system and the Hybrid Regenerative Water Recovery System (HRWRS). The potential applications for these candidate dual use technologies are quite different as are the mechanisms for transfer. In the case of zeoponics, a variety of commercial applications has been suggested which represent potentially lucrative markets. Also, the patented nature of this product offers opportunities for licensing to commercial entities. In the case of the HRWRS, commercial markets with broad applications have not been identified but some terrestrial applications are being explored where this approach has advantages over other methods of waste water processing. Although these potential applications do not appear to have the same broad attraction from the standpoint of rapid commercialization, they represent niches where commercialization possibilities as well as social benefits could be realized.

INTRODUCTION

While the primary mission of NASA is typically articulated in terms of space science and exploration, historically the innovations in these fields have found broad terrestrial applications in improving the lives of the general population. Thus, the dual use concept has existed, albeit passively, as an adjunct to NASA activities either when commercial applications or some other benefit appears identifiable to people inside or outside of the space community. A myriad of examples are available in areas such as computers, materials science, and life sciences to support evidence of technology commercialization directly attributable to NASA development. Even when not immediately identifiable as having broad commercial appeal, applications with seemingly more narrow benefits often, over time, achieve success in a commercial niche; provide improvements in quality of life; or advance the state of science in a way that lays the groundwork for future progress. The current emphasis on Dual-Use applications for NASA developed technologies is an effort to forge a closer relationship between the work performed for space exploration and its relevance to the lives of the general population. Research in Advanced Life Support Systems (ALSS) at the Johnson Space Center is one field where this relationship is actively being explored.

OVERVIEW OF ADVANCED LIFE SUPPORT SYSTEM PROGRAM AT JSC

The ALSS program is an appealing candidate for identifying dual-use opportunities in view of its focus on fundamental human needs such as air, water, and food production, in environments where maximum efficiency and recycling are absolute necessities. NASA has established as one of its major strategic goals, the expansion of human presence beyond Earth.¹ Possible targets include bases on the Moon and Mars. The pursuit of this goal will entail research and development of technology necessary to support very long duration human missions.² Existing space life support technologies, currently in use for space missions, are typically non-regenerative and limited to missions with short durations and small crew size. These non-regenerative systems utilize expendables which require resupply from Earth. Technologies necessary to support very long duration human missions will require systems in which the air, water, and waste are recycled and food is produced rather than stored. Emphasis must be placed on the efficiency, durability, and reliability of such systems to enable their operation for long periods with minimum maintenance and resupply from Earth.³ Therefore, the ALSS Program is an important linchpin to achieving NASA's goal given the critical importance of developing these regenerative life support systems for future long duration missions.

The basic approach to Advanced Life Support is the development of a closed loop environmental system. Like the earth itself this closed system would contain and utilize all necessary constituents

to sustain life, in the presence of the sun's energy. However, various technologies will be necessary to accomplish this task. Identifying, developing, and demonstrating these technologies and integrating them into subsystems and systems is the thrust of current activities in the ALSS Program planning.

Two distinct life support development megatrends are emerging. The first is an Advanced Environmental Control and Life Support System in which components of the most advanced physicochemical systems are made compatible with and may be augmented by the growth of plants. Other features of this advanced ECLSS system would be enhanced air regeneration, enhanced water recycling, solid waste recycling, and regeneration of process catalysts. This system would be most applicable to early lunar missions. The second development need identified is for a fully Integrated Regenerative Life Support System in which there would be major dependance on biomass (plants) for food production and significant contribution for overall life support. This system would be dependant on the biological components to sustain the mission and the air, water, and waste management systems would be integral with biomass production. This system would be the choice for the later more mature lunar outpost and for permanent operation on Mars.⁴ The ALSS Program is currently in place to meet these technical challenges.

As technology development proceeds, in this area, a number of potential dual use applications are being identified as part of a Life Support technology transfer effort. Advances in atmosphere contaminant control, liquid-gas separation, thermal control, water recycling, and plant growth have each yielded technologies with potential terrestrial applications. Among these, two promising candidates, the Hybrid Regenerative Water Recovery System and the Zeoponic Growth System, are currently being considered for further study. They each represent distinct approaches to dual-use technology transfer.

HYBRID REGENERATIVE WATER RECOVERY SYSTEM

Space Based Application

For long duration exploratory space missions, the reclamation of water for potable and hygiene uses from waste water is of vital importance. Approximately 230 kg of waste water are generated from a four person crew during the course of normal daily activities. Efforts to recover this resource have focussed primarily on physicochemical methods such as phase change and/or membrane processes. It is likely, however, that because of the additional waste components generated in a Lunar/Mars habitat, some biological component may have sufficient advantages (low temperature, low pressure operation, etc.) over such physicochemical methods.⁵ Also, systems for

these applications must minimize the use of expendables such as pretreatment chemicals, adsorption media, etc. Thus, long duration manned space missions will require integrated biological and physicochemical processes, such as the Hybrid Regenerative Water Recovery System (HRWRS), for recovery of resources from waste water.

The current HRWRS is sized for a four-person crew and consists of the following components: 1) a two-stage, aerobic, trickling filter bioreactor; 2) a reverse osmosis system; and 3) a photocatalytic oxidation system. The system was designed to accommodate high organic and inorganic loadings and a low hydraulic loading.⁶ These loadings are typical of what would be found in a space application where water resources are conserved. The bioreactor was designed in-house to oxidize organics to carbon dioxide and water; the reverse osmosis system reduces inorganic content to potable quality; and the photocatalytic oxidation unit removes residual organic impurities (part per billion range) and provides in-situ disinfection.⁷

It has been demonstrated that a hybrid biological and physicochemical system is capable of treating waste water from shower, urinal, hand wash, laundry, and dishwasher sources and can produce potable water from such waste water.⁸ However, in developing this integrated system for space applications, the major difficulties encountered thus far, have been mechanical in nature; specifically, pump design for such low flow rates with such high suspended solids levels. Also, while the principle of operation has been proven, the batch nature of the photocatalytic oxidation system makes it unsuitable for complete system integration. A substitute post-treatment system is under construction at this time. Finally, system closure, for space based application, will be maximized by 1) recovery of the 15% brine from the reverse osmosis system; 2) closure of the gas loop from the bioreactor effluent; and 3) the development of a methodology for the treatment of bioreactor solids.⁹

Dual-use Features and Issues

Although challenges remain in creating a fully integrated system for space applications, the current system and its component technologies have dual use applications suited to terrestrial settings. The reverse osmosis membrane component is one example of a component technology in wide use. Non-space applications for reverse osmosis membranes are well documented, with these membranes being used in desalinization long before consideration in space use. There is currently wide commercial use of membrane technology for liquid as well as gas separation. So, in this respect reverse osmosis represents a spin-in technology adapted for space use. However, the requirement to produce an adequately small configuration for space use, which is deployable and allows for water recovery in remote locations, has resulted in NASA supported research of reverse

osmosis to meet these needs and has contributed to expanding its terrestrial applicability. In fact, the HRWRS is unique with respect to its small size and its ability to treat waste water, produced by only a few people, to high water quality over a long duration of time. These features and others incorporated into the HRWRS, which are essential to space application, such as low power requirements, sizing flexibility/modularity, and high reliability, result in attractive characteristics for remote terrestrial applications, as well.

Proposed terrestrial applications for the HRWRS include use for water recovery in remote locations such as:

- health clinics in remote Alaskan villages
- desert communities or drought areas
- isolated mountain areas
- in conjunction with military field hospitals for disaster relief or military exercises

Preliminary feasibility discussions have begun on ways to adapt the HRWRS to some of these situations. Related to these applications is the possibility of using such a system for single family or larger community developments in lieu of access to municipal facilities. In the case of the former, a company based in Colorado developed a system similar in concept to the the HRWRS for the single family residence. In the case of the latter, there are indications that advantages may exist for providing HRWRS type systems to land development either during the development process, prior to municipal system availability, or as a community based system.

The bioreactor component of the HRWRS is another area where wider terrestrial applications could develop. The concept of using biological processes to address industrial applications has grown in popularity, particularly as a result of advances in genetic engineering and the recognition of the usefulness of microbes for performing processing tasks. However, more work is needed to identify specific dual use applications in this area.

Among the issues which must be considered in applying the dual-use concept to HRWRS are the transfer mechanism and technology development gaps to commercialization. With respect to the transfer mechanism there are currently no patents or patents pending on this technology. Therefore, licensing this technology is not an option. As further development takes place, particularly focused on dual use applications, patents with commercial potential may be identified. However, current efforts to apply this technology have centered on niches geared to solving quality of life issues such as sanitation problems in remote Alaskan villages and use in deployable field hospitals for disaster relief. These are areas where needs exist. The type of market need necessary

for broader commercialization has not yet been identified. The transfer activity has involved discussions with other government entities and universities to define viable applications rather than with commercial entities interested in further development for profitable ventures. While these efforts may not immediately result in rapid commercialization, the solution to quality of life problems is recognized as a justifiable mechanism for transferring government based research and may eventually lead to applications with greater commercialization appeal.

Thus, the current efforts to transfer this technology may help in addressing the issue of technology commercialization gaps by acting as a vehicle to further terrestrial development work and expanded acceptance of the concept for terrestrial use. Some of the gaps which have been identified for the system include scale-up feasibility versus size requirements, extended reliability testing, cost reduction/standardization, autonomous operation, and design ruggedness. Once a clear set of requirements has been defined around a specific terrestrial application, further work will need to take place to address these gaps.

ZEOPONIC PLANT GROWTH SYSTEM

Space Based Application

Plants are being considered as an important component of regenerative life support systems for long-duration missions (e.g., Space Station, planetary outposts). In addition to supplying food, plants have the capability to regenerate air by converting CO₂ back into O₂ and, through evapotranspiration, plants can convert wastewater into potable water. However, the microgravity environment of space presents several problems for plant growth. One problem is providing a favorable root medium for plant growth. Nutrient delivery systems (i.e., hydroponic systems) are complex and require pumps and sophisticated control and monitoring systems. Therefore, it would be desirable to develop a substrate that provides the plant essential elements in a static watering system which eliminates circulating pumps and monitoring systems required for hydroponic systems.

To address this need, a highly-productive, synthetic soil (or substrate) has been developed for plant growth experiments. The synthetic soils have been termed zeoponic plant growth substrates. Zeoponics is defined as the cultivation of plants in zeolite mineral substrates that contain loosely-bound ions (e.g., Ca²⁺, K⁺, Mg²⁺) within their crystal structures. Zeolites have the capability to freely exchange some of their constituent ions with ions in solution without change to their structural framework. In addition to zeolites, the zeoponic plant growth substrates developed at the Johnson Space Center also consist of a synthetic calcium phosphate mineral (i.e., apatite). The

synthetic apatite has several essential plant growth nutrients incorporated into its structure (e.g., Ca, Mg, S, P, Fe, Mn, Zn, Cu, B, Mo, Cl). The substrate (consisting of the zeolite and synthetic apatite) has been designed to slowly release these plant-growth nutrients into 'soil' solution where they become available for plant uptake and supply all nutrients for many growth seasons with only the addition of water.

The latest research on zeoponic growth systems indicates success in meeting this objective. Wheat plants have been successfully grown in zeoponic substrates. Dry weight production of wheat was highest in the zeoponic substrates as compared to control substrates which were watered with nutrient solutions. The zeoponic substrates provided either adequate or slightly higher amounts of the primary nutrients required for plant growth and supplied all of the plant growth nutrients for extended growing seasons.

Dual-use Features and Issues

Zeoponic plant growth substrates have numerous possibilities for terrestrial applications particularly in fertilization and soil conditioning. In its form as a zeoponic growth system, combining zeolite and apatite based nutrients, specific applications which have been identified include use:

- as potting medium for greenhouses and house plants
- as an additive fertilizer for gardens
- on residential and commercial lawns (e.g., golf courses)
- in large scale agricultural production

From a commercial standpoint, the combined properties of zeoponics are attractive. As a fertilizer it has been shown to be very effective in supplying necessary plant nutrients. The Zeoponic approach appears to be an extremely efficient mechanism for providing plant nutrients by releasing those nutrients to the plant as a function of the plant needs. This attribute reduces the amount and frequency of nutrient application. Furthermore, its slow release capability prevents the leaching and run-off problems often associated with traditional fertilization methods. Thus, zeoponics has the advantage over traditional fertilizers of less frequent application, which can eliminate material and labor cost, as well as prevention of environmental damage.

In addition to these advantages, the zeolite acts as a soil conditioner. Once zeolite nutrients are depleted they can be replenished by the addition of chemical fertilizers which are adsorbed back into the zeolite substrate and subsequently slow released. Thus, besides preventing environmental

damage due to its slow release properties, zeolites can potentially be used in environmental clean-up. For example, it may be possible to use zeolite as an uptake medium for high nitrogen waste products and in turn apply it to an horticultural/agricultural situation as a fertilizer. Moreover, zeolite uptake characteristics may be applicable to a broader range of remediation problems.

The commercialization potential for zeoionics is promising based on indications from a number of sources in both the user and producer community. In general, fertilizers are a \$20 billion annual business. While it is still in early stage development, the zeoionic technology has captured the attention of several specialty market segments within this broad category including commercial greenhouses, nursery companies, golf course servicing, and others. Each of these represents a sizeable market by themselves which could benefit from the unique characteristics of zeoionics. Furthermore, the continuing concern for environmental effects on fragile ecosystems, such as the Everglades, due to the fertilization of adjacent agricultural areas will focus additional attention on strategies which can prevent these problems and eliminate the associated clean-up cost. Zeoionic technology is well positioned to take advantage of this focus. With patents pending on the Zeoionic system and the synthetic apatite micronutrient substance a licensing arrangement is a likely avenue for transferring this technology. A strong licensing agreement could help further the commercialization process by encouraging the additional R&D, and marketing necessary to create products which meet market needs that zeoionics is capable of satisfying.

Johnson Space Center is currently working on transferring this technology to the private sector through the licensing process. However, in spite of its attractiveness, zeoionics has some commercialization gaps which must be addressed. Key among these gaps is the fact that zeoionic growth medium has never been produced in sufficiently large amounts to determine production costs for commercial quantities. In fact, zeoionic growth medium is currently produced in small quantities, manually in the lab, with laboratory grade chemicals. An effort is underway for scale-up of production to increase the process automation and decrease material costs. The scale up is necessary simply to meet the need for increasing quantities of zeoionic growth medium to support ongoing research in Advanced Life Support. However, estimates indicate that scaling up the process to meet projected need by the Space agency and associated research, alone, will result in at least a 10 fold decrease in material cost through bulk processing. Such substantial decreases in cost for producing zeoionic growth medium can benefit the research activities as well as contribute to filling commercialization gaps for technology transfer. Increasing the available quantity of this material will also contribute to improved process control for material producibility, as well as increased research on the effectiveness and applicability of this material. Such activity will prove beneficial to both the long duration space missions and widespread terrestrial application.

Summary

As NASA's Advanced Life Support Systems Program continues its research and technology development to provide highly reliable human life support systems for future long-duration missions, the opportunities for technology transfer will grow and benefits to life on Earth will accrue. Human habitation systems for use in hostile or ecologically-sensitive environments; reducing ground water contamination from fertilizers applied to agricultural lands; providing pleasing and healthy atmosphere in tightly sealed buildings; and gaining insight into the only, continuously- operating regenerative life support system -- Earth, are only examples of opportunities for technology transfer.

Acknowledgements

The authors would like to express appreciation to Marybeth Edeen, Nigel J. C. Packham, Michael Johnson, and Charles Galindo for their help in preparing this paper and their technical contribution in the areas of Hybrid Regenerative Water Recovery Systems and Zeoionic Plant Growth.

Endnotes

1. NASA-Johnson Space Center, *Plan for a Regenerative Life Support System Development Program*, February 1991, 4.
2. Ibid.
3. Ibid.
4. Ibid., p. 28
5. Charles E. Verostko, Marybeth A. Edeen, and Nigel J. C. Packham, "A Hybrid Regenerative Water Recovery System for Lunar/Mars Life Support Applications", 22nd International Conference on Environmental Systems, *SAE Technical Paper Series #921276* (1992): 1
6. Ibid.
7. Ibid.
8. Ibid., 7
9. Ibid., 7-8

References

1. NASA-Johnson Space Center, *Plan for a Regenerative Life Support System Development Program*, February 1991.
2. Verostko, Charles E. , Edeen, Marybeth A., and Packham, Nigel J. C., "A Hybrid Regenerative Water Recovery System for Lunar/Mars Life Support Applications", 22nd International Conference on Environmental Systems, *SAE Technical Paper Series #921276* (1992).
3. Banks, Gary, and Rose, Susan, "A Phased Approach to the Evaluation of CELSS for Long Term Space Habitation", University of Colorado, Department of Engineering Sciences, Boulder, Colorado.
4. National Sanitation Foundation, "Pure Cycle: Performance Evaluation of the Pure Cycle Total Water/Waste Water Recycle System. Interim Report", for Pure Cycle Corporation, July 1982.
5. Ming, Douglas W., Henninger, Donald L., and Galindo, C. Jr, "Solid-Support Substrate for Plant Growth at a Lunar Base", Controlled Ecological Life Support System: CELSS '89 Workshop, *NASA Technical Memorandum 102277*, March 1990: 409-434.
6. Allen, E. R., Hossner, L. R., Henninger, D. L., Ming, D. W., and Galindo, C. Jr, "Growth and Nutrient Uptake of Wheat in a Zeoponic System", *Memoirs of the 3rd International Conference on the Occurrence, Properties, and Utilization of Natural Zeolites*. Edited by G. R. Fuentes and J. A Gonzalez, 1993, Volume 2: 14-19.
7. Allen, E. R., Hossner, L. R., Henninger, D. L., Ming, D. W., " Solubility and Cation Exchange in Phosphate Rock and Saturated Clinoptilolite Mixtures", *Soil Science Society of America Journal* , Volume 57 (September - October 1993): 1368-1374.

ant

Session L3: MEDICAL CARE

Session Chair: J. Homick

A Feasibility Study For Perioperative Ventricular Tachycardia Prognosis and Detection And Noise Detection Using a Neural Network and Predictive Linear Operators

T.A. Moebes, SAIC

Section 0.0 Abstract

To locate the accessory pathway(s) in preexcitation syndromes, epicardial and endocardial ventricular mapping is performed during anterograde ventricular activation via accessory pathway(s) from data originally received in signal form. As the number of channels increases, it is pertinent that more automated detection of coherent/incoherent signals is achieved as well as the prediction and prognosis of ventricular tachycardia(VT). Today's computers and computer program algorithms are not good in simple perceptual tasks such as recognizing a pattern or identifying a sound. This discrepancy, among other things, has been a major motivating factor in developing brain-based, massively parallel computing architectures. Neural net paradigms have proven to be effective at pattern recognition tasks. In signal processing, the picking of coherent/incoherent signals represents a pattern recognition task for computer systems. The picking of signals representing the onset of VT also represents such a computer task. We attacked this problem by defining four signal attributes for each potential first maximal arrival peak and one signal attribute over the entire signal as input to a back propagation neural network. One attribute was the predicted amplitude value after the maximum amplitude over a data window. Then, by using a set of known (user selected) coherent/incoherent signals, and signals representing the onset of VT, we trained the back propagation network to recognize coherent/incoherent signals, and signals indicating the onset of VT. Since our output scheme involves a true or false decision, and since the output unit computes values between 0 and 1, we used a Fuzzy Arithmetic approach to classify data as coherent/incoherent signals.

Furthermore, a Mean-Square Error Analysis was used to determine system stability. The neural net based picking coherent/incoherent signal system achieved high accuracy on picking coherent/incoherent signals on different patients. The system also achieved a high accuracy of picking signals which represent the onset of VT, that is, VT immediately followed these signals. A special binary representation of the input and output data allowed the neural network to train very rapidly as compared to another standard decimal or normalized representations of the data.

Section 1.0 Introduction

Human beings are extremely adept at pattern recognition tasks. For example, human response time for simple tasks such as recognizing a face or identifying a sound is on the order of 0.5 s. On the other hand, traditional computers using traditional programs, while extremely good at number crunching tasks, can

not approach the human response time in solving these simple pattern recognition problems. This discrepancy between the ability of the human brain and the ability of traditional computers has motivated researchers to develop computing paradigms that are massively parallel and brain-based. The neural network (NN) paradigm is one such massively parallel paradigm that has proved extremely good at pattern recognition tasks. NNs have been found to be applicable in a variety of application areas: pattern recognition (Lippmann, 1987, Fukushima, 1988; Carpenter and Grossberg, 1988), remote sensing (Benediktsson et al., 1990), signal processing (Widrow and Winter, 1988; Widrow and Stearns, 1985; Lapedez and Farber, 1987; Palmer, 1987), control systems and robotics (Psalids et al., 1987; Barhen et al., 1989), image processing (Daugman, 1988; Cottrell et al., 1987; Zhou et al., 1988), speech processing (Waibel et al., 1989; Lang, 1989; Sejnowski and Rosenberg, 1987), associative memory (Kohonen, 1979; Kosko, 1987), and optimization problems (Hopfield and Tank, 1985; Ramanujam and Sadayappan, 1988). Most of the commercial NN applications have been in pattern recognition and in particular many deal with determining the coherent versus the incoherent data (Urnes, Hoy, and Ladage, A Neural Based Intelligent Control System For The NASA F - 15 Research Aircraft June, 1992, NASA International Workshop On Neural Networks and Fuzzy Logic '92, and Desai and Jansen, Wave form Classification Using Artificial Neural Networks, 1992, NASA Third International Workshop On Neural Networks and Fuzzy Logic '92). Urnes reports that the problem has not been satisfactorily solved. In feature extraction, sensor fusion, image processing, communication, and docking and capture mechanisms we encounter a number of pattern recognition tasks as data is processed (e.g., first break picking, identification of sequence boundaries). Some researchers in the robotics systems and signal processing community have already started attacking some of these problems with NN techniques (McCormack, 1990; Poulton et al., 1989; Huang et al., 1989; Liu et al., 1989). Since NNs have been shown to be good at pattern recognition tasks, we explored the use of NNs for coherent/incoherent data picking on a computerized system for intraoperative cardiac mapping. This paper describes how we defined and trained a NN to pick coherent/incoherent signals on signal data from the cardiac mapping system, and were able to pick signals representing the onset of VT. First we give a brief introduction of the computerized system for intraoperative cardiac mapping and NNs. Then we detail how we set up the coherent/incoherent data picking problem for an NN solution using Fuzzy Arithmetic approach, indicate how the results of applying the NN picker to real data could be displayed and tested for stability using least-means squares analysis, and finally, offer conclusions on our initial proposal study with further applications to predicting VT. The cardiac mapping system used in this study was the Baylor College of Medicine's Intraoperative Cardiac Mapping System. The neural network systems used in this study was Science Applications International Corporation's (SAIC) Artificial Neural Systems Simulation (ANSim) program and NASA's NETS system..

Section 2.0 The Intraoperative Cardiac Mapping System

To locate the accessory pathway(s) in preexcitation syndromes, epicardial and endocardial ventricular mapping is performed during anterograde ventricular activation via accessory pathway(s) from data originally received in signal form. A sock array for epicardial ventricular recording, a balloon array for monitoring the endocardium, and atria plaque recording from the atria epicardium are used on the retired heart in locating geographically on and in the heart muscle for electrical pulses indicating ventricular tachycardia (VT). During VT the patient's heart beats erratically and very quickly results in the heart's failure to adequately pump blood.

When an array is positioned by the surgeon, the operator and cardiologist quickly view each data channel signal on the computer's monitor to ensure proper data collection indicating induced VT. As events of interest occur, pressing a single key causes storage to hard disk of a buffer containing the most recent 16 s of data. After the signals are collected, the computer system automatically aligns signals and eliminates bad signal data based on weak signals and other data qualification algorithms. The cardiologist then eliminates incoherent signal data by viewing the signals on the monitor and turning incoherent signal data off by hand operation. Then the signal data left as coherent is mapped by the system to an isochronal map where contours represent electrode locations on the heart connected according to the same time the VT signals arrived. Thus, the earliest arrival of the VT signals represent the geographical location area on the heart of the initiation of the VT. This area is then abated by the surgeon.

If the isochronal (contour) map is viewed as unsatisfactory by the cardiologist, part or the entire procedure is repeated before ablation. Also, several areas on the heart may need ablation. The whole mapping procedure may take 45 minutes of the two hours allowed for the patient on the heart-lung machine with the present technology and 200 channels. A newly designed system ready to go into operation has over 1000 channels. Thus, new technology to automate coherent/incoherent data detection is needed for the new system to be practical. However, the judgmental calls of the cardiologist must be preserved. This feasibility study showed that a NN approach to this problem would satisfy all of the new requirements. The research here also showed that certain outputs of the NN indicated the onset of VT. This results could be used to prognosis potential VT patients who have not experienced actual VT occurrences. The reader should see the paper "Design and implementation of a new computerized system for intraoperative cardiac mapping", by Carl F. Pieper, et al, for a detailed explanation of intraoperative cardiac mapping system.

Section 3.0 Neural Networks

NN research started at least four decades ago (for a collection of classic papers refer to Anderson and Rosenfeld, 1988) and culminated in the development of

the perceptron by Rosenblatt (1962) in the late 1950s. Subsequently, during the mid-1960s, interest in the field waned for a number of reasons, primarily due to the limitations of the perceptron as pointed out in Minsky and Papert (1969). In recent years, NNs have reemerged as an active research area not only within the neuro-science disciplines but also among computer scientists, physicists, mathematicians, and engineers. The main reason for the re-emergence is that new network paradigms of great practical significance have overcome some of the limitations of the perceptron. NNs are no longer an esoteric topic confined to academic settings but an emerging commercial technology. Many major oil companies are actively developing applications although very few appear in public domain (e.g., Arehart, 1989; Baldwin et al., 1989; Veezhinathan and Wagner, 1990).

Section 4.0 Coherent/Incoherent Data Picking Using a NN

The problem of coherent/incoherent data picking is cast as a pattern recognition problem. The neural net classifies signals as coherent or incoherent. Each maximal peak (or trough) over a data window is characterized by a set of signal attributes computed from a window of five samples centered at the peak (Taner, 1988). To avoid processing all peaks in a trace in a window, we restricted our data search to a window centered on the first maximal arrival and classified only the peaks within this window. The attributes selected for characterizing a peak are

(1) Predicted next amplitude value after the maximum amplitude of the peak (or trough); Prediction accomplished by application of a linear prediction operator designed by R.E. Lane using a method of minimum square variation (Lane, R.E., 1959).

(2) Mean power level in the five sample window. If $f(t)$ denotes a signal and $g(t)$ its Hilbert transform, then we compute the envelope $E(t)$ as

$E(t) = [f(t) + g(t)]^{1/2}$ and from this envelope determine the mean power level (MPL) as

$$MPL(t) = (1/5) \sum_{t-2}^{t+2} E(t) \Delta t;$$

(3) Power ratio (PR) between a forward and reverse sliding of the five sample window is given by $PR(t) = MPL(t+2)/MPL(t-2)$; and

(4) Envelope slope (ES) at the peak value is given by $ES(t) = \Delta [E(t)] / \Delta t$.

(5) The sum (SUM) of all normalized power spectral density values over the entire window domain of $f(t)$. If $Y = FFT(f(t), 512)$, then the power spectral density, a measurement of the energy at various frequencies, is given by:

$$P_{yy} = Y * conj(Y) / 512.$$

Whereupon, $SUM = \sum_i (Pyy(i) / MAX(Pyy(i)), Pyy(i) > 0$.

Some of these attributes (2-4), referred to as Hilbert attributes, were among those originally proposed in Taner et al. (1979). Hilbert attributes have been used by Amizadeh and Chatterjee (1984) and others for clustering (classifying) signal data. Attributes 1-4 are considered to have good discriminatory power to differentiate first maximal arrival peaks of coherent data and first maximal arrival peaks of incoherent data (Taner, 1988). Attribute 1 proved to be an improvement over using just the maximum amplitude. Lanes' operators are applied by cross correlation methods and are designed for fast, real time applications. Attribute (5) has good discriminatory power to differentiate coherent data from incoherent over the entire signal domain.

In our scheme, we let the network process three adjacent peaks at a time in order to decide if the center peak is a first maximal arrival relative to coherent or incoherent data. Decision is based on the observation that peaks adjacent to a maximal peak on coherent data provide a spatial context (or correlation) to identify coherency. Incorporation of this limited spatial information proved to be extremely important for good performance on signal data. Without the adjacent peaks, the network will probably achieve poor results. Taner uses only one peak for classification and employs the nearest neighbor technique to pick first arrivals. We suggest a BPN that employs his data attributes as input data. Using similar attributes, Geerling and Berkhout (1989) track seismic horizons with a combination of linear discriminant analysis and heuristic search. After observing open heart surgery and the application of the electrical sock to the heart, it was noticed that incoherent signals occurred when the sock's electrodes were in contact over scar tissue or not in good contact with heart muscle. Poor contact lead too very low amplitudes while scar tissue or partial scar tissue lead to data that was front or back loaded with a particular frequency. This last observation lead to the inclusion of the power spectral density data as one of the input attributes. Our BPN consists of 15 input units (5 attributes per peak times 3 peaks) for the 15 data values representing the peaks, 5 hidden units, and a single output unit as shown in Figure 1. If input traces were input in their original form from the cardiac system as the decimal integers, the NN trained very slowly (e.g., one hour for 30 sets of input numbers). Later all input trace amplitudes were scaled to the range -1 to 1. This gave the NN the ability to train fast. However, the interpretation of the output was difficult since the output numbers had to be investigated down to four or five digits. This was probably due to the fact that the original input data amplitudes from the cardiac mapping system differed by as much as 1000 units from trace to trace and patient to patient.

Since our proposed output scheme involves a "yes" or "no" decision, and since our output unit computes values between 0 and 1, we classify data as a coherent or incoherent with a Fuzzy Arithmetic approach. When the output exceeds 0.5 we classify data as coherent and as incoherent when the output falls between 0 and

0.5. We chose five hidden units after a certain level of experimentation. A BPN with too many hidden units "memorizes" the training data rather than generalizing from it and thereby does not perform well on test data. A BPN with too few hidden units lacks the degrees of freedom required to learn from the training patterns. Given this empirical result, experimentation with different numbers of hidden units is required to arrive at the minimum number of hidden units that produce an acceptable performance on the test data. In our case, for discussion purposes, say five hidden units enabled us to achieve good test-case performance.

It was found that converting the data in a special binary format where -0.5 corresponds to binary 0 and +0.5 corresponds to binary 1 achieved a fast training and very usable network. From a numerical analysis approach, the same topology as in Figure 1 was kept, except that each input unit in Figure 1 was represented by 10 (binary) units. Thus, each decimal (integer) was converted to binary form with the special format. For example the number eight would become: -.5, -.5, -.5, -.5, -.5, +.5, -.5, -.5.

The actual topology submitted to the NN system is given in Figure 2. There were 15 x 10 input units, 10 hidden units and 6 output units. The output was limited to a six digit binary representation.

The decimal representation given in Figure 1 either would not train at all, or in one experiment took over one hour to train. Normalizing the decimal input integers between -1 and +1 achieved a fast training network, but since the original data differed from signal to signal by over 1000 units, interpretation of the output was too tedious or cumbersome. The special binary representation-tation given in Figure 2 appears to be a major find in the design of the topology for NNs that simulate the requirements for relatively small NNs as given in Fig. 1.

Section 5.0 Methodology

This section outlines the processing steps used in the feasibility study.

Step 1) The C Language program EXTRACT was executed to de-multiplex data on files from the cardiac mapping system to a form suitable for data processing and obtaining the Hilbert and Fourier attributes. All 128 channels were sampled and held 1000 times per second.

Step 2) The C-like MATLAB program HF_ATTRIBUTES using the output from EXTRACT as input was executed to generate decimal integer data in a format suitable for input into the NN system.

Step 3) The C Language program CONVERT using the output from HF_ATTRIBUTES was executed to generate data in the special binary format for input into the NN system as Figure 2.

Step 4) The output of CONVERT was input into the NN system. An example of the output system is given in Figure 3.

It was the final output of Step 4) that determined if the NN was successful in detecting coherent/incoherent signals.

Section 6.0 Results

The NN has been used to pick coherent/incoherent signals on a number of different patients. In one survey, the network was trained using 20 traces from a single patient profile and achieved 100% accuracy. The network achieved 95% accuracy on a second patient profile before any more training with the second patient's data and 99% accuracy after further training with the new data. The network was then tested with 200 traces from five different patient profiles. The network achieved 97% accuracy before post additional training and 99% accuracy after additional training with only 20 additional signals used in the additional training. The NN also gave a proportional measurement of the degree to which data was coherent or incoherent. If the final output answer special form binary digit from the NN was (-.498, . . . -.498,+.478) in one case and (-.498, . . . -.498,+.278) in another case, it was observed that the first signal was of better quality than the second. This gives the cardiologist an opportunity to set thresholds during surgery for different patients, circumstances, and requirements. The turn-around time for obtaining an isochronal map during surgery using an NN approach should be decreased by an order of magnitude, if NN coherent/incoherent data picking was implemented in the cardiac mapping system.

In processing a second set of data, a modified network was used. The first attribute of maximum amplitude was replaced by a sum of all unit normalized amplitudes (between 0 and 1) over the signals domain. This gave the network a greater ability to classify weak-amplitude data as incoherent. However, some weak amplitude data upon signal gain and amplification are seen to be a strong indicator of VT activity. Thus, using this data gives a more complete picture of VT arrival areas on the heart surface. The weak amplitudes in these cases may be caused by granulations of scar tissue in the epicardial. Using the original NN, gave an output reading of -.498, . . . -.498, +.0178. Using the modified NN, gave an output reading of -.498, . . . -.498, +.3178. Usually weak amplitude data is incoherent and can be detected by other initial signal processing techniques. In the case cited above with the +.0178 output, thresholds could be used to save such data.

It was also observed that coherent data whose last NN out put node value of x , with $-.1 < x < +.35$, was not only weakly coherent but such data was a good predictor of upcoming VT. This is probably due to the heart data becoming more noisy as VT approaches. This suggests a non-invasive method of predicting VT with standard ECG data.

Figure 4 shows a collection of plots of VT signals indicating data picked for NN processing to determine if the signals are coherent or incoherent. Figure 5 shows the corresponding NN processing indicating how the NN picked the corresponding signals (as labeled) as coherent or incoherent, and predictive of upcoming VT. The accuracy shown here is 100 percent. The training

was accomplished using different patient data profiles. Not all of the data is shown.

All indications are that NN embodiments would achieve greater than 97% accuracy on patient profile surveys and save valuable, critical time in the operating room. Each patient profile survey apparently does not require a separate training. However, as fast as the binary form of the NN trains, all indications are that separate training could be done in the operating room. We consider the fast training technique of using a binary form of the NN represents a significant find in NN technology for our applications.

Section 7.0 Summary

NNs lend themselves to solving pattern recognition tasks. We have shown how to construct a BPN to pick coherent/incoherent data for an intraoperative cardiac mapping system. Our decimal input form network consists of four signal attributes computed from three potential first maximal arrival peaks as input units and one Fourier spectral density input. These input units connect to five hidden units which, in turn connect to a single output unit. Our special binary form of the NN consisted of $10 \times 5 \times 3 = 150$ input units, 10 hidden units, and 6 output units. Using a BP learning algorithm, we trained our NN to identify coherent/incoherent data based on a user-selected training set of known data.

An important point of our proposed methodology is that of our seemingly simplistic approach to incorporating space-time information for coherent/incoherent selection, as well as a method for predicting VT in patients. Our proposal also indicates that training a network once from one patient profile is enough to pick coherent/incoherent data for another patient survey. Also, since our NN is small and in the special binary form, training and processing time should not be an issue even if the training needs to be done during surgery.

Coherent/incoherent data picking for cardiac systems has conventionally been a difficult task to automate. Traditional approaches have achieved a limited success. We have indicated here that the NN technique has the capacity to learn how to make the picks on different kinds of data sets for applications in an intraoperative cardiac mapping system. This approach also shows promise for other data picking.

Much work has been done in the application of expert systems to signal processing problems as evidenced by the articles in Simaan and Aminzadeh (1989). While expert systems are good for emulating high-level cognitive tasks, they are unsuitable for automating low-level perceptual tasks such as pattern perception, either auditory or visual. Neural networks, on the other hand, are good at learning perceptual tasks. Therefore, a combination of these two technologies leading to hybrid systems will be immensely beneficial to tackling complex signal processing problems. We predict that more of these hybrid systems will be seen in the future not only in the medical field but in other areas of signal processing including radar and the oil industry.

References

- Aminzadeh, F., Ed., 1987, Pattern recognition and image processing: Handbook of geophysical exploration, Vol. 20, Geophysical Press.
- Aminzadeh, F., and Chatterjee, S. L. 1984, Application of clustering in exploration seismology, *Geoexploration*, 23, 147-159.
- Anderson, J. A., and Rosenfeld, E., Eds. 1988., *Neurocomputing: Foundations of research*, MIT Press/Bradford Books.
- Arehart, R. A., 1989, Drill bit diagnosis using neural networks, *Proc. 1989 Conf. on AI in Petr. Expl. and Prod.*, 133-144.
- Baldwin, J. L., Otte, D. N., and M. Bateman. R., 1989, Computer emulation of human mental processes: Application of neural network simulators to problems in well log pattern recognition, *Proc. 1989 Conf. on AI in Petr. Expl. and Prod.*, 145-176.
- Barhen, J., Gulad, S., and Zark, M., 1989, Neural learning of constrained nonlinear transfer Vol. 22. No. 6, 67-76.
- Benediktsson, J. A., Swain, P. H., and Ersoy, O. K. 1990, Neural network approaches versus statistical methods in classification of multisource remote sensing data, *Inst. Elect. Electron. Eng. Trans. Geoscience and remote sensing*, 28, 540-552.
- Carpenter, G. A., and Grossberg, S., 1988, The ART of adaptive pattern recognition by a self-organizing neural network, *Computer*, 21, No. 3, 77-88. 1987(a), ART 2: Self-organization of stable category recognition codes for analog input patterns: *Applied Optics*, 26, No. 3, 4919-4946. 1987(b), A massively parallel architecture for a self-organizing neural pattern recognition machine, *Computer Vision Graphics, and Image Processing*, 37, 54-115.
- Cottrell, G. W., Munro, P., and Zipser, D., 1987, Image compression by back propagation: An example of extensional programming, ICS Report 8702, Institute for Cognitive Science, Univ. of California, San Diego.
- Daugman, J. G., 1988, Complete discrete 2-D gabor transforms by neural networks for image analysis and compression, *Trans. Inst. Elect. Electron. Eng., Acoustics, Speech and Signal Processing*, 36, No. 7, 1169-1179.
- Desai and Jansen, *Wavedorm Classification Using Artificial Neural Networks*, 1992, NASA Third International Workshop On Neural Networks and Fuzzy Logic '92.
- Doyen, P. M., Guidish, T. M., and de Buyl, M. H., 1989, Seismic discrimination of lithology and sand/shale reservoirs: A Bayesian approach: 59th Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 719-722.
- Fukushima, K., 1988, Neocognitron: A hierarchical neural network capable of visual pattern recognition in *Neural network*, Vol. 1, No. 2, 119-130.
- Fuller, B. N., and Smithson, S. B., 1989, Multivariate pattern recognition of seismic data for exploitation: 59th Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 723-725.
- Geerlings, A. C., and Berkhout, A. J., 1989, Heuristic event tracking linked to linear discriminant analysis, in Simaan, M., and Aminzadeh, F., Eds., *Advances in geophysical processing: JAT Press Inc.* Vol. 3. 201-233.
- Gorman, R.P., and Sejnowski, T. J., 1988, Learned classification of sonar targets using a massively parallel network, *Inst. Elec. Electron. Eng. Trans. Acoustics, speech and signal processing*, Vol. 36, No. 7, 1135-1140.
- Hecht-Nielsen, R., 1987, Kolmogorov's mapping neural network existence theorem, *Proc. Inst. Elec. Electron. Eng. First Int. Conf. on neural networks: Vol. m*, 11-13.
- Hopfield, J. J., and Tank, D. W., 1985, Neural computation of decisions in optimization problems: *Biological cybernetics*, Vol. 52, 141-152.
- Hopfield, J. J., 1982, Neural networks and physical systems with emergent collective computational abilities, *Proc. Nad. Acad. Sci., USA*, Vol. 79, 255-2558.
- Huang, K., Liu, W. H., and Chang, I. C., 1989, Hopfield model of neural networks for detection of bright spots, 59th Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 444-446.
- Huang, W. Y., and Lippmann, R., 1987, Comparisons between neural net and conventional classifiers, *Proc. 1st Inter. Conf. on Neural Networks*, *Inst. Elec. Electron. Eng.*, IV, 485-493.
- Kohonen, T., 1987, *Self-organization and associative memory*, 2nd Edition: Springer Verlag New York, Inc.
- Kosko, B., 1987, Adaptive bidirectional associative memories, *Applied Optics*, Vol. 26, No. 23, 4947-4960.
- Lane, Ralph E., *Mathematical processes, linear operators using few data points*, Military Physics Research Laboratory, The University of Texas, Austin, Texas, 1959.
- Lang, K. J., 1989, A time-delay neural network architecture for speech recognition: Dissertation, CMU-CS-89-185, Carnegie Mellon University.
- Lapedez, A., and Farber, R., 1987, Nonlinear signal processing using Neural networks: Prediction and system modeling, *Tech. Rep. LAUR-87-2662*, Los Alamos National Lab.

Lee, Y., and Lippmann, R. P., 1989, Practical characteristics of neural network and conventional pattern classifiers on artificial and speech problems, in Touretzky, D. S., Ed., *Advances in neural information processing systems 2*: Morgan Kaufmann Pub.

Lippmann, R. P., 1987, An introduction to computing with neural nets: *Inst. Elec. Electr. Eng., ASSP Magazine*, April, ~n.

Liu, X., Xue, P., and Li, Y., 1989, Neural network method for tracing seismic events, 59th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 716-718.

McCormack, M. D., 1990, Seismic trace editing and first picking using neural networks, 60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 321-324.

Minsky, S. A., and Papert, S. A., 1969, *Perceptrons*, MIT Press.

Palmer, D., 1987, Removing random noise from EKG signals using a backpropagation network, informal paper from Hecht-Nielsen Neurocomputer, Inc.

Pieper, Carl F., Parsons Dan, Lawrie, Gerald M., Lacy, Jeffery, Roberts, Robert, and Pacifico, Antonio, Design and implementation of a new computerized system for intraoperative cardiac mapping, *American Physiological Society*, 1991, 1529-1539.

Poulton, M. M., Glas, C. E., and Sternberg, B. K., 1989, Recognizing EM ellipticity patterns with neural networks, 59th Ann. Internat. Mtg., Soc. Expl. Geophys., Extended Abstracts, 208-212.

Psaltis, D., Sideris, A., and Yamamura, A., 1987, Neural Controllers, *Proc. Inst. Elec. Electron. Eng., 1st Internat. Conf. on Neural Networks*, Vol. IV, 551-558.

Ramanujam, J., and Sadayappan, P., 1988, Optimization by neural networks, *Proc. Inst. Elec. Electron. Eng. 2nd Internat. Conf. on Neural Networks*, Vol. II, 325-332.

Roeth, G., and Tarantola, A., 1989, Retrieving background velocity model using neural networks, *Groupe de Tomographie Geophysique, Periodical Report No. 16*, 73-98.

Rosenblatt, F., 1962, *Principles of neurodynamics*, Spartan Books.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J., 1986, Learning internal representations by error propagation, in *Parallel distributed processing*, MIT Press, Chapter 8.

Sejnowski, T. J., and Rosenberg, C. R., 1987, Parallel networks that learn to pronounce English text, *Complex Systems*, Vol. 1, 145-168.

Shea, P. M., and Lin, V., 1989, Detection of explosives in checked airline baggage using artificial neural system, *Proc. of Inter. Joint Conf. on Neural Networks*, Vol. II, 31-34.

Simaan, M., and Aminzadeh, F., Eds., 1989, *Advances in geophysical data processing*, Vol. 3, Artificial intelligence and expert systems in petroleum exploration, JAI Press.

Taner, M. T., Koehler, F., and Sheriff, R. E., 1979, Complex seismic trace analysis, *Geophysics*, 44, 119-1212.

Taner, M. T., 1988, The use of supervised learning in first break picking in Bielanski, E., Ed., *Proc. Symposium Geophysical Society of Tulsa*.

Tou, J. T., and Gonzales, R. C., 1974, *Pattern recognition principles*, Addison-Wesley.

Urnes, Hoy, and Ladage, A Neural Based Intelligent Flight Control System For The NASA F - 15 Flight Research Aircraft June, 1992, NASA- Third International Workshop On Neural Networks and Fuzzy Logic '92.

Veezhinathan, J., and Wagner, D., 1990, A Neural network approach to first arrival picking, *Proc. Int. Joint Conference on Neural Networks*, Vol. I, 235-240.

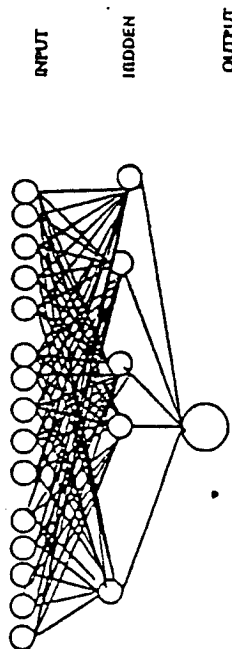
Waibel, A., Hanazawa, T., Hinton, G., Shikano, K., and Lang, K., 1989, Phoneme recognition using time-delay neural networks, *Inst. Elec. Electron. Eng. Acoustics, Speech, and Signal Processing*, Vol. 27, No. 3, 328-339.

Werbos, P., 1974, Beyond regression: New tools for prediction and analysis in the behavioral sciences, dissertation, Applied Mathematics, Harvard Univ.

Widrow, B., and Winter, R., 1988, Neural nets for adaptive filtering and adaptive pattern recognition, *Computer*, Vol. 21, No. 3, 25-39.

Widrow, B., and Stearns, S., 1985, *Adaptive signal processing*, Prentice Hall.

Zhou, Y. T., Chellappa, R., Vaid, A., and Jenkins, B. K., 1988, Image restoration using a neural network, *Inst. Elec. Electron. Eng., Trans. Acoustics, Speech, and Signal Processing*, Vol. 36, No. 7, 1141-1151.

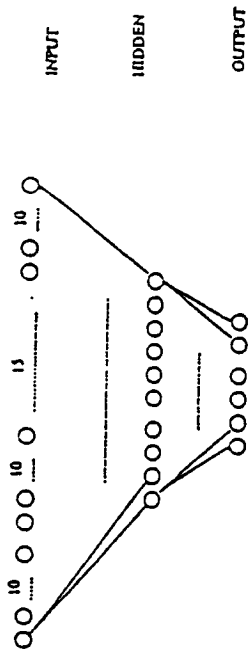


LEFT PEAK CENTER PEAK RIGHT PEAK

ATTRIBUTES

- PREDICTED VALUE AFTER
- MAXIMUM AMPLITUDE
- MEAN POWER
- POWER RATIO
- ENEVELOPE SLOPE
- POWER DENSITY SPECTRUM

Figure 1. Coherent/Incoherent data picking neural network



ATTRIBUTES

- PREDICTED VALUE AFTER
- MAXIMUM AMPLITUDE
- MEAN POWER
- POWER RATIO
- ENEVELOPE SLOPE
- POWER DENSITY SPECTRUM

Figure 2. Topology of the coherent/incoherent data picking neural network for the special binary format

Figure 3: Output From The NN System

RECALL 6, 1

/* Input vector 1 : */

-0.478658, -0.475836, -0.469284, -0.475090, -0.470809, -0.437245,

/* Input vector 2 : */

-0.478641, -0.473466, -0.470729, -0.476208, -0.470877, -0.390641,

/* Input vector 3 : */

-0.475842, -0.476121, -0.464263, -0.470580, -0.468157, -0.374276,

/* Input vector 4 : */

-0.478803, -0.471997, -0.469595, -0.475888, -0.470676, -0.390361,

/* Input vector 5 : */

-0.464500, -0.468994, -0.473949, -0.470307, -0.471562, 0.422108,

/* Input vector 6 : */

-0.463685, -0.474797, -0.473476, -0.468343, -0.473971, 0.371821,

/* Input vector 7 : */

-0.463561, -0.466970, -0.474205, -0.470065, -0.473598, 0.451783,

/* Input vector 8 : */

-0.466506, -0.468845, -0.466982, -0.465439, -0.472578, 0.430247,

/* Input vector 9 : */

-0.466802, -0.468839, -0.473975, -0.471927, -0.474043, 0.416889,

/* Input vector 10 : */

-0.469446, -0.470509, -0.474085, -0.473532, -0.474110, 0.418382,

/* Input vector 11 : */

-0.475303, -0.474494, -0.469675, -0.474660, -0.467105, -0.206562,

/* Input vector 12 : */

-0.460718, -0.463495, -0.475494, -0.467562, -0.465668, 0.429976,

/* Input vector 13 : */

-0.474842, -0.470675, -0.467179, -0.474065, -0.462884, -0.152269,

/* Input vector 14 : */

-0.453162, -0.455138, -0.466574, -0.458578, -0.464784, 0.455315,

/* Input vector 15 : */

-0.462081, -0.471533, -0.472086, -0.467531, -0.471830, 0.416687,

/* Input vector 16 : */

-0.448438, -0.459818, -0.458182, -0.442758, -0.459059, 0.390698,

/* Input vector 17 : */

-0.462817, -0.463901, -0.459881, -0.462252, -0.460812, 0.397687,

Figure 4. VT Cardiac Mapping System Plots from a Patient Profile

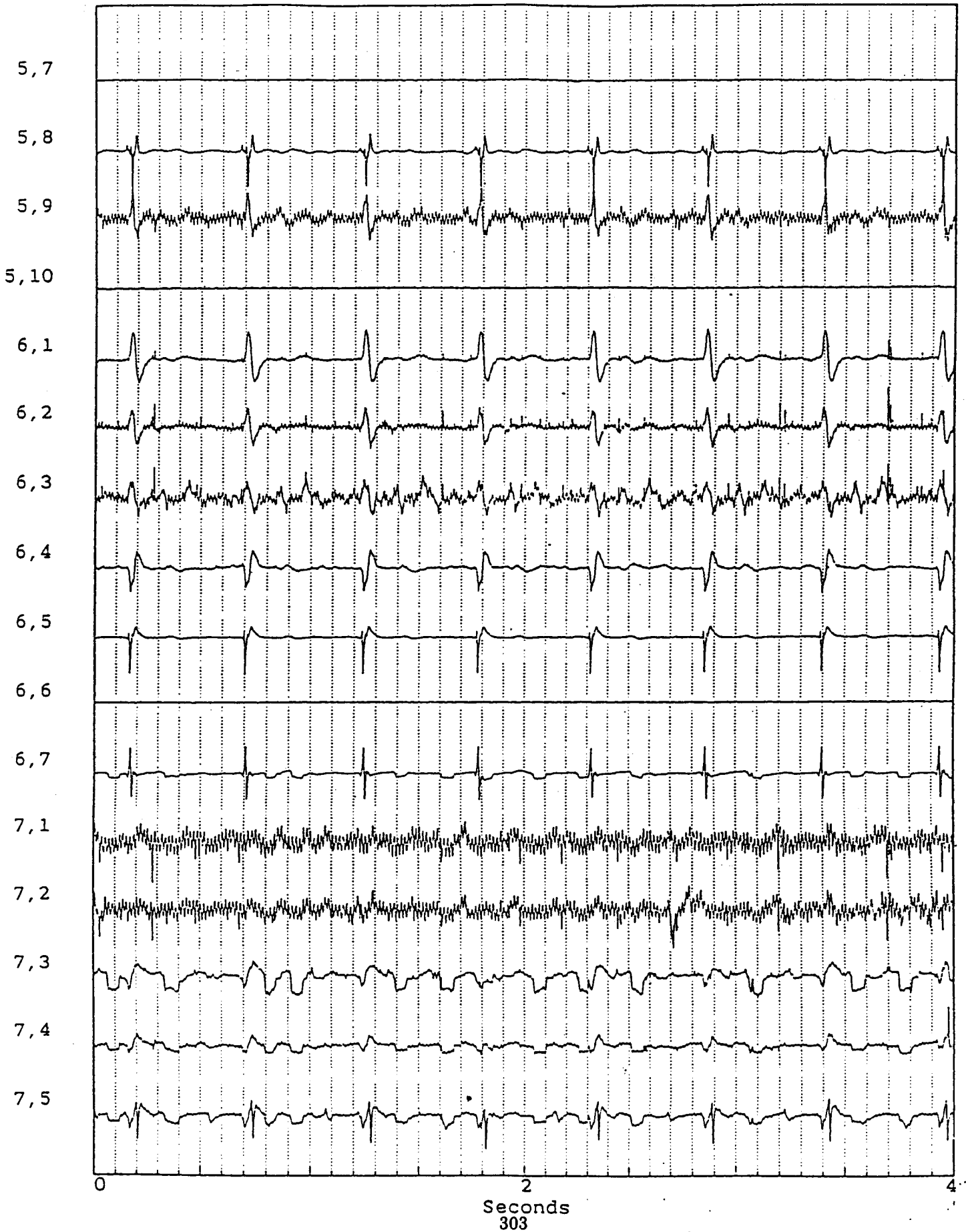


Figure 5. Output From The NN System Corresponding To The Plots In Figure 1.

The Signal (m,n) labels correspond to the (m,n) labels on the plots of the signals in Figure 1.

RECALL 6, 1	Label	Classification
/* Input vector 1 : */	Signal (7,1)	
-0.478658, -0.475836, -0.469284, -0.475090, -0.470809, -0.437245,		Incoherent
/* Input vector 2 : */	Signal (7,2)	
-0.478641, -0.473466, -0.470729, -0.476208, -0.470877, -0.390641,		Incoherent
/* Input vector 3 : */	Signal (9,1)	
-0.475842, -0.476121, -0.464263, -0.470580, -0.468157, -0.374276,		Incoherent
/* Input vector 4 : */	Signal (11,1)	
-0.478803, -0.471997, -0.469595, -0.475888, -0.470676, -0.390361,		Incoherent
/* Input vector 5 : */	Signal (6,1)	
-0.464500, -0.468994, -0.473949, -0.470307, -0.471562, 0.422108,		Coherent
/* Input vector 6 : */	Signal (6,2)	
-0.463685, -0.474797, -0.473476, -0.468343, -0.473971, 0.371821,		Coherent
/* Input vector 7 : */	Signal (6,3)	
-0.463561, -0.466970, -0.474205, -0.470065, -0.473598, 0.451783,		Coherent
/* Input vector 8 : */	Signal (6,4)	
-0.466506, -0.468845, -0.466982, -0.465439, -0.472578, 0.430247,		Coherent
/* Input vector 9 : */	Signal (6,5)	
-0.466802, -0.468839, -0.473975, -0.471927, -0.474043, 0.416889,		Coherent
/* Input vector 10 : */	Signal (7,5)	
-0.469446, -0.470509, -0.474085, -0.473532, -0.474110, 0.418382,		Coherent
/* Input vector 11 : */	Signal (7,3)	
-0.475303, -0.474494, -0.469675, -0.474660, -0.467105, -0.206562,		Coherent
/* Input vector 12 : */	Signal (7,7)	
-0.460718, -0.463495, -0.475494, -0.467562, -0.465668, 0.429976,		Coherent
/* Input vector 13 : */	Signal (10,2)	
-0.474842, -0.470675, -0.467179, -0.474065, -0.462884, -0.152269,		incoherent(or weak coherent)
/* Input vector 14 : */	Signal (11,2)	
-0.453162, -0.455138, -0.466574, -0.458578, -0.464784, 0.455315,		VT occurred after this data Coherent
/* Input vector 15 : */	Signal (11,3)	
-0.462081, -0.471533, -0.472086, -0.467531, -0.471830, 0.416687,		Coherent
/* Input vector 16 : */	Signal (7,4)	
-0.448438, -0.459818, -0.458182, -0.442758, -0.459059, 0.390698,		Coherent
/* Input vector 17 : */	Signal (7,5)	
-0.462817, -0.463901, -0.459881, -0.462252, -0.460812, 0.397687,		Coherent

5-9-39
44438

**VIBRATION CONTROL OF DEPLOYABLE ASTROMAST BOOM -
PRELIMINARY EXPERIMENTS**

M. Swaminadham
Assoc. Prof., Dept. of Engrg. Techn.
Texas A&M University
College Station, Texas 77843-3367

David A. Hamilton
Loads and Structural Dynamics Branch
NASA - L.B. Johnson Space Center
Houston, Texas 77058

ABSTRACT

This paper deals with the dynamic characterization of a flexible aerospace solar boom. The modelling issues and sine dwell vibration testing to determine natural frequencies and mode shapes of a continuous-longeron deployable ASTROMAST lattice boom are discussed. The details of the proof-of-concept piezoelectric active vibration experiments on a simple cantilever beam to control its vibrations are presented. The control parameters like voltage to the controller crystal and its location are investigated, to determine the effectiveness of control element to suppress selected resonant vibrations of the test specimen. Details of this experiment and plans for its future adaptation to the prototype structure are also discussed.

INTRODUCTION

Lightweight lattice structures with automatic deployment and retraction capabilities offer potential benefits for space applications. Typical applications of such structures include instrument booms, gravity-gradient booms, antennas, sensor masts, mobile transporters, and solar sail masts. An added feature is that these structures can be quickly designed and fabricated to meet the requirements of a variety of space applications. The present technology allows the design of these space masts in two configurations -- one with continuous longeron members which are elastically coiled to their stowed position, and the other with articulated, segmented longerons which are folded in to stowed configuration. The ratio of its retracted (stowed) -to-deployed lengths is typically 1/50. The size and shape to which these masts can be fabricated is virtually unlimited; however, continuous masts are limited to small radius due to the longeron overall elastic bending strength, whereas the articulated masts with segmented members do not have size and length limitations.

All space masts are susceptible to vibrations while in flight. On-orbit manipulations induce severe oscillations and the dynamic characterization of such structures is an integral part of its design for mission. The natural frequencies, operational deflection shapes and modal damping are the key parameters that determine the suitability of the structure for a specific mission.

In the present study, a $\frac{1}{4}$ segment of a 100-ft long, 29.5-inch diameter continuous longeron solar boom manufactured by ASTROMAST Corporation is investigated for its vibration characteristics. The size and geometric configuration of this lattice structure with rotational joints, double-braced battens, and diagonal stiffeners pose additional burden on the analysis and testing tasks. In addition, the future goal of suppressing vibrations of booms and adaptation of piezoelectric devices to accomplish this objective presents challenging opportunities for research.

This research has two objectives: First, model the real-world ASTROMAST boom by a basic finite-element method and test the boom by sine dwell test for natural frequencies and mode shapes; second, demonstrate the proof-of-concept of piezoelectric crystals as vibration actuating devices on a simple test specimen.

ASTROMAST VIBRATIONS

The finite-element model of the 15 bay, 21-ft long, 29.5-inch diameter boom consisted of 81 nodes and 378 rod elements. The basic model treated all battens as straight circular rods. Friction at the joints was neglected, the aluminum rotating joints were considered as concentrated masses. The material properties of S-glass supplied by ASTROMAST Company were used [1]. The analytical model applied partial rotational constraints for the joints where the three longerons meet with the aluminum end plates. ABAQUS finite-element software was used to determine the first 15 frequencies and modes. At first glance, it was difficult to define the mode shapes. The mode participation factors which define the dominant component of motion helped identify the mode shapes of this boom. Three distinct motions seemed possible -- torsion about the length (x) axis, and bending motions about two perpendicular (x-y) axes. Insignificant model participation factors associated with some frequencies indicated coupled motions between x and z axis and also x and y axis.

The mast has been deployed and supported at both ends. A combined electrodynamic-piezoelectric shaker with matching network provided a sine forcing function at the mid-length of the boom. A rowing accelerometer was used and the responses were analyzed in time and frequency domains to identify resonances. Lissajous figures were also generated to confirm the resonances. Figure 1 shows the set up for vibration test of the ASTROMAST. The mast was excited by a harmonic forcing function at one of the resonance frequencies and tuned to obtain maximum response. A comparison of the measured and predicted frequencies of the mast is provided in Table I.

A reasonable agreement is evident with the predicted and measured frequencies. It is difficult, however, to match the predicted mode shapes with the measured, because of limited instrumentation. The test indicated distinct resonances clustered around 62 and 86 Hz. The analysis did not delineate these modes. The analysis and tests identified the frequency regimes of interest and gave confidence to adapt crystal controllers on the test boom to control resonant vibrations.

PIEZOELECTRIC ACTUATOR EXPERIMENTS

The piezoelectric phenomenon - an effect that produces mechanical motion of piezoelectric materials such as quartz when subjected to electrical potential, and its converse effect are well known. However, the applications of synthetic piezoelectric materials such as barium titanate, lead zirconate titanate for actuation and control of motion of mechanical and structural components are little known or understood, and the present research explores the effectiveness of lead zirconate titanate crystals [2] for controlling resonant vibrations of a test cantilever beam.

The experiment consisted of exciting the cantilever beam in its flexural mode by an exciter crystal and inducing a negative effect on the same beam by a controller crystal. Figure 2 shows the exciter crystals near the fixed end of two specimens. Control crystal is cemented to the bottom surface of the second beam. A small accelerometer at the tip measured the resonant response of the beam. Exciter and control crystals were bonded directly to the surface of the test specimen by a 50:50 epoxy. An inverter was placed ahead of the controller crystal power amplifier circuit to affect 180° phase shift in the control voltage. Figure 3 shows the schematic of the instrumentation used to excite and control vibrations of a cantilever beam.

Two experiments were conducted -- the first with the controller crystal at two-thirds length and the second with the controller crystal near the root of the beam. Table II summarizes the results. Based on the observations, the following general comments are made: The crystals were effective in controlling the vibrations; the crystals worked well at higher frequencies compared to lower frequencies; appreciable phase between the excitation and response was evident; the beam response became unstable at higher control voltages; larger control voltages and custom design crystal were perhaps, required for real-world structures.

The fundamental mode of the beam at 29 Hz was difficult to excite because of two reasons -- the motion at low frequencies was transferred to the base, and the crystal is not effective at such low frequencies. The vibration reduction at the second, third and fourth flexural modes was substantial. A typical oscillograph signals at 184 Hz in Figures 4 (a) and (b) with and without control crystal indicate the effectiveness of controller in reducing the vibration. Similar signatures were recorded for other resonances at 29, 522, 1021, and 1680 Hz. With the change in exciter location, especially the crystal at two-thirds length altered the natural frequency to a small degree. More beneficial effects were observed when the exciter crystal was located near the fixed end of the beam. With the inverter in place, it was possible to introduce 180° phase shift and achieve vibration reduction, however, there is a reason to believe that a precise matching of phase between the excitation and response will completely reduce the resonant vibrations of the test beam.

CONCLUSIONS

The deployable ASTROMAST solar boom was finite-element modelled and sine dwell tested. A reasonable agreement between the predicted and measured frequencies

was observed. In addition to research, this amazing structure roused the interest of students in space structures, structural dynamics and vibration control. Piezoelectric crystals showed distinct promise for their application for the vibration control of structural components. Higher voltages and custom designs will perhaps, function well to control the ASTROMAST type flexible space structures. The present research will attempt to adapt large size crystals to achieve vibration reduction in the test ASTROMAST.

ACKNOWLEDGEMENTS

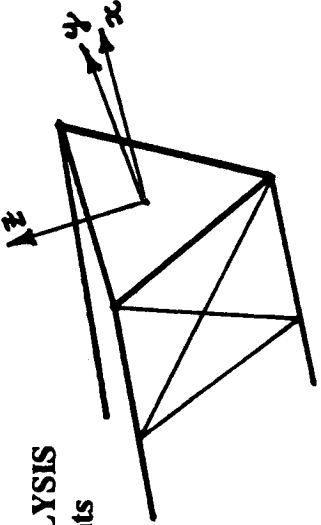
The first author acknowledges with gratitude the financial assistance received from John Space Center (JSC), Houston, under the JSC - Regional University Research Program for 1993-94 to do this research. He also thanks Dr. John A. Weese, Head of the Department of Engineering for his constant support and encouragement.

BIBLIOGRAPHY

1. ASTROMAST for Space Applications, Astro Aerospace Corporation Report, AAC-B-004, July 1985.
2. Some Design Considerations in the Use of Bimorphs & Motor Transducers, Morgan Matroc, Inc. Report, TP-237

TABLE I SOLAR BOOM VIBRATION ANALYSIS
FEM MODEL: 81 Nodes, 378 Rod Elements

MATERIAL PROPERTIES: $E = 8.84E + 06$ psi, Poisson's Ratio = 0.25,
 Mass Density = $2.6E - 04$ lbm/in³;
 Radius of longeron = $\frac{5}{8}$ inch; Batten = 0.314 in.
 Diameter, $D = 29.5$ in



Mode Participation Factors									
MODE	Frequency (ABAQUS)	Frequency Experimental (Hz)	X	Y	Z	X _R	Y _R	Z _R	COMMENTS
1	23.9	20.7				18.04			Torsion about x
2	29.2	27.4						154.4	Bending about z
3	32.6	30.6					-154.8		Bending about y
4	33.1	37.4						-2.5	Combined x and z
5	58.2	45.1							Insignificant Motion
6	61.65	62.4						-86.5	Bending about z
		64.5					84.9		Bending about y
		65.1							
7	86.6	82.5					5.5		Torsion about x
		86.6							
		89.3							
8	94.9	95						-68.1	Bending about y
9	97.1								

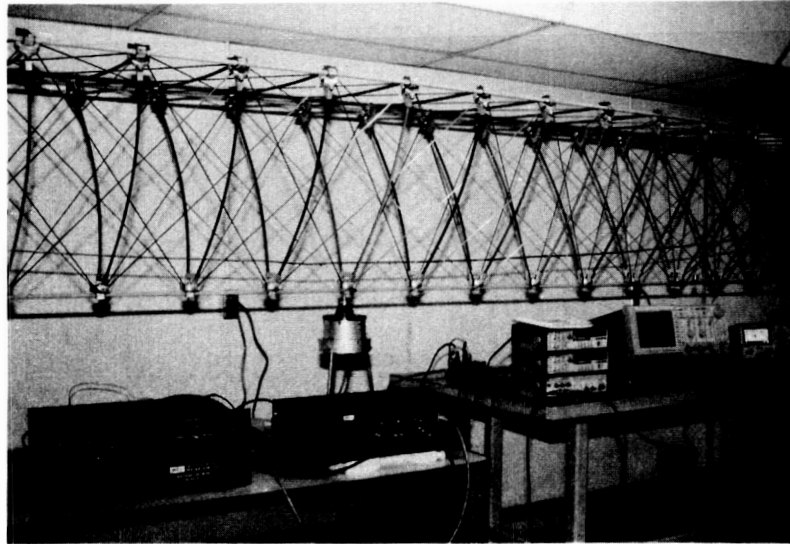


Figure 1. A Photograph of ASTROMAST Under Vibration Test.

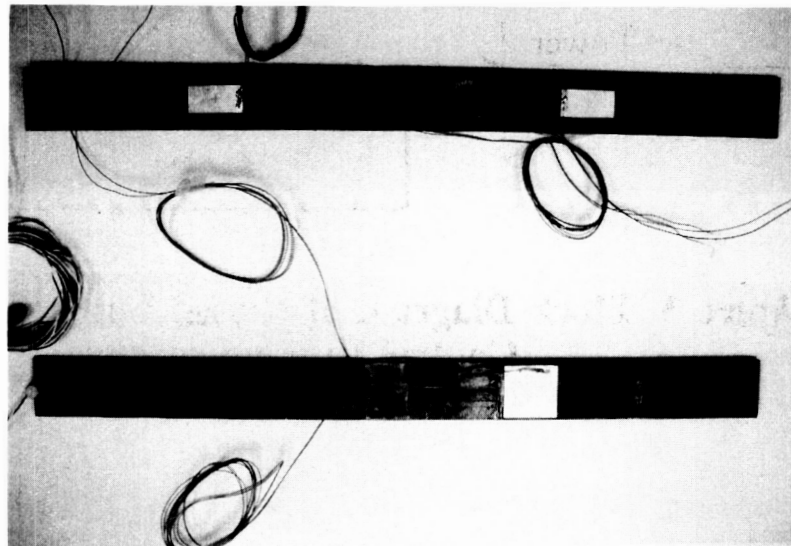


Figure 2. Piezoelectric Crystals on Two Test Specimens.

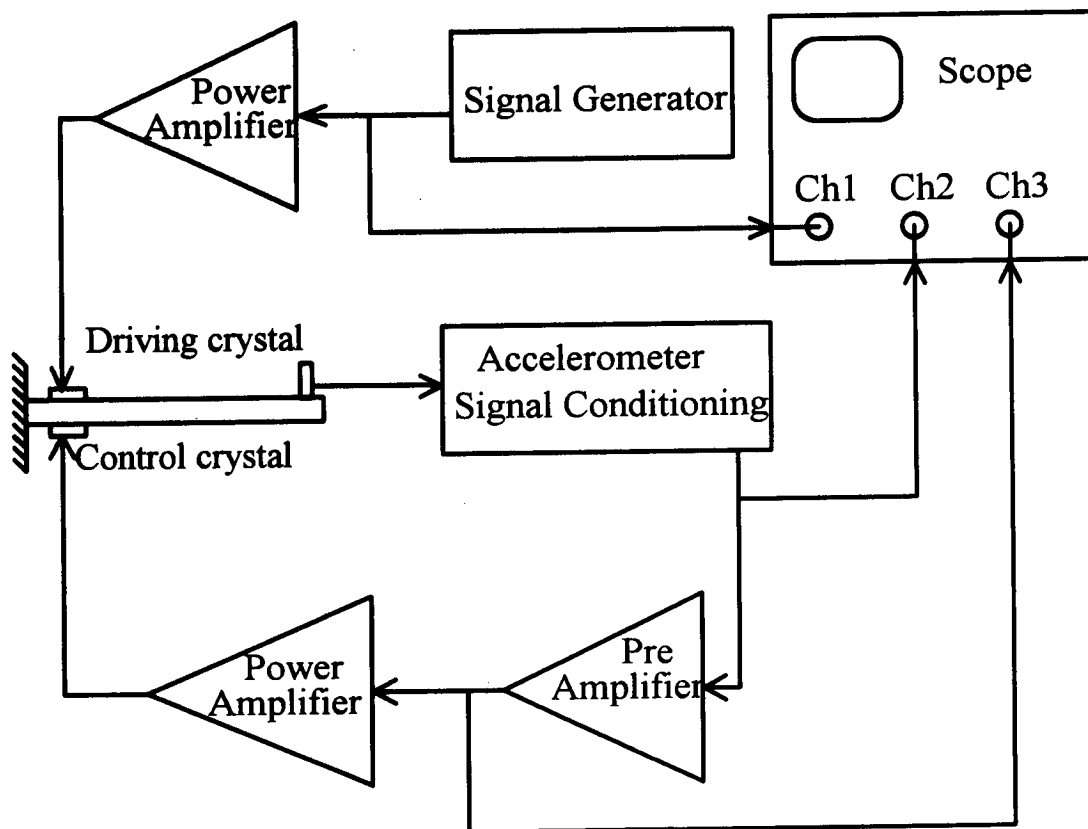
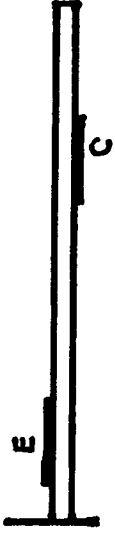
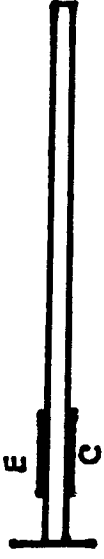
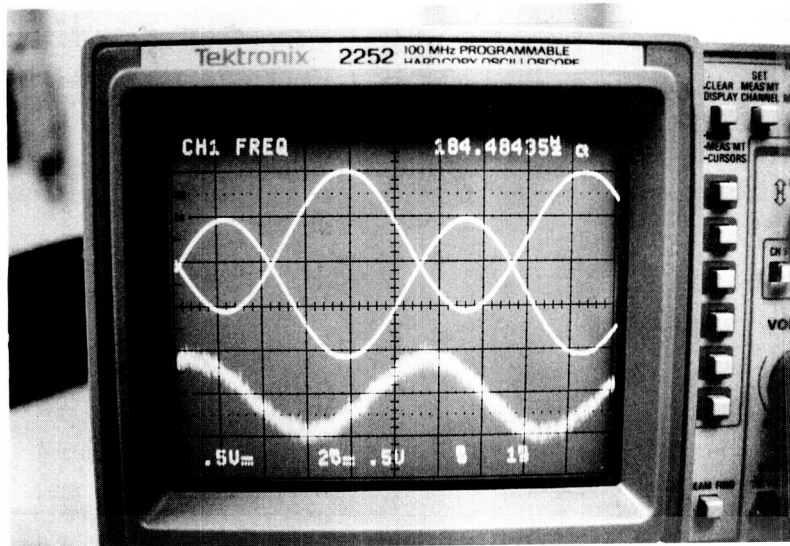


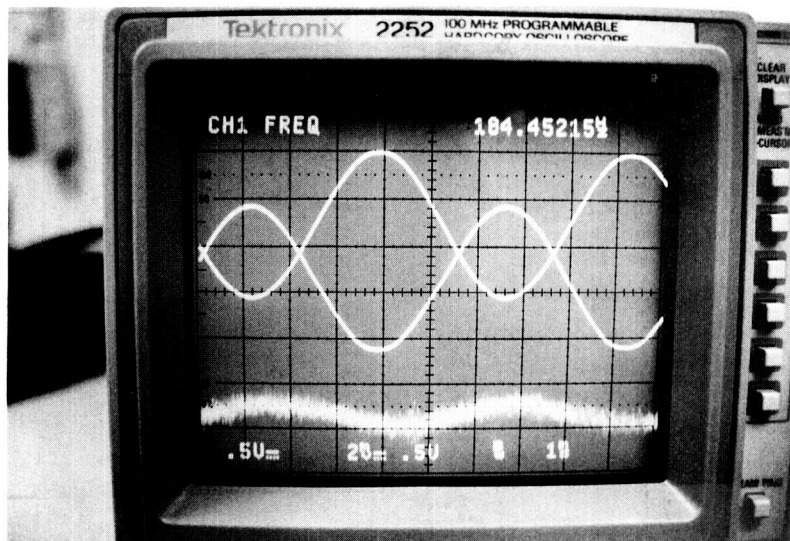
Figure 3. Block Diagram of Closed Loop Vibration Control Instrumentation.

TABLE II
PIEZO ELECTRIC CRYSTAL (PZT 5, 1 x 0.5 x 0.040 inch)
ACTUATOR EXPERIMENTS ON A STEEL CANTILEVER BEAM (12 x 1.14 x 0.139 inch)

		EXPERIMENT I				EXPERIMENT II				
MODE		Frequency (Hz)	Exciter (E)	Control (C)	BOTH (E+C)		Frequency (Hz)	Exciter (E)	Control (C)	BOTH (E+C)
1		29.7	7.8 mv	3 mv	--		29.2	7.02 mv	6.0 mv	2 mv
2		187.6	160 mv	160 mv	50		182.8	78 mv	78 mv	35 mv
3		522.2	300 mv	302 mv	71		506.5	150 mv	150.5 mv	29 mv
4		1021	508 mv	506 mv	385		984.3	820 mv	665 mv	205 mv
5		1680	254 mv	200 mv	123		1622	1290 mv	1300 mv	205 mv



**Figure 4a. Beam Response Signal (Bottom) at 184 Hz
With Control Crystal Off.**



**Figure 4b. Beam Response Signal (Bottom) at 184 Hz
With Control Crystal ON.**

530-29

STRUCTURAL HEALTH MONITORING OF LARGE STRUCTURES

Hyoungh M. Kim* and Theodore J. Bartkowicz†
McDonnell Douglas Aerospace - Space Systems, Houston, Texas

Suzanne Weaver Smith‡
University of Kentucky, Lexington, Kentucky

David C. Zimmerman§
University of Houston, Houston, Texas

ABSTRACT

This paper describes a damage detection and health monitoring method that was developed for large space structures using on-orbit modal identification. After evaluating several existing model refinement and model reduction/expansion techniques, a new approach was developed to identify the location and extent of structural damage with a limited number of measurements. A general area of structural damage is first identified and, subsequently, a specific damaged structural component is located. This approach takes advantage of two different model refinement methods (optimal-update and design sensitivity) and two different model size matching methods (model reduction and eigenvector expansion). Performance of the proposed damage detection approach was demonstrated with test data from two different laboratory truss structures. This space technology can also be applied to structural inspection of aircraft, offshore platforms, oil tankers, bridges, and buildings. In addition, its applications to model refinement will improve the design of structural systems such as automobiles and electronic packaging.

INTRODUCTION

Space structures intended for long-term operation such as the International Space Station (ISS) may be damaged by micro-meteoroid and orbital debris impacts, long exposure to the space environment, fatigue, or accidents with spacecraft docking and berthing operations. Comprehensive on-orbit inspection of the ISS structure using traditional non-destructive evaluation (NDE) methods requires extensive extravehicular activity (EVA) and does not always reveal structural anomalies. Robotic inspections are also possible, but

there are limitations on areas of the spacecraft which can be inspected visually by robots or EVA astronauts.

Offshore structures for oil and gas production also must survive for thirty years or more in remote and hostile environments. They are subject to damage from tropical storms, hurricanes, corrosion, fatigue, or docking accidents. Offshore structures are inspected below the waterline by divers who, much like EVA astronauts, have limitations on the time their life support systems can sustain them for structural inspection. Robotic or human visual inspections are hampered by murky waters, barnacles, and other subsea plant life attached to the structure.

Vibration measurements have been used with system identification algorithms to produce quantitative damage detection results [1]. A math model of the undamaged structure, usually correlated with test data of the undamaged structure, is used with vibration information measured from the damaged structure. This damage detection approach is equivalent to verification of structural properties in specific locations (called model refinement). This non-traditional use of vibration measurements shows promise for application to on-orbit, remote NDE for both manned and unmanned space structures.

The proposed on-orbit damage detection method monitors the location and extent of structural damage through periodic system identification during the service life of the spacecraft. Modal parameters such as natural frequencies and mode shapes from a damaged structure can be related to the system mass and stiffness matrices of the undamaged structure. The feasibility of identifying mode shape, frequency, and damping of large space structures through on-orbit modal tests has been addressed in previous studies [2]. In addition, the cross-correlation random data processing strategy was developed to identify modal parameters using ambient responses from spacecraft [3].

However, the outcome of the research will have much wider applications. Offshore structural inspections can benefit from the same technology which utilizes vibration responses from the structure to locate the specific area and extent of suspected damage. Oil tankers, bridges, and buildings can also utilize the same tech-

*Principal Engineer, Space Station Div., Member AIAA

†Engineer Specialist, Space Station Div., Member AIAA

‡Asst. Professor, Eng. Mechanics, Senior Member AIAA

§Assoc. Professor, Mechanical Eng., Member AIAA

nology for structural inspection. Serious accidents can be avoided through the use of a real-time, early warning system implemented by a small number of sensors on those existing structures. In addition, applications in model refinement will improve the design of structural systems such as automobiles and electronic packaging.

This paper describes a damage detection and health monitoring method that was developed for large and complex structures with a limited number of measurements. Theoretical background is provided for the optimal-update method, design sensitivity analysis technique, and a model reduction/expansion technique. In the new two-step damage detection approach, a general area of structural damage is first identified and, subsequently, a specific damaged structural component is located. Performance of the proposed damage detection approach was demonstrated with testing and analysis of the NASA-LaRC 8-bay and McDonnell Douglas 10-bay truss structures. Further detailed theoretical and technical information can be found in References [4, 5, 6].

INITIAL DAMAGE DETECTION

System identification techniques can be directly applicable to structural damage detection and categorized into the following four classifications: classical approach [7], design sensitivity approach [8], eigenstructure assignment approach [9], and optimal-update approach [10]. The optimal-update approach adjusts system mass and/or stiffness matrices to match the test frequencies and mode shapes by solving a constrained optimization problem. In addition to minimum changes of the initial system matrices, other physical properties can be imposed such as symmetry, positive definiteness, and structural connectivity. By comparing the adjusted model to the initial model, structural damage can be identified and located. This approach was selected for initial damage detection because it is applicable to large and complex structures.

Theoretical Background

Figure 1 shows an analysis flow diagram for the initial damage detection approach. An analytical (undamped, linear) model of the undamaged structure can be generated using the finite element method and written as

$$M_{au}^{n \times n} \ddot{x}_{au}^{n \times 1} + K_{au}^{n \times n} x_{au}^{n \times 1} = 0 \quad (1)$$

where subscripts a and u represent analytical and undamaged, respectively, and superscript n represents n-degree of freedom (dof). Frequencies and mode shapes of the damaged structure obtained from modal testing

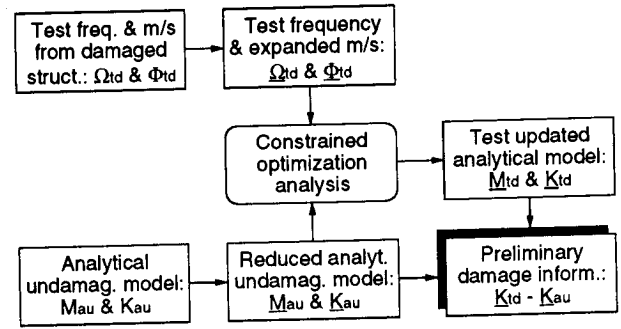


Figure 1: An Analysis Flow Diagram for the Initial Damage Detection Approach.

will satisfy the following modal equations

$$-M_{td}^{m \times m} \Phi_{td}^{m \times p} (\Omega_{td}^{p \times p})^2 + K_{td}^{m \times m} \Phi_{td}^{m \times p} = 0 \quad (2)$$

where subscripts t and d represent test and damaged, respectively; and superscripts m and p represent the number of measurements and the number of identified modes, respectively ($m \ll n$).

To compare the undamaged analytical model and damaged test model for damage detection, they must have the same number of dofs. Therefore, the order of the analytical model (system mass and stiffness matrices) must be reduced or the order of test results (mode shapes) must be expanded. A new reduction/expansion approach was developed by selecting an intermediate dof set, i , which is small enough for computational efficiency but large enough to describe the physical model in sufficient detail to locate the damage. The orders of equations 1 and 2 will become the same, allowing comparison to locate the structural damage:

$$\hat{M}_{au}^{i \times i} \hat{x}_{au}^{i \times 1} + \hat{K}_{au}^{i \times i} \hat{x}_{au}^{i \times 1} = 0 \quad (3)$$

$$-\hat{M}_{td}^{i \times i} \hat{\Phi}_{td}^{i \times p} (\Omega_{td}^{p \times p})^2 + \hat{K}_{td}^{i \times i} \hat{\Phi}_{td}^{i \times p} = 0 \quad (4)$$

where $(\hat{\cdot})$ indicates new variables through the coordinate transformations [4].

The basic optimal-update approach, developed by Baruch and Bar Itzhack [10], adjusts the analytical stiffness matrix using the modal parameters obtained from modal testing. The undamaged analytical stiffness matrix is adjusted by using the measured frequencies and mode shapes to estimate the damaged stiffness matrix, i.e. (superscripts are dropped)

$$\text{minimize } \theta = \frac{1}{2} \|\hat{M}_{au}^{\frac{1}{2}} (\hat{K}_{td} - \hat{K}_{au}) \hat{M}_{au}^{-\frac{1}{2}}\| \quad (5)$$

$$\text{subject to } \hat{K}_{td}^T \bar{\Phi}_{td} = \hat{M}_{au} \bar{\Phi}_{td} \Omega_{td}^2, \hat{K}_{td} = \hat{K}_{td}^T$$

where $\bar{\Phi}_{td}$ is the orthogonalized modal matrix. Solu-

tions to Equation 5 result in

$$\hat{K}_{id} = \hat{K}_{au} - \hat{K}_{au} \bar{\Phi}_{id} \bar{\Phi}_{id}^T \hat{M}_{au} - \hat{M}_{au} \bar{\Phi}_{id} \bar{\Phi}_{id}^T \hat{K}_{au} + \hat{M}_{au} \bar{\Phi}_{id} \bar{\Phi}_{id}^T \hat{K}_{au} \bar{\Phi}_{id} \bar{\Phi}_{id}^T \hat{M}_{au} + \hat{M}_{au} \bar{\Phi}_{id} \Omega_{id}^2 \bar{\Phi}_{id}^T \hat{M}_{au} \quad (6)$$

By comparing differences in the stiffness matrices of the test-adjusted *damaged* model and the initial analytical *undamaged* model, i.e. $\hat{K}_{id} - \hat{K}_{au}$, the location and extent of structural damage can be identified.

Case Studies

Performance of the proposed initial damage detection approach was demonstrated using test data from the NASA-LaRC 8-bay Dynamic Scale Model Technology (DSMT) model. Figure 2 shows a schematic of the cantilevered truss structure which was fully instrumented at 96 translational dofs and tested for 15 different damage conditions. Full instrumentation allowed the study

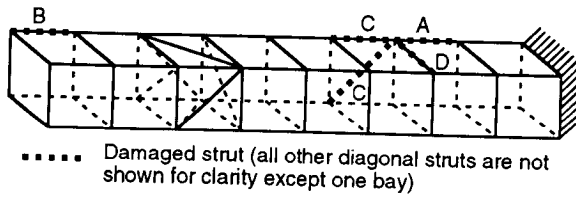


Figure 2: A Schematic of the NASA-LaRC 8-Bay DSMT Model.

of sensor placement, model reduction, and eigenvector expansion techniques. Types of damage include full, multiple, and partially damaged struts from (easy-to-locate) high-strain longerons to (difficult-to-locate) low-strain battens. During the study, several techniques were employed to enhance the damage detection capability: sensor placement, identification of bad sensors, an orthogonal procrustes (OP) technique [11] for mode shape expansion, and a parallel technique based on minimum rank perturbation theory (MRPT) [12].

The performance improvement from each enhancement for all 15 damage cases was examined and quantified [6]. Table 1 shows part of the damage analysis results, with two different enhancements, for four damage cases in which the stiffness loss of a removed member is 100%. For this study, five test modes at all 96 instrumented dofs were used without model reduction or eigenvector expansion. The baseline approach located two of five damaged members. The enhanced baseline process using iterative sparsity preservation improved the performance considerably except for the case with damage in a low-strain batten. The use of MRPT successfully located all the damaged members and provided an improved assessment of the stiffness loss in each case.

Case	Damage Type	Baseline	Enhanced Baseline	Enhanced MRPT
A	high-strain longeron	*	46%	91%
B	low-strain longeron	4%	37%	77%
C	multiple members - longeron/diagonal	*/21%	64/48%	95/93%
D	low-strain batten	*	*	62%**

* not located ** located with difficulty

Table 1: Damage Detection Results with Enhancements for the DSMT Model.

DETAILED DAMAGE DETECTION

The optimal-update approach, even with enhancements, identifies only a general area of structural damage when the instrumentation is limited. The design sensitivity approach provides damage detection information directly related to physical properties of the individual structural members by minimizing analytical/modified and test modal parameters. Without *a priori* information, however, this method requires extensive computational resources for large and complex structures. This approach was selected to identify a specific damaged structural component based on the initial damage detection results.

Theoretical Background

Figure 3 shows an analysis flow diagram for the detailed damage detection approach. The design vec-

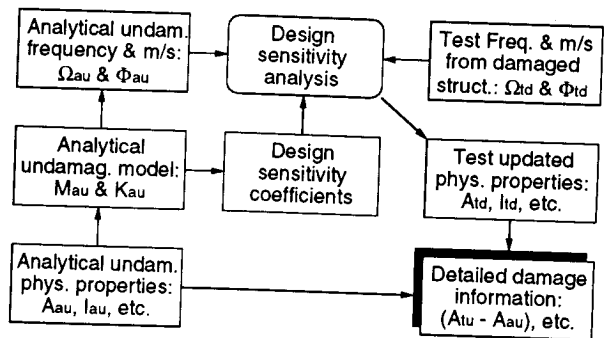


Figure 3: An Analysis Flow Diagram for the Detailed Damage Detection Approach.

tor, which includes physical properties such as cross-sectional area and moment of inertia, is defined as

$$\mathbf{u} = [u_1 \ u_2 \ \dots \ u_d]^T \in \mathbb{R}^{d \times 1} \quad (7)$$

The state vector, which includes frequencies and mode

shapes, is defined as

$$\mathbf{v} = [v_1 \ v_2 \ \dots \ v_s]^T \in \mathbb{R}^{s \times 1} \quad (8)$$

Using linear perturbation theory, the change of state vector, $\Delta \mathbf{v}$, due to the normalized change of design vector, $\Delta \mathbf{u}$, can be approximated without solving the eigenvalue problem, i.e. [5]

$$\Delta \mathbf{v} \doteq \mathbf{Q} \Delta \mathbf{u} \quad (9)$$

where \mathbf{Q} is called a sensitivity matrix and defined as

$$[Q_{ij}] = \left[\frac{\partial v_i}{\partial u_j} \right] \in \mathbb{R}^{s \times d} \quad (10)$$

The modified state vector, $\mathbf{v}^m \in \mathbb{R}^{s \times 1}$, with respect to the change in design vector can be calculated using Equation 9

$$\mathbf{v}^m = \mathbf{v}^o + \Delta \mathbf{v} \doteq \mathbf{v}^o + \mathbf{Q} \Delta \mathbf{u} \quad (11)$$

where $\mathbf{v}^o \in \mathbb{R}^{s \times 1}$ is the original state vector calculated from the undamaged analytical model.

The state vector error, \mathbf{e}^s , and the magnitude of design vector change, \mathbf{e}^d , are defined by

$$\mathbf{e}^s = \left| \frac{\mathbf{v}^t - \mathbf{v}^m}{\mathbf{v}^t} \right| \in \mathbb{R}^{s \times 1} \text{ and } \mathbf{e}^d = \left| \frac{\Delta \mathbf{u}}{\mathbf{u}^o} \right| \in \mathbb{R}^{d \times 1} \quad (12)$$

where $\mathbf{u}^o \in \mathbb{R}^{d \times 1}$ is the original design vector used in the undamaged analytical model and $\mathbf{v}^t \in \mathbb{R}^{s \times 1}$ includes frequencies and mode shapes obtained from modal testing of the damaged structure. The optimal design vector can be found by minimizing the difference between the test results and the modified analytical results with least changes to the analytical model, i.e.

$$\text{minimize } \theta = \mathbf{g}^T \mathbf{e}^s + \mathbf{h}^T \mathbf{e}^d \quad (13)$$

where $\mathbf{g} \in \mathbb{R}^{s \times 1}$ and $\mathbf{h} \in \mathbb{R}^{d \times 1}$ are weighting vectors for the state vector error and design vector change, respectively. Now the damaged structural component can be identified by examining the resulting design vector change. During the study, several techniques were employed to enhance the damage detection capability such as using a multiple set of design sensitivity coefficients.

Case Studies

Performance of the proposed detailed damage detection approach was demonstrated using test data from the 5.1-meter long hexagonal truss (HexTruss) which has limited instrumentation. Figure 4 shows the location and sensing direction of 48 accelerometers mounted on the 10-bay (5 full- and 5 half-hexagonal bays) HexTruss which was tested for 20 different damage conditions. A finite element model of 432 dofs

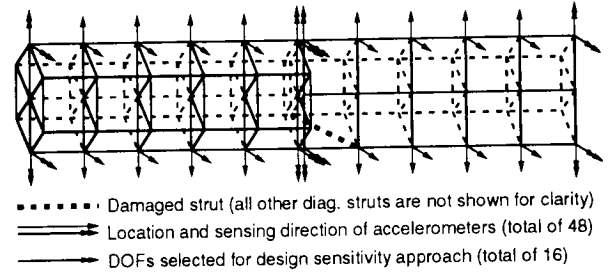


Figure 4: A schematic of the 10-Bay HexTruss.

was created for the undamaged configuration for use in the damage detection analysis. Analysis results for one damage case, which had a broken diagonal truss element in the half-hexagonal section, are discussed in the following.

For initial damage detection, test mode shapes of the intentionally damaged HexTruss were expanded to 144 dofs from 48 measured dofs using the dynamic expansion method [13]. The undamaged analytical system matrices were reduced to the same 144 dofs from 432 dofs using the Guyan reduction method [14]. Nine frequencies and mode shapes of the damaged structure along with the undamaged analytical model were used to detect and locate the structural damage using the simple Baruch method. The initial damage detection approach located the damaged truss bay, but did not locate the specific damaged truss member due to the limited number of measurements.

For detailed damage detection, based on the initial damage information, all 22 truss members contained in the suspect truss bay were selected (out of 224 members in the entire truss) as potential damaged components. This resulted in 88 design variables for consideration as potential damaged properties, out of at least 896 physical properties without preselection. Mode shapes were included in the state vector, because the specific damaged member could not be identified with frequencies alone. To further reduce the computational effort, two modes and 16 measurements were selected for the detailed damage detection among nine modes and 48 measurements used for the initial damage detection, respectively. As shown in Figure 5, the damaged member of the HexTruss was identified within one iteration.

CONCLUSIONS

A damage detection and health monitoring method for large and complex structures with limited instrumentation was presented. The method first identifies a general area of structural damage using optimal model update and a hybrid model reduction/eigenvector expansion technique. Then, it locates a specific dam-

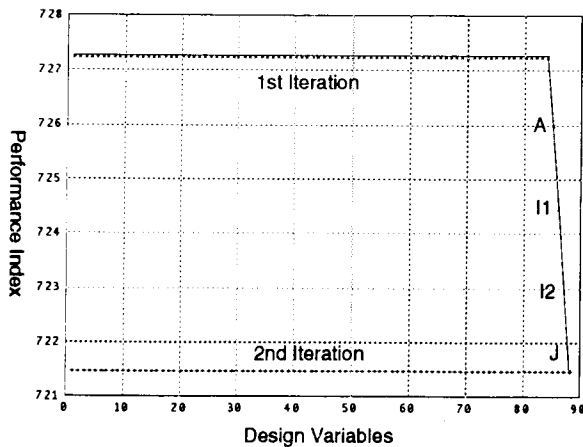


Figure 5: Convergence of Performance Index for the HexTruss.

aged structural component using a design sensitivity technique based on the initial damage information. Through the experimental validation, it was found that the proposed two-step damage detection approach can locate the specific damaged component of large structures with limited instrumentation. The same technology developed for space structures has a potential for structural inspection of aircraft, oil platforms, oil tankers, bridges, and buildings.

ACKNOWLEDGEMENT

The authors wish to express their sincere thanks to H.H. Doiron of MDA for many helpful discussions, T.A.L. Kashanggaki of NASA-LaRC for providing the DSMT test results, and D.A. VanHorn of MDA for extracting modal parameters from the HexTruss test data.

REFERENCES

- [1] Zimmerman, D.C. and Smith, S.W., "Model Refinement and Damage Location for Intelligent Structures," *Intelligent Structural Systems*, H.S. Tzou and G.L. Anderson, Editors, pp. 403-452, Kluwer Academic Publishers, The Netherlands, 1992.
- [2] Kim, H.M. and Doiron, H.H., "On-Orbit Modal Identification of Large Space Structures," *Sound & Vibration*, Vol. 26, No. 6, pp. 24-30, June 1992.
- [3] Kim, H.M., VanHorn, D.A., and Doiron, H.H., "Free-Decay Time-Domain Modal Identification for Large Space Structures," *Proc. 33rd AIAA SDM Conference*, Dallas, TX, pp. 1204-1211, Apr. 1992; to appear in *J. GUIDANCE*.
- [4] Kim, H.M. and Bartkowicz, T.J., "Damage Detection and Health Monitoring of Large Space Structures," *Sound & Vibration*, Vol. 27, No. 6, pp. 12-17, June 1993.
- [5] Kim, H.M. and Bartkowicz, T.J., "A Two-Step Structural Damage Detection Approach with Limited Instrumentation," to be presented at the *35th AIAA SDM Conference*, Hilton Head, SC, April 1994.
- [6] Zimmerman, D.C., Smith, S.W., Kim, H.M., and Bartkowicz, T.J., "An Experimental Study of Structural Damage Detection Using Incomplete Measurements," to be presented at the *35th AIAA SDM Conference*, Hilton Head, SC, April 1994.
- [7] West, W., "Structural Fault Detection of a Light Aircraft Structure Using Modal Technology," JSC Loads and Structural Dynamics Branch Report, Apr. 1988.
- [8] Flanigan, C.C., "Test/Analysis Correlation Using Design Sensitivity and Optimization," *Aerospace Technology Conference and Exposition*, SAE Paper No. 871743, Long Beach, CA, Oct. 1987.
- [9] Minas, C. and Inman, D.J., "Correcting Finite Element Models with Measured Modal Results Using Eigenstructure Assignment Methods," *Proc. 6th IMAC*, pp. 583-587, Feb. 1988.
- [10] Baruch, M. and Bar Itzhack, I.Y., "Optimal Weighted Orthogonalization of Measured Modes," *AIAA Journal*, Vol. 16, No. 4, pp. 346-351, Apr. 1978.
- [11] Smith, S.W., "Simultaneous Expansion and Orthogonalization of Measured Modes for Structure Identification," *Proc. 31st AIAA SDM Conference - Dyn. Spec.*, pp. 261-270, Apr. 1990.
- [12] Kaouk, M. and Zimmerman, D.C., "Structural Damage Assessment Using a Generalized Minimum Rank Perturbation Theory," *Proc. 34th AIAA SDM Conference*, La Jolla, CA, pp. 1529-1538, Apr. 1993.
- [13] Berman, A. and Nagy, E.G., "Improvement of a Large Analytical Model Using Test Data," *AIAA Journal*, Vol. 21, No. 8, pp. 1168-1173, Aug. 1983.
- [14] Guyan, R.J., "Reduction of Stiffness and Mass Matrices," *AIAA Journal*, Vol. 3, No. 2, pp. 380, 1965.

531 -39
44496

An Analytical Tool For Fracture Control

Raymond M. Patin & Royce G. Forman
NASA Johnson Space Center , Mail Code ES5
2101 NASA Road 1
Houston, TX 77058

ABSTRACT

The structural integrity of critical NASA space flight hardware is ensured through a damage tolerance program which is dictated in the fracture control requirements^[1-3]. The goal of a fracture control program is to ensure proper material selection and control, control of working stress levels, use of fracture resistant design concepts, manufacturing process control, and the use of qualified inspection procedures. The damage tolerance philosophy assumes that all components contain initial defects and the damage accumulation process is thus entirely crack propagation. The service capability of fracture critical components is determined by analysis and/or test. A large percentage of NASA fracture critical hardware is certified by analysis only and it is therefore imperative that the analysis code utilized contain the latest state of the art technology. Not all space contractors and vendors are able to develop and maintain an analysis code which meets this requirement. A response to this need has been fulfilled by the development of the NASA/FLAGRO (NASGRO) fracture mechanics code here at the Johnson Space Center. The NASGRO code is based primarily in Linear Elastic Fracture Mechanics (LEFM) and is continually updated to meet to needs of the user community. Development assistance comes from other NASA centers as well as the European Space Agency (ESA). The code currently has a vast materials data file (over 350 different entries) which defines the material strength, fracture toughness, crack growth rate, and crack growth threshold. The code also allows for user defined inputs of data from other sources. The code also provides a large library of stress intensity factor solutions for numerous geometries and loading conditions.

INTRODUCTION

The fatigue failure of a component occurs due to the application of cyclic stresses in which the maximum applied stress is significantly less than the ultimate and/or yield strength of the material. The fatigue life is typically characterized into three phases; (1) crack nucleation or initiation, (2) crack growth or propagation, and (3) final failure or instability. The problems associated with fatigue and methods of analyzing the phenomenon have been in place for quite some time. The stress-life (S-N), strain-life (ϵ -N), and LEFM

approaches have been utilized for approximately 100, 30, and 20 years respectively[4]. Each of these analysis methodologies has been characterized as being suited to particular industrial applications, and all are in wide use today. LEM addresses the crack propagation and instability stages of the component life and provides physical insights into the actual mechanisms of the fatigue process. Recent advances in fracture mechanics research[5-6] is resulting in the successful application of fracture mechanics for modeling the crack initiation portion of the fatigue life; therefore, fracture mechanics may soon become the sole analytical tool required for service life predictions. NASA/JSC has developed a computer code (NASGRO) which performs LEM analysis. Complete details about the NASGRO code can be found in the users manual[7].

THE STRESS INTENSITY FACTOR

The first correlation of failure load to the defect size present in a material was made by Griffith[8] in 1920 by means of an energy balance formulation. Irwin[9] later demonstrated the equivalence of the stress-intensity approach to the energy balance method. The stress intensity factor (denoted as K) is a parameter which reflects the redistribution of stress in a body due to the introduction of a crack; K represents the intensity of the stress field in the vicinity of the crack tip. The stress-intensity factor is expressed mathematically as:

$$K = \sigma \sqrt{\pi c} \cdot \beta(c/W) \quad (1)$$

where, σ is the applied stress, 'c' is the crack size, 'W' is the specimen width, and β is a non-dimensional factor which is dependent on the geometry and type of loading which is applied. NASGRO has a library of stress-intensity factor solutions for various geometries and loading conditions. There are ten (10) thru-thickness crack case solutions, one (1) embedded crack case, three (3) corner crack cases, ten (10) surface/thumbnailed crack cases, and seven (7) standard specimen crack case solutions. A schematic of a thru-thickness edge crack (TC02) is given in Figure 1. NASGRO currently has weight function solutions for 3 crack cases with more under development. A weight function solution allows for any stress gradient to be accurately modelled in the analysis.

Equation (1) is utilized in determining the crack size and static load required to fail a component once the fracture toughness is known. This type of analysis is called a residual strength assessment and it defines the maximum permissible defect the structure can tolerate. NASGRO also utilizes a net-section yield criteria in conjunction with LEM principles to provide realistic residual strength predictions for ductile materials.

FRACTURE TOUGHNESS

The fracture toughness is a measure of the crack extension resistance of a material. The level of elastic constraint at the crack tip, which is a function of the thickness of the component, dictates what the fracture toughness of a particular material will be. The case of maximum crack tip constraint (plane strain) provides a lower bound measure of fracture toughness which is denoted as the plane strain toughness (K_{Ic}). As the component thickness is reduced below the plane strain requirement a greater amount of lateral contraction at the crack tip occurs, and the fracture toughness increases due to the additional plastic work which takes place on loading. The variation of fracture toughness with thickness is depicted in Figure 2 along with the relationship NASGRO uses to model the phenomenon. The fracture toughness for the part-thru crack case is not fully plane strain nor is it fully plane stress. In the absence of fracture toughness data for such cases the effective toughness (K_{Ie}) is approximated with an empirical relationship.

FATIGUE CRACK GROWTH

When a cracked component is cyclically loaded such that the maximum stress intensity factor is less than the fracture toughness, stable (subcritical) crack propagation will occur. Paris et al.[10] established the correlation of the applied stress intensity range ($\Delta K = K_{max} - K_{min}$) to crack growth rate (dc/dN). The crack growth rate relationship used in NASGRO is:

$$\frac{dc}{dN} = \frac{C(1-f)^n \Delta K^n \left(1 - \frac{\Delta K_{th}}{\Delta K}\right)^p}{(1-R)^n \left(1 - \frac{\Delta K}{(1-R)K_c}\right)^q} \quad (2)$$

where, N is the number of cycles, ' c ' is the crack length, ' R ' is the stress ratio (minimum stress/maximum stress), ' f ' is the crack opening function, and C , n , p , and q are empirically derived constants. NASGRO has over 350 different materials which have been curve fit to equation (2) in both English and Metric units.

Numerical integration of equation (2), with regard to a prescribed geometry and loading spectrum, results in a prediction of the cycles required for the crack to grow from an initial size to the critical size.

SUMMARY

Within NASA, and many other organizations, the structural integrity of fracture critical hardware is ensured through the use of damage tolerance principles. The basic assumption in damage tolerance is that all critical hardware contains initial crack like defects. The size of the initial defect is quantified by Non-Destructive Evaluation (NDE), proof test, or by sensitivity test programs which determine the initial manufacturing defect sizes present in the components. The basic ingredients of a damage tolerance analysis are the residual strength calculation (critical crack size) and the crack growth prediction (service interval). The results of a typical damage tolerance analysis, which provides the relationship between crack size and life expended, are depicted in Figure 3.

NASGRO provides the capability to perform a complete damage tolerance assessment. The stress intensity factor library contains 31 different geometries which can be utilized for residual strength assessments and service life predictions. The crack growth rate equation utilized in NASGRO incorporates crack closure principles to model the stress ratio influence in the data and also accurately models the threshold and the instability regions of the data. The NASGRO equation is also fitted to over 350 different materials.

The NASGRO code is continually being updated with the latest fracture technology to maintain state of the art capability. An interagency working group consisting of researchers from the NASA (Johnson, Langely, and Marshall), FAA, NLR (Netherlands), Air Force, and Navy has been formed to coordinate the enhancements of NASGRO to meet the current and future needs in damage tolerance analysis. The current development plan includes crack growth retardation, two and three dimensional boundary element routines for stress analysis and stress intensity factor solution generation, elastic-plastic crack instability, and environmental effects on crack growth. An example of the boundary force stress analysis coupled with a weight function solution is provided in Figure 4.

REFERENCES

1. SD73-SH-0082A, Space Shuttle Orbiter Fracture Control Plan, Revised September 1974.
2. NHB 8071.1, Fracture Control Requirements For Payloads Using The National Space Transportation System (NSTS), September 1988.
3. SSP 30558, Fracture Control Requirements for Space Station, August 1991.
4. Bannantine, J.A., Comer, J.J., Handrock, J.L., Fundamentals Of Metal Fatigue Analysis, Prentice Hall, 1990.
5. Newman, J.C., Phillips, E.P., Swain, M.H., and Everett, R.A., "Fatigue Mechanics: An Assessment of a Unified Approach to Life Prediction", *Advances in Fatigue Lifetime Predictive Techniques*, ASTM STP 1122, Mitchell and Landgraf Eds., 1992, pp. 5-27.
6. Newman, J.C., "Fracture Mechanics Parameters for Small Fatigue Cracks", *Small Crack Test Methods*, ASTM STP 1149, Larsen and Allison Eds., 1992, pp. 6-33.
7. JSC-22267A (Draft), Fatigue Crack Growth Computer Program NASGRO Ver. 2.0, 1993.
8. Griffith, A. A., "The Phenomena of Rupture and Flow in Solids", Transactions, Royal Soc. London, Vol. 221, 1920.
9. Irwin, G.R., "Analysis of Stresses and Strains Near the End of a Crack Traversing a Plate", Transactions, Am. Soc. Mechanical Engrs., Journal of Applied Mechanics, 1957.
10. Paris, P.C., Gomez, R.E., and Anderson, W.E., 'A Rational Analytic Theory of Fatigue', The Trend In Engineering, Vol. 13, No. 1, University of Washington, 1961.

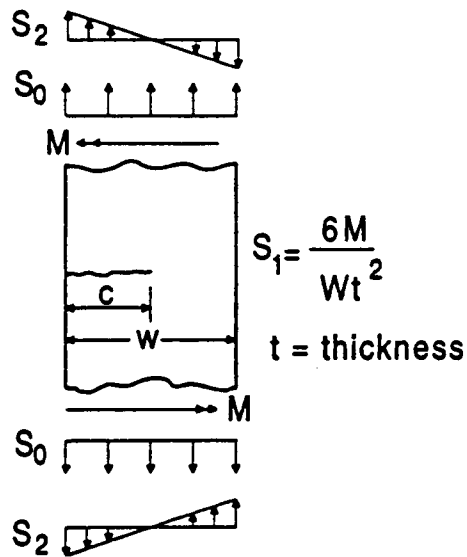


Figure 1: Thru-thickness edge crack case TC02.

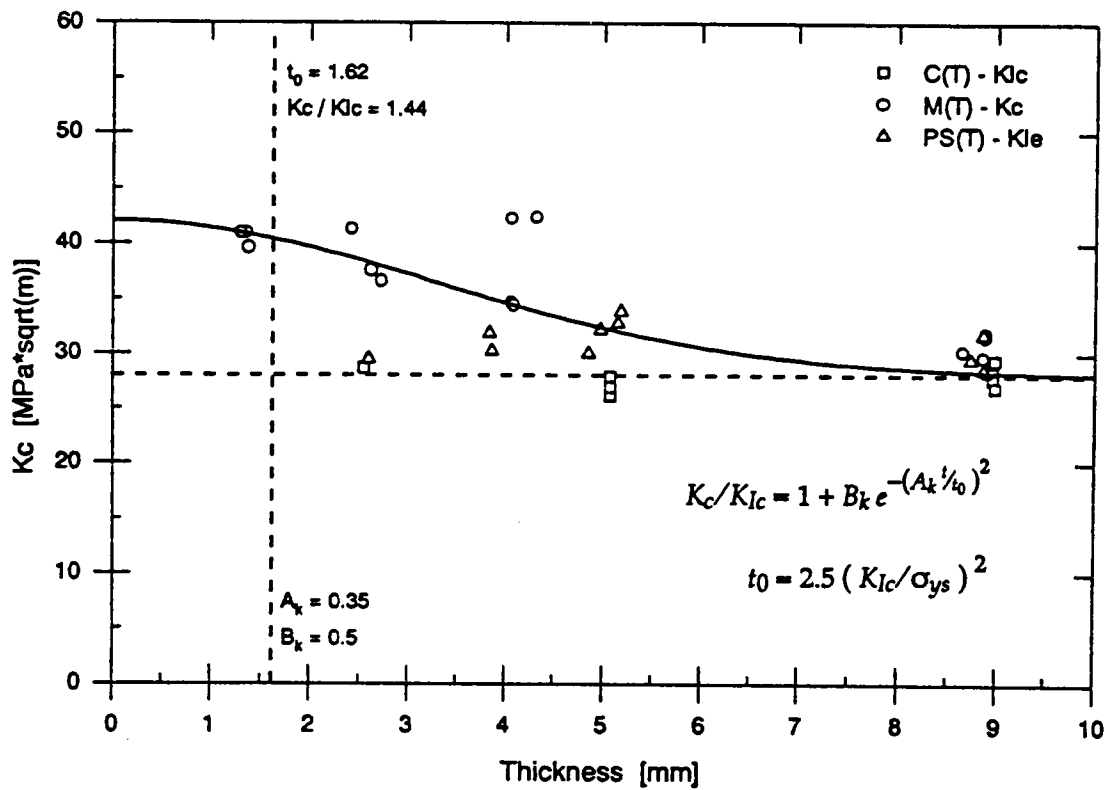


Figure 2: Variation of fracture toughness with specimen thickness.

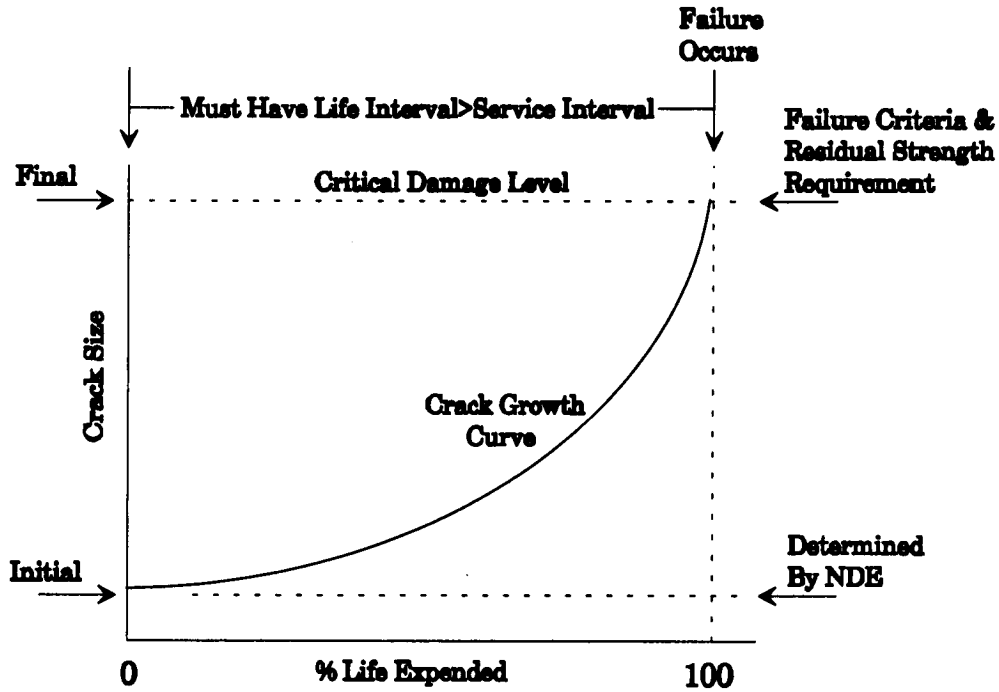


Figure 3: Schematic representation of damage tolerance analysis result.

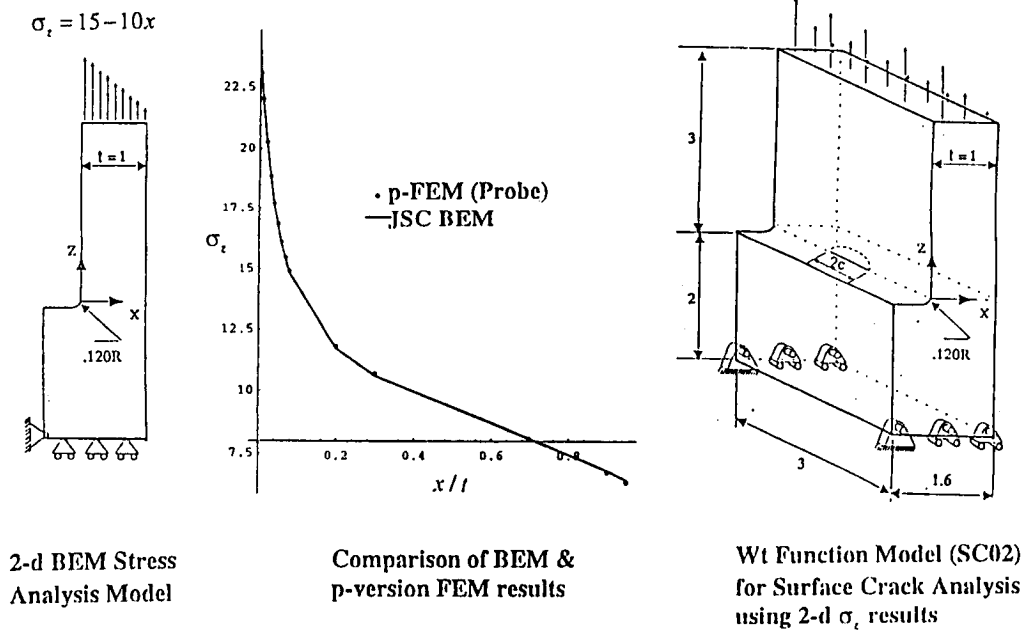


Figure 4: Example of boundary force stress analysis coupled with weight function stress intensity factor solution.

DESIGN AND FABRICATION OF NDE STANDARDS FOR FRACTURE CONTROL

Dual-Use Space Technology Transfer Conference And Exhibition

Charles L. Salkowski
NASA Johnson Space Center
Mail Code ES5
Houston, Texas 77586

ABSTRACT

Analysis of space flight hardware to determine safety, life, or damage tolerance, requires assumptions concerning flaws which may remain in a part after inspection. Flaw assumptions are based on Probability of Detection (POD) estimates for various geometry's and inspection methods. Acquisition of flaw POD data has historically been limited due to the cost of fabricating realistic flaw specimens. The NDE Laboratory at the Johnson Space Center has developed improved methods for fabricating flaw specimens. NASA has recently given approval for general industry to obtain flaw specimen sets at low cost.

INTRODUCTION

The structural integrity of critical NASA space flight hardware is ensured through a damage tolerance program in which component service capability is determined by analysis and/or test. A large percentage of the fracture critical hardware is certified by analysis only. The performance of deterministic fracture mechanics analysis to assess component life requires assumptions about the existence of initial flaws which may remain after inspection at the manufacturer or after inspection in-service. Initial flaw assumptions depend on the nominal capability of inspectors to detect flaw using Nondestructive Testing (NDT) techniques.

NASA HISTORY OF QUANTITATIVE NDE (QNDE)

NASA first incorporated fracture mechanics technology during the design of the space shuttle orbiter¹. Analysis of preliminary designs indicated the need for QNDE to determine initial flaw size assumptions. Three different efforts were performed to collect QNDE data. Due to the high cost of flaw specimens, each effort was performed using the same specimen set. However different analysis techniques were used to meet specified reliability levels (90% Reliability, 95% Statistical Confidence). The final effort was performed by the primary orbiter contractor, Rockwell International. Most flaw size assumptions were based on detection data from penetrant inspections, since it was by far the most widely used NDE method.

ACQUISITION AND ANALYSIS OF QNDE DATA

Inspector capability may be inferred by studying individual inspection results from tests using a set of specimens containing flaws with known dimensions. The nominal capability of the general population of inspectors (also called Probability of Detection, POD) is approximated by

grouping a large number of individual test results. This nominal capability is referred to in specifications as standard NDT. A standard NDT flaw size is determined for different flaw geometry's and NDT techniques. Occasionally, flaws smaller than the standard NDT capability must be detected in order to meet life cycle requirements. The capability to find these smaller flaws is called special NDT and must be demonstrated in person to a NASA representative.

A significant number of methods for analysis of QNDT data have been proposed over the last 20 years. NASA has traditionally used an approach based on a binomial distribution which results in a point estimation of POD². It can be shown that if 29 events with two possible outcomes are observed, and all 29 events result in the same outcome, the probability that the same outcome will be observed during any subsequent event under identical conditions is 90% with a corresponding 95% statistical confidence. Another approach was developed by the University of Dayton Research Institute (UDRI) under the direction of the U.S. Air Force³. This method is based on the principle of maximum likelihood. A statistical model of POD vs. flaw size was produced based on a large volume of inspection data. Subsequent data is evaluated by calculating input parameters to the model using maximum likelihood equations. The UDRI method may be used to generate a complete POD vs. flaw size curve for each inspector.

Each of these methods are useful under certain conditions. The binomial method is quick and easy to apply, particularly since only a limited number of specimens are required. However, each specimen set is designed to prove the POD of a specific flaw size. The UDRI method is a more complete evaluation of the capability of the inspector. Once the test is completed, the POD for each inspector can be determined for any flaw size. However, at least 80 specimens are required for each UDRI test. If the distribution of flaw sizes in the test set is not adjusted properly, equations to determine inputs to the model may not be solvable. Unfortunately, the adequacy of the flaw distribution is never known until after the test. The U.S. Air Force engine program is based on a calculated risk assessment, therefore knowledge of the inspectors capability to detect any flaw must be known. The UDRI method is called out specifically in the new Air Force Inspection Reliability Standard. By contrast, NASA has previously determined cutoff points for standard NDT and special NDT, therefore the binomial method is used most often for capability demonstrations.

FABRICATION OF POD SPECIMENS

A practical limitation all POD analysis methods have in common is the requirement for large sets of accurate flaw specimens. Fatigue cracks are generally considered the most difficult flaw to detect for most inspection methods and therefore are used most often in POD specimens. The process for growing a fatigue crack in a controlled environment can be expensive. Commercial costs have typically averaged \$1000 per cracked specimen. Consideration of different materials, part geometry's, flaw aspect ratios, and inspection methods, requires huge numbers of specimens. Previous POD data was collected using simple flat plate geometry's, one crack type, and limited inspection methods. POD for other geometry's and inspection methods was extrapolated from this data.

The NDT Laboratory at the Johnson Space Center has developed unique facilities for fabricating realistic flaw specimens, and accumulating and evaluating POD data. New methods for controlled growth of fatigue cracks have dramatically reduced time and cost investment. The primary innovation has been the development of the "scoop-rib" method for initiating fatigue cracks". This idea involves counter-sinking a notched, thin, rib of metal in the surface a flat plate which serves as an initiation point for a fatigue crack. The scoop-rib

method produces rapid initiation of a fatigue flaw of known dimensions. Continued growth under specific cycling conditions results in a flaw of known length and aspect ratio.

Significant efforts have also been expended in producing flaws in complicated geometry's such as holes, threads, and rods. One example involves producing a fatigue flaw in a large piece of plate stock from which a finished part is then machined. The finished part may be fabricated so that the flaw ends up at location desired. A continuous effort is in progress to produce large sets of flaw specimens in a large variety of materials and geometry's. POD data is collected on these specimens from inspections performed by NASA contractors. This data is used to update initial flaw size assumptions for fracture analysis.

SUMMARY

These innovations have led JSC to promote the widespread distribution of NDE specimens throughout the NASA community. Under this philosophy, all contractors are encouraged to demonstrate POD capability by requesting the loan of specimens from JSC. Test results remain confidential, and inspectors who do not meet the required confidence level can be retested after additional training. JSC standards are also available to general industry. Standards may be requested from the JSC NDT Laboratory by any organization on the condition that NASA is allowed to add the inspection data collected to the JSC flaw detection database. Under certain conditions, flaw specimens may be acquired permanently by outside organizations.

REFERENCES

1. "Space Shuttle Orbiter Fracture Control Plan", SD73-SH-0082A, Rockwell International Space Division, Revised September 1974.
2. Yee, B. G., Chang, F. H., Couchman, J. C., Lemon, G. H., and Packman, P. F., "Assessment of NDE Reliability Data", NASA CR-134991, NASA Lewis Research Center, October 1976.
3. Berens, A. P., Hovey, P. W., "Statistical Evaluation of NDE Reliability in the Aerospace Industry", Proceedings of the Review of Progress in Quantitative NDE", Williamsburg, Virginia, June 1987.

533-38
44518

MODAL TEST TECHNOLOGY AS NON-DESTRUCTIVE EVALUATION OF SPACE SHUTTLE STRUCTURES

Micheal S. Grygier
ES43
NASA Johnson Space Center
Houston, Texas

ABSTRACT

Modal test and analysis is being used for non-destructive evaluation of Space Shuttle structures. The purpose of modal testing is to measure the dynamic characteristics of a structure to extract its resonance frequencies, damping, and mode shapes. These characteristics are later compared to subsequently acquired characteristics. Changes in the modal characteristics indicate damage in the structure. Use of modal test technology as a damage detection tool was developed at JSC during the Shuttle acoustic certification program and subsequent test programs. The Shuttle Modal Inspection System was created in order to inspect areas that are impossible or impractical to inspect with conventional methods. Areas on which this technique has been applied include control surfaces, which are covered with thermal protection tiles, and the Forward Reaction Control Module, which is a frame structure that supports various tanks, thrusters, and fluid lines, which requires major disassembly to inspect. This paper traces the development of the technology, gives a status of its implementation on the Shuttle, explains challenges involved in implementing this type of inspection program, and suggests future improvements in data analysis and interpretation. Dual-use applications of the technology include inspections of bridges, oil-platforms, and aircraft.

1.0 INTRODUCTION

The rotating machinery industry has long used vibration characteristics of machines to monitor the machine health. Changes in vibration characteristics show wear. Consistent monitoring of these changes allows replacement of parts before a major breakdown occurs. Now NASA is using a similar technique to inspect Space Shuttle structures for damage. The technique is based on modal testing technology. Modal testing as applied to the Shuttle consists of exciting the orbiter vehicle with a random vibration signal and measuring its response at 350 locations. The data, as frequency response functions, is post processed to obtain resonance frequencies, damping, and mode shapes. This data set characterizes the baseline state of the structure. After several missions, the Shuttle is tested again. The second data set is compared to the

baseline. This comparison yields information that can indicate mass or stiffness changes in the structure. If no known changes were made, then the changes can be attributed to damage or wear.

NASA engineers have proved and developed the application of this type of testing for inspection purposes over a period of several years at the Johnson Space Center Vibration and Acoustic Test Facility. This paper will chronicle this development and the current implementation on the Orbiter.

2.0 TECHNOLOGY DEVELOPMENT

2.1 ORBITER ACOUSTIC FATIGUE CERTIFICATION TESTS

The idea for using modal test data for inspection purposes originated during the Shuttle Acoustic Fatigue Certification Tests during the late 70's and early 80's. The objective of the certification was to certify various components of the Orbiter for 100 missions. Modal testing was conducted on each component before acoustic testing to better locate strain gauges for the acoustic tests. A modal test engineer could look at the mode shapes and determine where high stress areas were, based on the location in the shape showing the most curvature. Each shape has its own characteristic stress state. Damage in the structure near the high stress areas for a particular mode would necessarily change the characteristics of that mode. Realization of this fact led to the practice of acquiring modal test data after each phase of acoustic testing for comparison to the pre-acoustic modal test results. This practice led to the discovery of various damage that was not found through conventional inspection processes. Refer to Table I for a summary of the tests where damage was indicated by modal test results[1],[2].

The most dramatic example of the utility of modal testing as an inspection tool occurred on the half-span body flap test article. The body flap is a control surface mounted on the aft of the Orbiter beneath the main engines and controls pitch during descent. Refer to Figure 1. It is a box assembly with aluminum ribs, a main spar, aluminum honeycomb skins, and a trailing edge wedge assembly comprised of a forward spar and full-depth honey-comb. The wedge assembly is

Table I. Orbiter Acoustic Certification Test experience showing damage found using modal testing.

Test Specimen	Summary of Results
WA-19 Wing Leading Edge	Initial Post-acoustic inspection showed no damage. Post-acoustic modal test showed significant changes from pre-acoustic modal test. Subsequent inspection disclosed cracked ribs.
AFA-15 Half-Span Body Flap (Stimulated Actuators)	Post-acoustic modal test showed decrease in cantilever mode frequency. Subsequent inspection revealed galling of bearings at actuator to rib interface.
AFA-15 Half-Span Body Flap (Flight Actuators)	Initial post-test inspection showed no damage. Post-acoustic modal test showed large decrease in spanwise bending mode frequency. Subsequent disassembly showed cracks in shear clips and hinge pins.
AFA-15 Half-Span Body Flap (Flight Actuators)	Repaired body flap was re-tested. Post-acoustic modal test showed increased damping of cantilever mode compared to earlier tests suggesting different actuators were installed. Actuators had been changed out with qual units with cracked gear teeth.
AFA-12 Aft Bulkhead	Visual post-acoustic inspection revealed wear of heat shield bearing surfaces and cracks in stiffeners. Post-acoustic modal test comparisons showed inconsistent mode shapes in areas of damage.

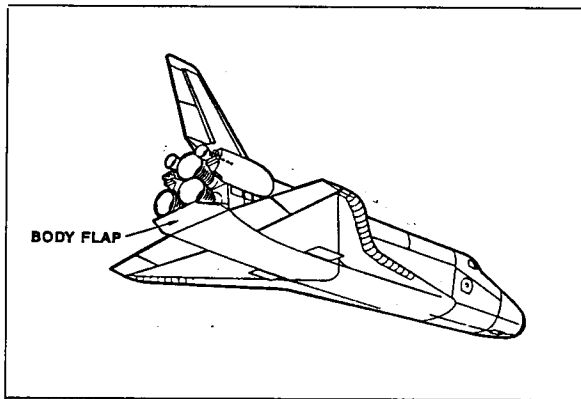


Figure 1. Location of body flap

connected to the box assembly by shear clips to the ribs and piano hinges top and bottom to the skins. There are four actuators that attach the body flap to the aft fuselage stub of the Orbiter. Refer to Figure 2. The test article consisted of a half-span body flap with two actuators attaching it to an aft fuselage stub structure, which in turn was mounted to a steel billet.

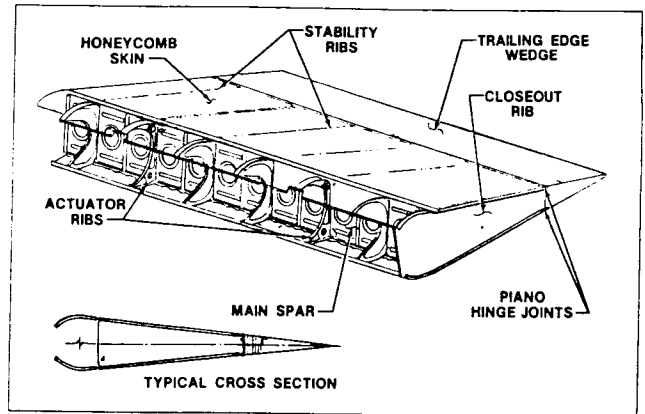


Figure 2. Half-span body flap assembly

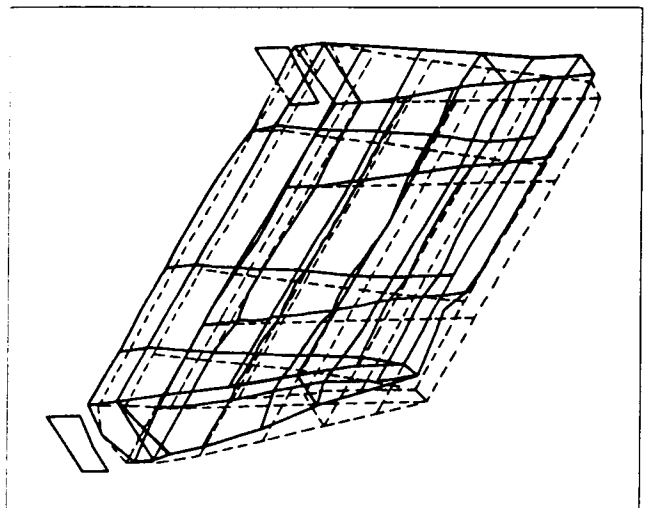


Figure 3. Half-span body flap mode 1 - first cantilever

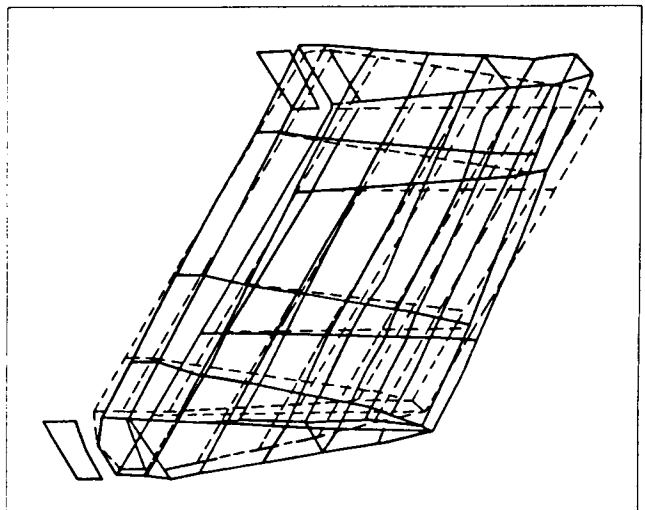


Figure 4. Half-span body flap mode 2 - torsion

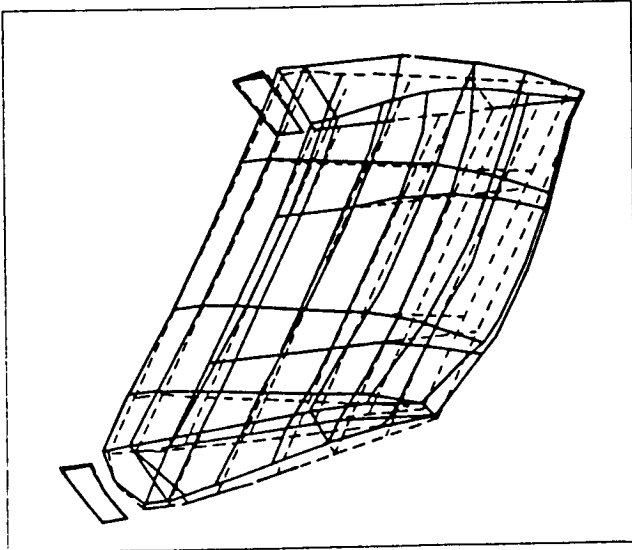


Figure 5. Half-span body flap mode 3 - spanwise bending

Modal testing of the half-span body flap test resulted in several mode shapes, three of which are important for this discussion. The first mode is a cantilever bending of the body flap about the actuators, which were mounted to a steel billet (Figure 3). The second mode is a torsion mode with a node line running front to back at the centerline (Figure 4). The third mode is the first spanwise bending mode with two node lines running front to back along each of the two ribs attached to the actuators (Figure 5). The frequencies of the first two modes are determined by the stiffness of the actuators and interface. The frequency of the third mode is determined by the spanwise stiffness of the assembly.

Three series of acoustic tests were conducted on the test article. During the first series, the actuators used were simulated non-flight-like solid steel actuators. After the first series of acoustic tests, the modal test data showed a decrease in the first cantilever mode frequency. This indicated that a loss of stiffness had occurred between the aft fuselage stub and the body flap. Initial inspections did not show any evidence of damage. A subsequent tear-down inspection, however, did uncover "galled" bearings in the actuator to body flap interfaces.

For the second series of tests, flight type actuators were used. Modal test data before and after the acoustic test showed a marked reduction in frequency of the spanwise bending mode. This was reported at the time with the following possible causes: 1.) honeycomb skin separation from core at the trailing edge and/or forward box assembly, 2.) failure of the trailing edge wedge assembly to forward box assembly

interface, 3.) failure of the spar of the trailing edge wedge assembly. A follow-on inspection using visual, x-ray, and ultrasonics did not yield any indication of damage except two small dents in the upper surface of skin. Consequently, a tear-down inspection was completed. This inspection showed extensive damage. Of the 7 shear clips connecting the trailing edge wedge assembly to the forward box assembly, 5 had cracks ranging from 0.25 to 0.875 inches long [3]. Also, the piano hinge pins were both cracked in 2 places. The 2 shear clips that were not cracked were the ones on the ribs attached to the actuators. Note that these were at node lines for the spanwise bending mode.

The third series of tests occurred after the test article had been repaired. This time the post-acoustic modal test showed an increase in damping of the cantilever mode. This was attributed to changes in the actuators or their interfaces. It was subsequently revealed that the actuators used had been also been used in qualification testing and contained cracked gear teeth.

2.2 ENTERPRISE MODAL TESTING

To further develop and evaluate modal testing as an inspection tool, a set of tests was conducted on the Orbiter Enterprise. The objectives of that test were to develop efficient means of modal testing and to evaluate modal property sensitivity to known damage. Thirteen cases of fault simulations were conducted. These simulations included removal of shear pins, wing truss tubes, various fasteners, web material, and honey-comb skin. Seven of the thirteen caused noticeable changes to the modal response data. Four of these caused significant changes. One explanation for not detecting more of the changes involves the non-stationarity of the data. There were certain environmental factors that could not be controlled which also affected the data. These included large temperature deviations from morning to night as well as ground vibrations induced by nearby quarry blasting[4].

2.3 CESSNA MODAL TESTING

Because of the problems associated with getting repeatable modal data in an uncontrolled environment, a modal test program was undertaken to establish the limits of repeatability. The test article for this series of tests was a used Cessna airframe. The temperature was maintained to 70.9 to 74.5 degrees Fahrenheit. Over a period of one month, modal response data was acquired 20 times. Overlays of the frequency response data shows little change. A similar experiment was conducted while varying the temperature 61 to 74.5 degrees. This time the overlays were not as repeatable.

A follow-on study was then conducted with the objective of further evaluating modal property sensitivity to known damage while in a controlled environment. Of nine damage cases tested, all showed changes in the frequency response and mode shape data. Two were not significant enough to flag as faults. Direct excitation of each control surface was suggested as a way to improve the fault detection process for these two cases. The conclusion from this program was that damage was detectable by modal test evaluation before it became critical[5].

2.4 COLUMBIA AND DISCOVERY BODY FLAP TESTS

Based on the empirical data and experience gained from all these tests, NASA engineers decided that the Orbiter structural inspection program could be improved by implementing a modal test inspection program. It was concluded that this type of inspection would augment the existing inspection program by monitoring large areas of structure without having to remove thermal protection tiles. This is particularly advantageous for the Orbiter program because some recommended inspections could not be done otherwise. The Orbiter control surfaces were identified as most applicable for this type of inspection since large areas of these surfaces are made up of honeycomb panels. These include the elevons, body flap, vertical tail, and rudder/speed brakes.

An early opportunity to carry out this type of inspection came following STS-28. An anomaly during ascent was identified involving large oscillations of the body flap. As a result of the investigation, a modal test was performed on the body flap of Columbia and Discovery in September 1989. Results from these tests showed that Columbia's body flap to fuselage interface was less stiff than Discovery's. It was also shown that Columbia's body flap interface was non-linear, more so than Discovery's. It was recommended that the actuators be inspected for wear [6]. Following the tests, two of the actuators were replaced on Columbia. The two actuators removed were disassembled and inspected. Some gear teeth were found to show plastic deformation on their surfaces. This phenomenon was not expected. Analysis showed that if this had progressed for 22 missions, the gear teeth would have failed.

2.5 SHUTTLE MODAL INSPECTION SYSTEM ACCEPTANCE AND TRAINING

A mobile modal test system was procured and delivered in January 1990. Acceptance testing was conducted at the Kennedy Space Center (KSC) using the Cessna

airframe as a test article. Subsequent procedure development, training, and certification were conducted at JSC using a quarter-scale orbiter model. The Shuttle Modal Inspection System (SMIS) was then moved to KSC.

3.0 CURRENT IMPLEMENTATION ON ORBITER

3.1 OVERVIEW OF SMIS TESTS PERFORMED

The SMIS is currently used for inspections of the orbiter control surfaces and the Forward Reaction Control System (FRCS) module. The control surface tests are performed once after the Orbiter maintenance period to establish the baseline data set and again before entering the next maintenance period. The maintenance periods are usually 3 years and 10 missions apart. The rationale for this spacing is to obtain inspection results that can help guide the detailed inspections scheduled for the maintenance period. If damage is found, then there will be enough time in the maintenance period to fix it.

To date nine SMIS tests have been performed. Three baseline control surface tests have been performed. No pre-maintenance tests have been performed yet. Three other tests have been performed on body flaps. Repeat tests on body flaps have been hampered by lack of consistent configuration and non-linearities inherent in the body flap to fuselage interfaces. To date, the repeat tests show that the body flap structure proper is not degrading. These tests have also shown that the data can be repeatable despite the time and missions elapsed between tests. Two tests have also been performed on FRCS modules.

3.2 OPERATIONAL CONSTRAINTS

There are certain operational constraints required to successfully perform a modal inspection test. These constraints include vehicle configuration, quiet vehicle during testing, environmental control, and set-up time.

Because an inspection based on modal test technology requires two data sets for comparisons, a consistent vehicle configuration is required for each test. A change in configuration, such as removal of a large component, changes the mass and possibly stiffness distribution of the structure. These changes manifest themselves in the modal data. Such configuration changes would then possibly mask changes in the data due to damage.

A quiet vehicle is required during data acquisition. This is required because the excitation level used for modal testing is very low. Any hands-on work on the vehicle

during data acquisition would contaminate the data. Because of this constraint, the SMIS system was designed to acquire, store, review, and reacquire any suspect data in an eight-hour window.

Two other constraints are environmental control and set-up time. The Orbiter Processing Facility (OPF) maintains a constant temperature and humidity environment within the tolerances needed for repeatable data. Set-up time is approximately ten days. This set-up can occur in parallel with most other OPF activities.

3.3 SYSTEM PERFORMANCE

Despite the operational constraints, the system has performed as designed. With one exception, the tests have all been completed in the time planned. During one test, an abnormally high number of suspect data points were identified and reacquired. This test took nine hours to accomplish. This performance required much hard work and preparation.

The following factors contributed to the successful implementation: proper design of the system, detailed procedure development, extensive training, and frequent scheduling meetings with the people in charge of vehicle operations.

4.0 GUIDELINES FOR APPLICATION TO NON-SHUTTLE STRUCTURES

Implementation of a modal-test-based inspection program requires an outlay of capital or use of specialized contractors who have the equipment and expertise. The equipment and software required to acquire and analyze the data is very specialized, although off-the-shelf equipment and software are readily available to conduct modal testing and analysis. A team of highly skilled analytical engineers is also required.

The first test on a structure will require two to three weeks to set-up and optimize the data acquisition parameters. The first test will set the standards for all subsequent tests. It is also a good time to finalize any configuration control issues. Remember, this first test does not constitute an inspection. The inspection technology is based on comparing a subsequently acquired data set to this first baseline. However, some anomalies in this first data set can indicate wear or damage in the structure.

Some special challenges that could be faced while implementing this type of inspection include the following. Structures with many components tend to

have many modes. High modal density will require more analysis time to extract the modal parameters and more judgement in interpreting changes in the modes. Structures which exhibit non-linear dynamic characteristics can also cause difficulty in data interpretation.

5.0 FUTURE DEVELOPMENTS

The current implementation of modal inspection for the Orbiter relies on the classic modal inspection techniques. That is, modal frequencies and damping values are compared as well as shape-to-shape comparisons. The Modal Assurance Criteria (MAC) is used as a quantitative measure of shape consistency. These comparisons will indicate if significant damage has occurred and isolate it to a particular area on a component. Additional analytical tools are being developed which will allow better resolution of damage location and damage assessment. There are many projects on-going to develop these methods even outside NASA. However, most of these methods require the use of a correlated finite element model [7-10].

6.0 CONCLUSION

NASA experience shows that modal test technology can be used as a non-destructive evaluation technique for structures. Modal inspection techniques are most suitable for structures that are inaccessible to other means of inspection. Modal inspection can also guide the application of more detailed inspection methods by screening large areas for damage. A modal inspection system can be integrated successfully into an operations environment.

Modal inspection does require a commitment to use specialized equipment, software, and personnel. There are certain constraints to modal inspection including structure configuration control, quiet structure during testing, environmental control, and set-up time.

The current damage detection algorithms are useful, but more advanced algorithms are being developed to improve the damage location and assessment capabilities of modal inspection.

7.0 ACKNOWLEDGMENTS

The author would like to thank Walter West for his dedication during the last 16 years to developing modal inspection technology; Stan Weiss for his commitment to implementing modal inspection on the Orbiter; Charles Mitchell for his hard work analyzing data and

developing systems for SMIS analysis, and Robert Coleman for ideas and technical direction.

REFERENCES

1. Hunt, D.; Weiss, S.; West, W.; Dunlap, T.; Freesmeyer, S., "Development and Implementation of a Shuttle Modal Inspection System," *Sound and Vibration*, Bay Village, Ohio, Vol. 24, Issue 8, August, 1990, p 34.
2. West, W., "Illustration of the Use of Modal Assurance Criteria to Detect Structural Changes in an Orbiter Test Specimen," *Proceedings of Air Force Aircraft Structural Integrity Conference*, Warner Robins Air Logistics Center, Robins AFB, Georgia, November 1984.
3. Cheng, J., "Life Certification Analysis for Orbiter Body Flap," STS 84-0128, Rockwell International, Contract NAS9-14000, May 1984, p 51.
4. West, W., "Fault Detection in Orbiter OV-101 Structure and Related Structural Test Specimens," NASA JSC Loads and Structural Dynamics Branch Report, September 1984, pp 35,36,44.
5. West, W., "Structural Fault Detection of a Light Aircraft Structure Using Modal Technology," NASA JSC Loads and Structural Dynamics Branch Report, April 1988, pp 27,28.
6. Brillhart, R., "Summary Report of Space Shuttle Orbiter OV-102 and OV-103 Body Flap Modal Survey Investigation For Damage Assessment," SDRC No. 41039, SDRC Engineering Services Division, Inc., San Diego, California, September 1989, p 9.
7. Stubbs, N.; Kim, J.; Topole, K., "Nondestructive Damage Analysis Using Modal Parameters of Orbiter OV-102 Body Flap," Techcom, Inc., College Station, Texas, NASA contract NAS 9-18605.
8. Stubbs, N.; Broome, T.; Osegueda, R., "Nondestructive Construction Error Detection in Large Space Structures," AIAA-88-2460, AIAA SDM Issues of the International Space Station, Williamsburg, Virginia, April 1988.
8. Ricles, J., "Nondestructive Structural Damage Detection in Flexible Space Structures Using Vibration Characterization," NASA Contract NGT-44-001-800, August 1991.
9. Kim, H.; Bartkowicz, T., "Damage Detection and Health Monitoring of Large Space Structures," AIAA SDM Conference, La Jolla, California, April 1993.
10. Zimmerman, D; Kaouk, M., "Eigenstructure Assignment Approach for Structural Damage Detection," *AIAA Journal*, Vol. 30, No. 7, 1992.

534-26
1-21

DEVELOPMENT OF ADVANCED ALLOYS USING FULLERENES

J. Sims, M. Wasz, J. O'Brien, D. L. Callahan, and E. V. Barrera
Department of Mechanical Engineering and Materials Science
Rice University, Houston, TX 77251

ABSTRACT

Development of advanced alloys using fullerenes is currently underway to produce materials for use in the extravehicular mobility unit (EMU). These materials will be directed toward commercial usages as they are continually developed. Fullerenes (of which the most common is C₆₀) are lightweight, nanometer size, hollow molecules of carbon which can be dispersed in conventional alloy systems to enhance strength and reduce weight. In this research, fullerene interaction with aluminum is investigated and a fullerene-reinforced aluminum alloy is being developed for possible use on the EMU. The samples were manufactured using standard commercial approaches including powder metallurgy and casting. Alloys have been processed having 1.3, 4.0 and 8.0 volume fractions of fullerenes. It has been observed that fullerene dispersion is related to the processing approach and that they are stable for the processing conditions used in this research. Emphasis will be given to differential thermal analysis and wavelength dispersive analysis of the processed alloys. These two techniques are particularly useful in determining the condition of the fullerenes during and after processing. Some discussion will be given as to electrical properties of fullerene-reinforced materials. Although the aluminum and other advanced alloys with fullerenes are being developed for NASA and the EMU, the properties of these materials will be of interest for commercial applications where specific Dual-Use will be given.

INTRODUCTION

The objective of this project is to produce fullerene-reinforced materials for space and commercial applications using conventional processing methods. These new materials are based on the use of fullerenes, of which the most common is C₆₀, which are lightweight, nanometer size, hollow molecules of carbon which can be dispersed in conventional alloy systems to enhance strength and reduce weight. The fullerene-reinforced materials would be used to manufacture components of a prototype Primary Life Support System (PLSS) of the Extravehicular Mobility Unit (EMU). The emphasis of this paper will be on the aluminum-fullerene system which in design is to have the following properties: fullerenes dispersed throughout the metal matrix that are undisturbed (unreacted) by the processing and with mechanical properties enhanced by their addition. Discussion will be given as to the motivation of this research, processing techniques used, characterization techniques applied for the study of fullerenes and possible industries where technology transfer could occur.

RESEARCH MOTIVATION

In 1985, Kroto *et al.*, [1] discovered a third form of carbon. This is the molecular form of carbon generically known as fullerenes. Fullerenes are carbon molecules that are arranged in spherically shaped pentagons and hexagons resembling a soccer ball as shown in Figure 1. In 1990, Krätschmer *et al.*, [2] discovered a way to produce macroscopic quantities in

the solid form. Once the fullerenes were relatively abundant, research into their properties and usefulness accelerated. Their properties such as nanometer size, density, and chemical stability are the motivation for this research because they are properties which could lead to enhancement of mechanical properties. C₆₀ forms an FCC crystalline lattice with a density of 1.67 g/cm³ and a lattice parameter of 1.42 nm [1,2,3]. These molecules in the crystalline form are separated by van der Waals forces. This separation gives the individual molecules an effective diameter of 1.0 nm. Each molecule is a truncated icosahedron with an open volume having a 0.35 nm diameter. The density of the C₆₀ molecule has been calculated to be 1.2 g/cm³ which makes it a very light alloying addition or composite reinforcement. Thermally, these fullerenes are structurally stable to 2500°C [4] and sublime in vacuum at temperatures as low as 350°C maintaining structural integrity [5]. Chemically, they have a range of reactions from no reaction (in some acids) to carbide like reaction depending on the metal they are in contact with [6-10]. Currently, aluminum alloys are being used in the PLSS of the EMU where lighter and stronger materials would reduce the payload and save in fuel cost for the Space Shuttle Program.

PROCESSING OF ADVANCED ALLOYS

Powder metallurgy (PM) and casting approaches have been used to process fullerene-reinforced aluminum alloys. For these materials two levels of fullerene concentration can be considered which would separate the materials into either alloys (fullerene concentration less than 10%) or composites (fullerene concentration greater than 10%). Alloys are strengthened by second phase strengthening mechanisms such as dispersion hardening [11,12]. Composites are strengthened by added reinforcements which generally show a rule of mixtures relationship [13]. For this phase of the project, the alloying approach was taken.

Samples of 1.3, 4.0 and 8.0 volume fraction (v/o) fullerenes were produced using both standard solid state powder metallurgy (PM) and aluminum melt techniques (casting). Powders of fullerene extract (C₆₀-C₇₀ mixture) were mixed with pure aluminum powder and pressed into green compacts. The compacts were sintered to achieve consolidation. A two stage heat treatment was performed on the compacts. In the initial stage, a temperature above the fullerene sublimation temperature (500°C) but below the normal sintering temperature of aluminum was used to promote interparticle dispersion of the fullerenes. In the final stage, the temperature was ramped to near the melting temperature of the aluminum (600-650°C) to promote aluminum consolidation.

For the casting approach, both powders and ingots were used. Temperatures for the melting process ranged from just over the melting temperature of 660°C to almost 1100°C. Heat treatment times varied from 15 minutes to 1 hour at the specified temperature. Samples were processed in both air and under flowing dry argon.

CHARACTERIZATION TECHNIQUES AND DISCUSSION

In this research, the final desired material would possess properties where fullerenes are dispersed throughout the metal matrix and undisturbed (unreacted) by the processing. Their dispersion would lead to enhanced mechanical properties such as high strength and stiffness. To investigate the fullerene-reinforced aluminum alloys optical and scanning electron microscopy (SEM) was used to observe the as-processed microstructures and wavelength dispersive spectrometry (WDS) using an electron microprobe was used to

evaluate fullerene dispersion and stability. Differential thermal analysis (DTA) was used to further confirm the condition of the fullerenes over an even broader temperature range.

For the microstructure, optical microscopy was the primary means of characterization because sample charging and matrix masking of the signal made SEM analysis less useful. Figure 2 shows the difference in microstructure of three samples processed by a first stage heat treatment of 500°C for 24 hours and a second stage heat treatment of 600°C for 24 hours, where different volume fractions of fullerenes were used. The grain boundaries are optically enhanced by the presence of the fullerenes. The large black patches in the micrograph for the 8.0 v/o sample are fullerene agglomerations due to the saturation of the aluminum matrix. These agglomerations disappear as the volume fraction is decreased. The clarity of the grain boundaries also decreases with volume fraction. As a means of quantifying fullerene dispersion, WDS line scans were performed. Figure 3 is representative of a line scan across a grain of a PM sample. The points at which the carbon content is highest are indicative of a grain boundary. Fullerene dispersion was limited to grain boundary dispersion in this case where exact concentrations within the grains are a subject of continued study. It is expected that some fullerenes are present within the grains. The actual concentration is low and inhomogeneous and has not been reasonably detected by WDS even though calculations of the volume of fullerenes in saturated grain boundaries suggest that fullerenes must also be present in the grains.

It was found that by using an electron microprobe, that the characteristic WDS peak of the fullerenes was clearly distinguishable from either the diamond or the graphite peak. Figure 4 shows the comparison of the WDS carbon peaks for diamond, graphite and fullerenes. To confirm that there was no reaction between the aluminum and the fullerenes, DTA was used. Figure 5 shows DTA data for fullerenes, aluminum, and an aluminum-fullerene alloy. The sharp dip at ~660°C is the melting of aluminum. Similar dips and peaks are indicative of thermal events such as phase changes and chemical reactions like carbide formation. Other than the dip for the melting of aluminum, there are no other thermal events. DTA analysis over a higher temperature range showed that there were no other thermal events until 1100°C was reached.

Vickers's microhardness was used to determine the extent of mechanical property enhancement. This preliminary testing method was used because only small samples have been produced in this research. Figure 6 is a microhardness profile across a PM sample. Note that there is a significant increase in the average hardness number ranging from Vickers hardness (VH) 24 for the monolithic aluminum encasing the alloy to about VH 43 for the alloy. It is evident that the presence of the fullerenes enhances the hardness of the sample. The scatter observed in the data is indicative of fullerene inhomogeneity. Hardness values for the melted samples are similar, however, the samples tend to have higher porosity. The result of these analyses was the confirmation of the observations made by microprobe analysis.

TECHNOLOGY TRANSFER

There are countless innovations from space technology that have already found their way into commercial aircraft and the advent of the Dual-Use program will further that directive. Development of new and advanced materials for structural applications in current technologies means that properties such as high strength, high stiffness, and enhanced toughness must be obtained where weight saving is of importance. The materials being developed in this research are proposed to replace conventional aluminum alloys used on the PLSS of the EMU where weight savings is of interest. This material may also be used

in any industry in which these properties are desirable. The aircraft industry relies heavily on lightweight materials with high strength-to-weight ratios where this material would be useful. The automotive industry also uses lightweight materials to reduce fuel consumption and would also have an interest in these materials. Currently both an end user and processor company are sought to both prototype the PLSS and commercially produce the fullerene-reinforced materials.

SUMMARY

In this paper, ongoing work by Rice University to develop fullerene-reinforced materials was presented. It was observed that limited fullerene dispersion was achieved by PM and casting processing. Dispersion was interparticle (in grain boundaries) in nature as evidenced by optical microscopy of the processed samples. The fullerenes were undisturbed by the processing as observed by WDS and DTA analysis. Microhardness was shown to increase because of the presence of the fullerenes but this was limited by the small extent of the fullerene dispersion within the aluminum grains. As this project continues, alloy optimization will be used to develop a prototype of components of the PLSS for the EMU. Technology transfer will become a focus and applications for fullerene-reinforced aluminum will be identified.

ACKNOWLEDGMENTS

The authors would like to acknowledge Shiou-Jyh Hwu, Kristen Etheridge, Cordelia Bucher of the Chemistry Department and Milton Pierson of the Geology Department at Rice University for their help. Funding for this research was provided by NASA-Johnson Space Center, Houston, Texas through contract no. NAG9-628. The Nasa project monitors have been Evelyn S. Orndoff, Frederic Dawn and Chin Lin.

REFERENCES

1. Kroto, H.W.; Heath, J.R.; O'Brien, S.C.; Curl, R.F. and Smalley, R.E.; "C₆₀: Buckminsterfullerene," **NATURE**, Vol. 318, 1985, pg. 162.
2. Krätschmer, W.; Lamb, L.D.; Fostiropoulos, K. and Huffman, D.R.; "Solid C₆₀: a New Form of Carbon," **NATURE**, Vol. 347, 1990, pg. 354.
3. Kortan, A.R.; Kopylov, N.; Glarum, S.; Gyorgy, E.M.; Ramirez, A.P.; Fleming, R.M. and Haddon, R.C.; "Superconductivity at 8.4 K in Calcium-doped C₆₀," **NATURE**, Vol. 355, 1992, pg. 354.
4. Zhang, B.L.; Wang, C.Z.; Chan, C.T. and Ho, K.M.; "Thermal Disintegration of Carbon Fullerenes," **PHYSICAL REVIEW B**, Vol. 48, No. 15, 1993, pg. 381.
5. Ismail, I.M.K. and Rodgers, S.L.; "Comparisons Between Fullerene and Forms of Well-known Carbon," **CARBON**, Vol. 30, No. 2, 1992, pg. 229.
6. Smalley, R.E.; "Great Balls of Carbon," **THE SCIENCES**, Mar./Apr., 1990, pg. 22.

7. Altman, E.I. and Colton, J.; "Determination of the Orientation of C₆₀ Adsorbed on Au(111)," **PHYSICAL REVIEW B**, submitted.
8. Altman, E.I. and Colton, J.; "Characterization of the Interaction of C₆₀ with Au(111)," **ATOMIC AND NANOSCALE MODIFICATION OF MATERIALS**, ph. Avoures, ed., NATO ASI Series, in press.
9. Ohno, T.R.; Chen, Y.; Harvey, S.E.; Kroll, G.H.; Weaver, J.H.; Haufler, R.E. and Smalley, R.E.; "C₆₀ Bonding and Energy Level Alignment on Metal and Semiconductor Surfaces," **PHYSICAL REVIEW B**, Vol. 44, No. 24, 1991, pg. 747.
10. Ohno, T.R.; Chen, Y.; Harvey, S.E.; Kroll, G.H.; Weaver, J.H.; Chibante, L.P.F. and Smalley, R.E.; "Metal-Overlayer Formation on C₆₀ for Ti, Cr, Au, La and In," **PHYSICAL REVIEW B**, Vol. 47, No. 4, 1992, pg. 2389.
11. Hertzberg, R.W.; "Strengthening Mechanisms in Metals," **DEFORMATION AND FRACTURE MECHANICS OF ENGINEERING MATERIALS**, Wiley and Sons, New York, NY, 1989, pg. 130.
12. Ashby, M.F.; "The Theory of Critical Shear Stress and Work Hardening in Dispersion Hardened Crystals," **OXIDE DISPERSION STRENGTHENING**, Ansell, G.S.; Cooper, T.D. and Lenel, F.V., eds.; Met. Soc. Conf., Gordon and Breach, 1966, pg. 143.
13. Jones, R.M.; "Micromechanical Behavior of a Lamina," **MECHANICS OF COMPOSITE MATERIALS**, McGraw-Hill, New York, NY, 1975, pg. 91.

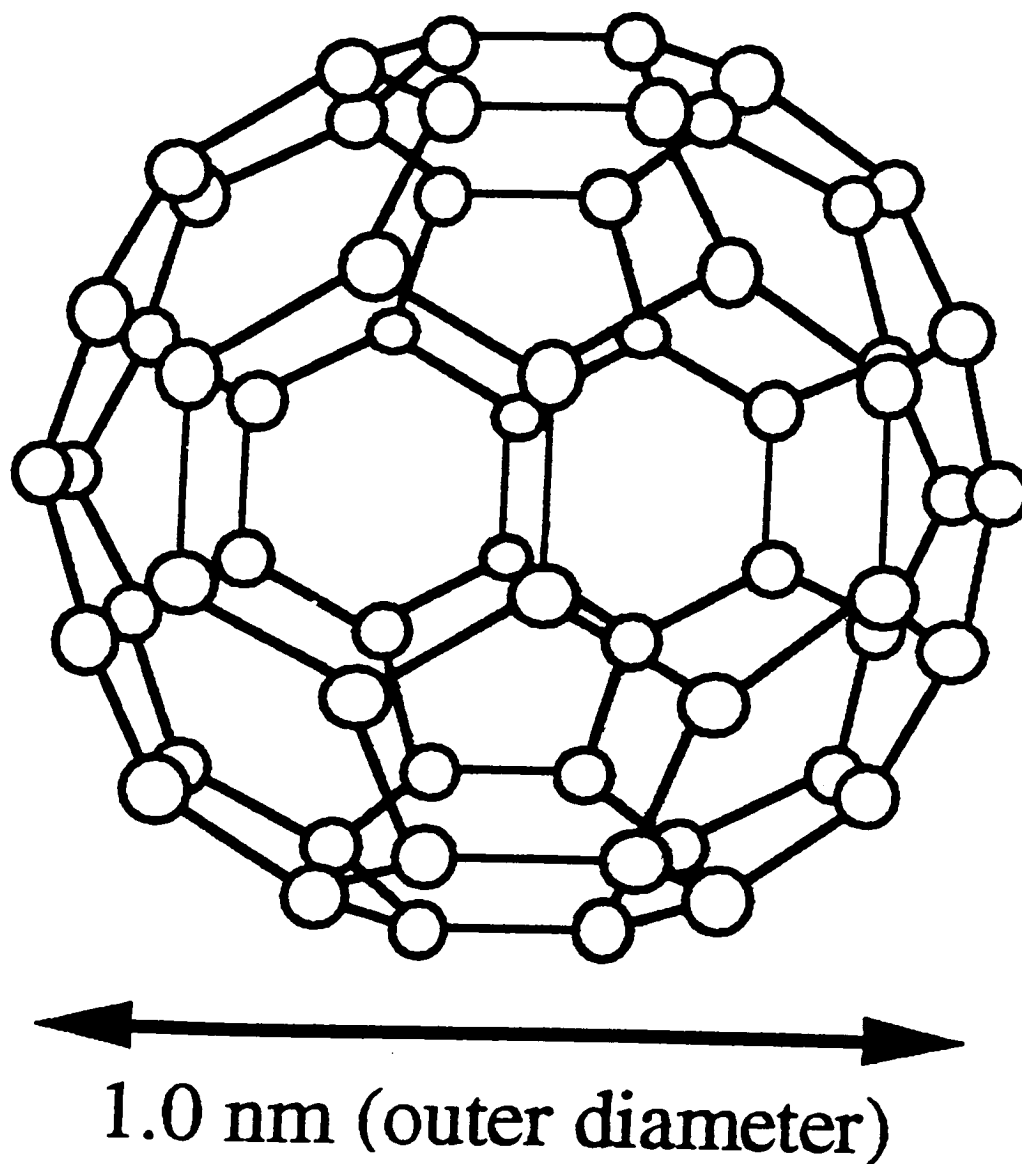


Figure 1. Truncated icosahedral structure of a C₆₀ molecule



a) b) c)

Figure 2. Micrographs of a) 8.0 v/o, b) 4.0 v/o, c) 1.3 v/o fullerene samples at 500 X magnification

Al and C WDS Line Scans of a Sample Processed in the Solid State

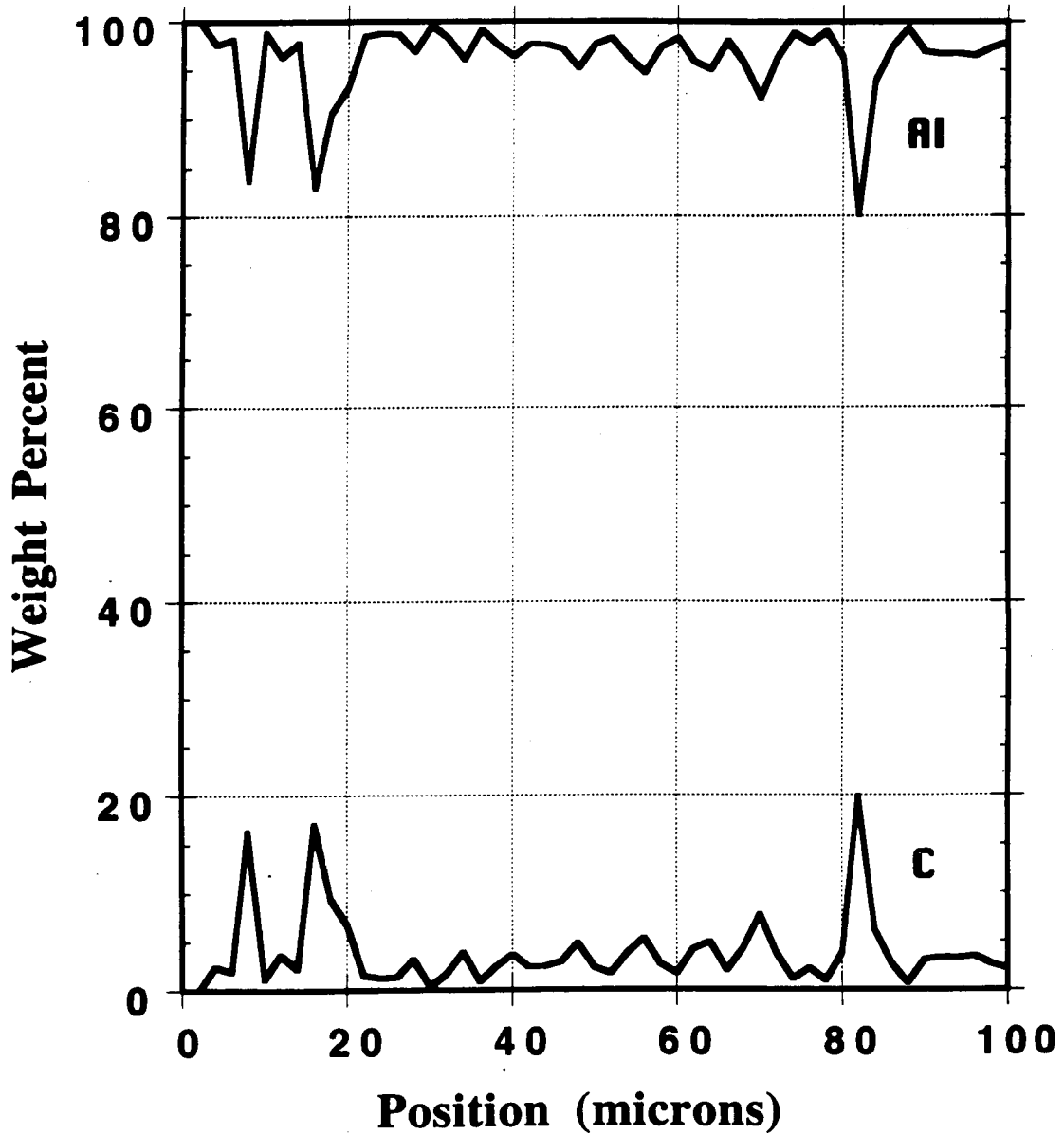


Figure 3. WDS line scans of aluminum and carbon across the grain of a 4.0 v/o fullerene sample

Spectra Comparison of Carbon Forms

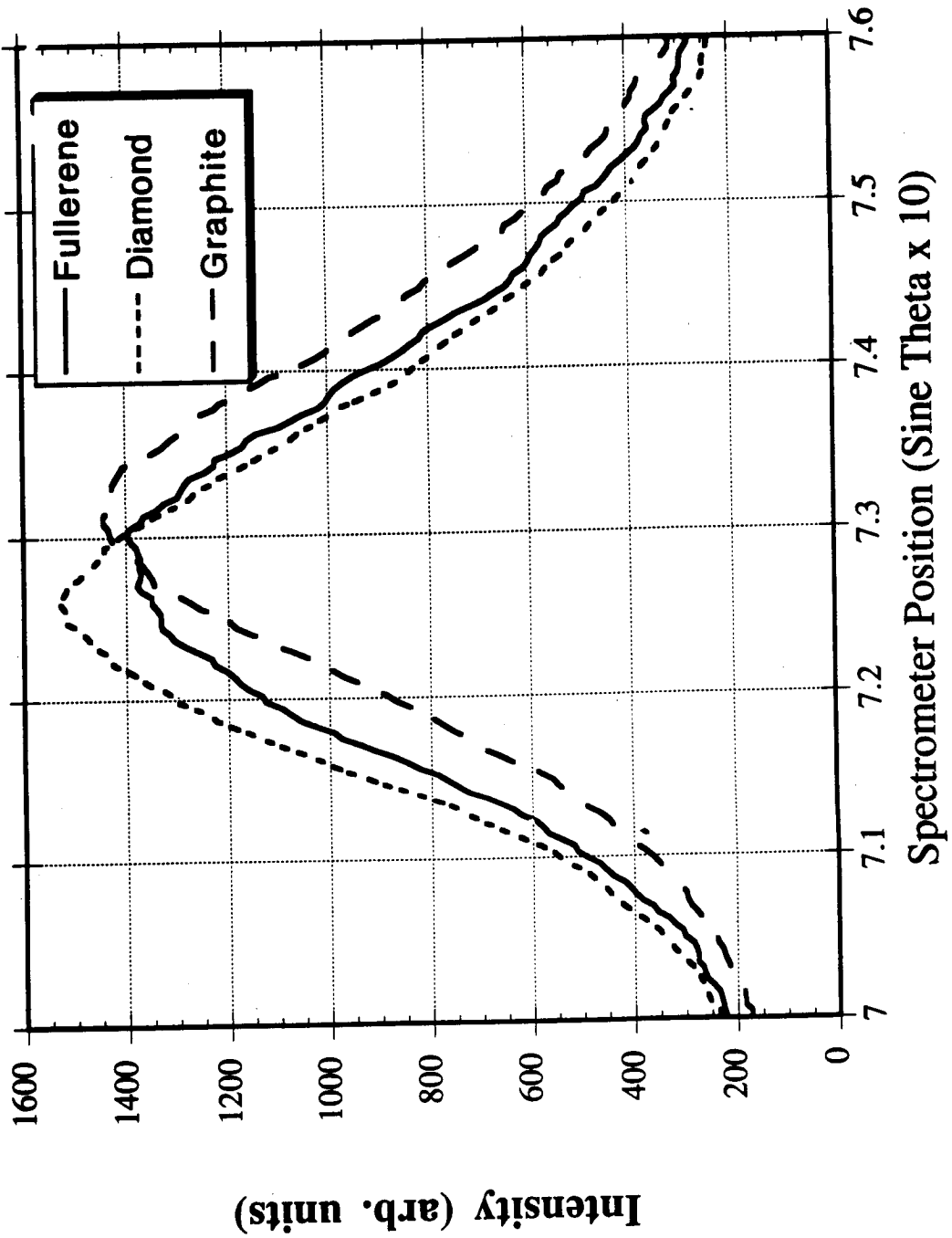


Figure 4. WDS spectra for the three forms of carbon

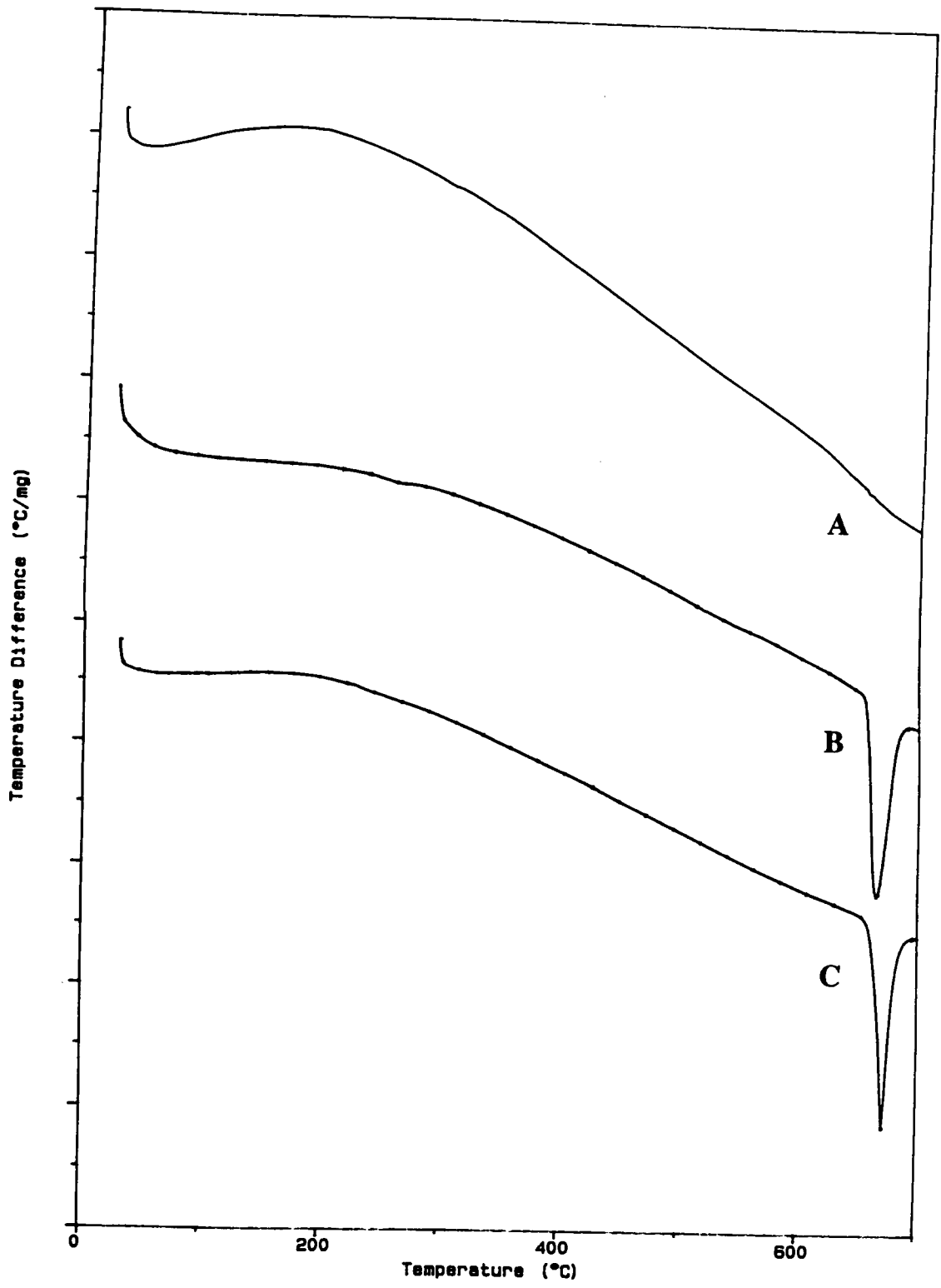


Figure 5. Differential thermal analysis of A) fullerene powder, B) pure aluminum, and C) an 8.0 v/o fullerene sample

Microhardness Profile

Sample 9.5AlF2.5-48/500-24/600-E3

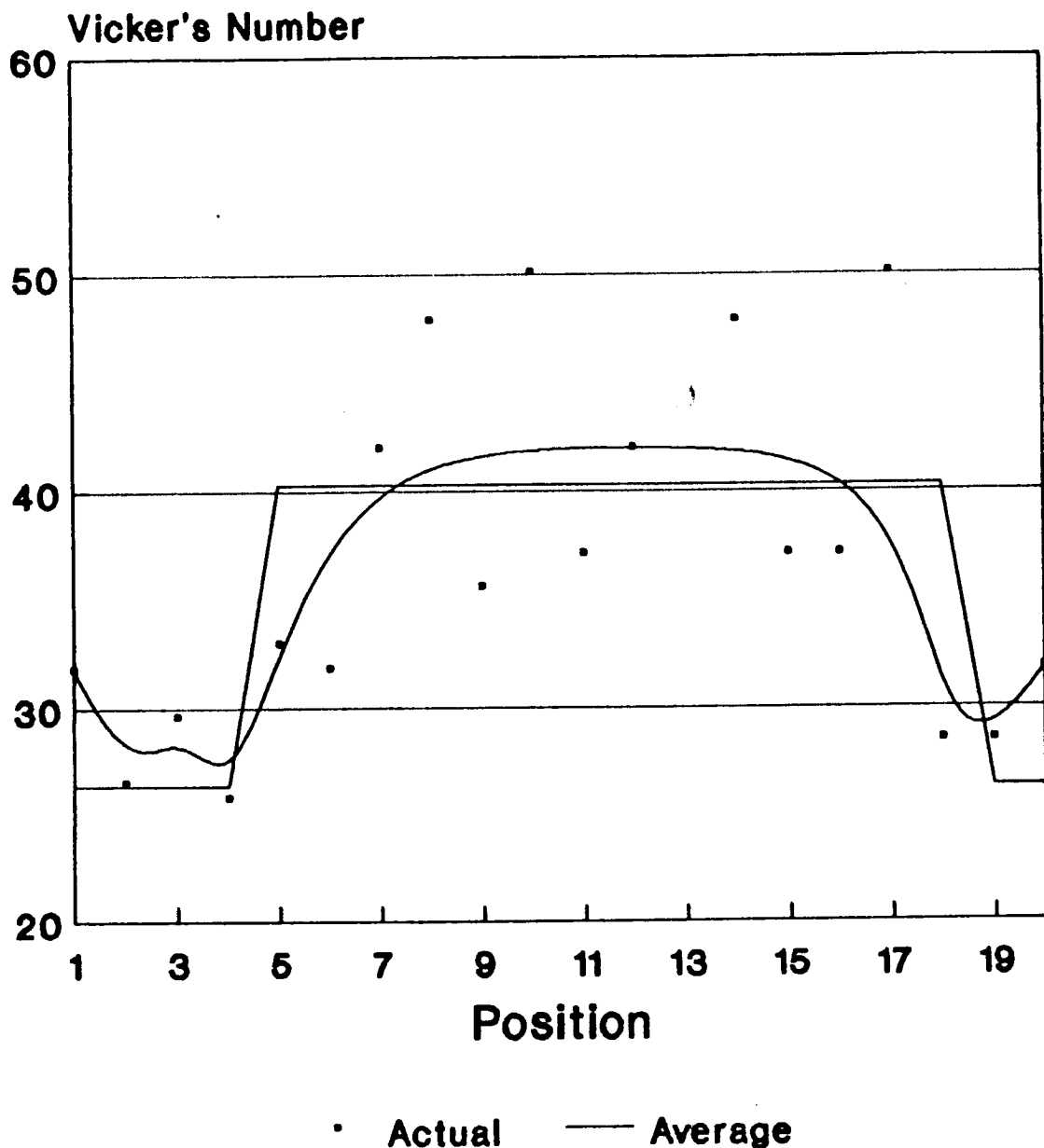


Figure 6. Vickers's microhardness profile across a 4.0 v/o fullerene sample. The actual line is a smoothed spline fit of the data points. The average line is actual average hardnesses for the two regions taken at many random points in each region

01218

Session M2: MATERIALS AND STRUCTURES II

Session Chair: Lubert Leger

ILMENITE AS A DUAL-USE MATERIAL

535-24
44523

A. A. Kumar[†], R. K. Pandey^{††}, T. N. Fogarty[†] & R. Wilkins[†]

([†])Prairie View A&M University, Center for Materials, Microdesign and Microfabrication & Department of Electrical Engineering, Prairie View, TX 77446-0397

(^{††})Texas A&M University, Center for Electronic Materials & Department of Electrical Engineering, College Station, TX 77843-3128

ABSTRACT

This paper addresses the subject of dual-use space technology transfer of a novel, non-traditional material termed ilmenite, found in a large percentage in the moon rocks brought back by NASA's APOLLO missions. The paper is somewhat premature in the sense that though the material as a mineral has been known for a long time, very little is known about pure single crystal ilmenite and hence few applications have been demonstrated. Yet, in another sense, it is very timely due to the fact that ilmenite promises to be a very interesting competition to silicon, silicon carbide and other compound semiconductors, especially those that are employed in high power, high temperature and large data storage/retrieval applications. It seems to be an excellent example of a small investment-high return situation. While some of the applications of this material - for production of oxygen, for instance - have been well-known, electronic applications have received relatively little attention. One reason for this was the fact that growth of single crystal ilmenite requires precise process conditions and parameters. We believe for the first time these have been determined in the Center for Electronic Materials, Texas A&M University. The work being done at Texas A&M University and Prairie View A&M University (supported by Battelle Pacific Northwest Laboratories and the Center for Space Power) indicates the excellent potential this material has in space as well as in terrestrial applications. To mention a few: • as a wide band gap semiconductor it has applications in high temperature, high power situations, especially when heat dissipation is a problem such as may occur in the Space Station; • the possibility of this material radiating in the blue region, it has immense applications in optoelectronics; • as a material with a high density of highly directional *d*-bands, it lends itself to novel processing conditions and perhaps even to "tunability" of physical parameters; • as a potential scintillating material, it has possible applications as a sensor in waste management; • as an oxygen sensor it has possible applications in automotive electronics; and • as a radiation resistant material, it has obvious applications in the space environment. Results - experimental and theoretical - obtained so far in our laboratories will be reported with particular emphasis on the transfer of technology involving this fascinating material.

I. Introduction

The present day state-of-the-art technology is such that new demands create the need for discovering new electronic materials with increasingly improved properties. Of particular interest are those materials that allow "dual-use" - space and commercial applications. NASA has been at the forefront of materials research, applications and technology transfer. NASA's research spans practically all aspects of materials, from processing and growth to testing and characterization to technology transfer. Some of this research has spun off from past space programs. For instance, NASA's involvement in photovoltaic energy conversion now spans more than three and half decades. This paper deals with a material - ilmenite - that has been found in large percentages in the moon rocks brought back by the APOLLO missions. Ilmenite has been identified as an attractive material for a large number of novel applications (1-7). In particular, the photovoltaic, optoelectronic, and the high radiation resistant characteristics of ilmenite seem to offer superior performance than the traditional semiconductors. For instance, silicon solar cell has been the mainstay power conversion source for the U.S. space program. However, the conversion efficiency of Si seems to be bound at an upper limit of ~ 30%. Furthermore, Si is difficult to use in optoelectronic applications due to its indirect bandgap. A competing material, SiC, has attracted significant attention owing to its large bandgap[8]. However, this bandgap is indirect in nature. Also, the processing of SiC is problematic; the single crystal has been grown with great difficulty. Ilmenite could serve as an alternate material that could replace semiconductors like Si, SiC and GaAs in energy conversion (1-4). Another application is in photoelectric cells that can be used for cleaning water and recovering gases like oxygen and hydrogen especially on the moon (7). Because of its large band gap (~ 2.58 eV) it can be used to develop optoelectronic devices like LEDs and solid state lasers in the visible spectral range. Ilmenite can be grown both an p-type and an n-type semiconductor making bipolar junction devices consisting of alternate n- and p-type materials (5-7) possible. These potential applications and the possibility that ilmenite might be "bandgap engineered" in novel ways, makes ilmenite a natural material for extensive study. The work carried out at Prairie View A&M University and Texas A&M University so far confirms this prediction.

II. Brief Overview of Large Band Gap Semiconductors

There has been a tremendous amount of interest in large band gap semiconductor materials. Well-known examples of these are - silicon carbide (SiC), GaAs heterostructures, diamond, and boron nitride (B₃N) [9]. While each of these materials has its own advantages, there is a growing tendency for exploring new and non-traditional semiconductor materials [9]. This paper focuses on one such material - **ilmenite** - from the family of iron titanates.

The principal driving force behind research in large band gap semiconductors is their potential use in high power, high frequency, high temperature microelectronic and optoelectronic (including photovoltaic) devices which are also resistant to radiation damage. Over the past two decades significant advances have been made in the growth, processing and doping of these materials. As a result this class of materials now holds a significant promise for semiconductor electronics in many areas of application. Typical

properties include: increase in electrical conductivity by two or three orders of magnitude relative to comparable silicon devices; a substantial improvement in the on-state resistance of these devices; larger band gap relative to silicon provides considerable reduction in off-state power loss; and enhanced maximum frequency of operation.

Heat is one of the most deleterious environments for the operation of semiconductor devices. The electrically generated heat load from silicon devices will continue to increase with increased integration. Such high temperature environments occur in: avionics (need for reduction of fuel consumed in cooling commercial and military aircraft), space (reduce launch weight of an unmanned satellite), automotive (under the hood control of ignition of fuel to air mixture), well-logging applications, and high temperature sensors and integrated circuits.

III. MAJOR APPLICATIONS "Dual-Use" Applications of Ilmenite

Ilmenite has a very high potential in offering dual-uses. Table I gives an overview:

Table I- Selected dual-use applications of ilmenite and its competitors

Factor of Interest	Most Relevant Characteristic	Ilmenite	SiC	GaAs	Si
Energy Conversion	High conversion efficiency; higher "light" to "dark" current ratio	Yes	Yes	Yes	No
Optoelectronics					
Solid State Laser	Direct bandgap	Yes*	No	Yes	No
Increased Data Storage	Smaller wavelength	Yes	Yes	No	No
Radiation Resistance	High carrier mobility	Yes	Yes	Yes	No
	Annealing ability	Yes	Not known	Yes	No
Gas sensors	Stability Sensitivity Selectivity	TBD High TBD	No High Not Known	Yes Low	No Low
Processing	Cost, complexity of apparatus,	Small	Large	Large	Large
High Temperature Operation	Oxidation Resistance	Yes	No	Yes	No
Waste Detection	Scintillation	Yes*	Not known	Not known	Not known
Material Modification - Heterostructures	"Seamless" interfaces	Possible	No	Yes	No

*Indication to this effect, more work required.

A. Space Applications Aspects

1. Weight Savings: Extensive cooling equipment is required for Si based electronic (maximum operating temperature 125°C) on spacecraft. Much of this equipment could be eliminated using a large band gap material such as ilmenite.

2. Photovoltaic Devices: There is a need for different types of materials for photovoltaic use in space and on earth. This is due to a difference in spectral intensities between terrestrial and space solar insolation. There is significantly higher amount of energy in the spectrum at shorter wavelengths in space. Photovoltaic conversion at short wavelengths has been particularly problematic owing to the fact that short wavelength light is absorbed very near the top surface of the cell. However the surfaces of silicon and gallium arsenide are replete with dangling bonds and interface states which can trap free electrons and holes, leading to a degradation in conversion efficiency. Ilmenite, owing to its bandgap being in the 2.5 eV range, will have a higher response to this spectral range and hence may be a more suitable material for a space solar cell.

3. Radiation Resistance: Solar cell radiation damage is the major life-limiting factor for an array, especially in a geosynchronous orbit. A typical silicon solar array may be oversized by as much as 40% to assure sufficient power over the mission [10] The large band gap of ilmenite makes it naturally radiation resistant.

B. Commercialization Aspects

While the research on ilmenite is still in an early stage to determine its eventual commercializability, the benefits it seems to offer make it worthwhile for substantial investment in its research and study. The following is a brief discussion of the type of applications anticipated and the possible advantages of ilmenite over the standard semiconductors.

1. Sensors: Semiconductor sensors have stability, low cost, and simple signal output [11]. The drawbacks of existing oxide sensors are mainly reproducibility and stability. Currently employed compressed powders are not noted for their predictability. Slight changes in particle size, intergranular contact area, and adsorbed impurities produce variations in the signal output. Also the bulk and surface properties of some oxides can change with time at elevated temperatures. For automotive applications, there is at present no accepted semiconducting oxide sensor for oxidants, except (perhaps) for TiO₂, for determining the air/fuel ratio in auto exhausts. Ilmenite should possess a high degree of stability and a high sensitivity (proportional to the bandgap) useful for sensor applications.

2. Heterostructures of Ilmenite: The interesting annealing properties of ilmenite (which effect the oxidation state of the iron) may provide the means for creating "seamless" interfaces between areas of n- and p-type material, perhaps leading to tunable bandgaps.

3. Optoelectronic Devices: Solid state blue-green lasers have immense applications in optical information storage industry [12]. Optical data storage systems are fundamentally limited by the wavelength of the laser light used in read and write operations. The data transfer rate in an optical recording system operating at a fixed rotational speed is inversely proportional to the laser wavelength and the capacity of a fixed-diameter optical disc is inversely proportional to its square [12]. Thus, a 430 nm based optical drive would increase the information storage on a disc by almost a factor of four when compared with a 780 nm based drive. The potential market for both entertainment and computer applications is estimated to be in the hundreds of millions of dollars [12]. Ilmenite, due to its possible direct bandgap being in the 2.5 eV range would be an ideal material for such a laser.

4. Thermoelectric Coolers: Currently, freon is a major component of cooling system, and is known to an environmental hazard. Substitutes for cooling system and cooling materials are needed. Initial measurements of Seebeck coefficients appear to make ilmenite attractive in refrigeration and cooling systems, especially when compact solid-state coolers are required for automotive and microelectronic applications.

IV. Basic Properties of Ilmenite

Table II- Comparison Between Ilmenite, Si, SiC and GaAs

Physical Property	Ilmenite FeTiO ₃	Si	SiC	GaAs
Bandgap	2.58 eV direct(?)	1.12 eV indirect	2.9 eV indirect	1.43 eV direct
Crystal Structure	Corundum (Al ₂ O ₃)			
Unit Cell	Hexagonal	Cubic	Hexagonal	Cubic
a	5.09 Å	5.43 Å	3.08 Å	5.65 Å
b		5.43 Å		5.65 Å
c	14.09 Å	5.43 Å	15.1 Å	5.65 Å
c/a	2.77	-	4.90	-
Resistivity	1.45 Ωm	2.3 * 10 ⁵ Ωcm	10.32 Ωcm	0.4 Ωcm
Hall coefficient	0.26 * 10 ⁻⁵ m ³ /C	432.4 m ³ /C	30*10 ⁻⁵ m ³ /C	.034 m ³ /C
Carrier Concentration	2.6 * 10 ²⁴ m ⁻³	1.445 *10 ¹⁶ m ⁻³	2.08 * 10 ²² m ⁻³	1.838 * 10 ²⁰ m ⁻³
Density	4.83 gm/cm ³	2.32 g/cm ³	3.21 g/cm ³	5.30 g/cm ³
Antiferromagnetic	Neel point = 55K	-	-	-
Ferrimagnetic	in solution with α-Fe ₂ O ₃	-	-	-
p-type semiconductor	pure ilmenite	on doping	on doping	on doping
n-type semiconductor	in solution with α-Fe ₂ O ₃	on doping	on doping	on doping
Melting point	1410°C	1685°K	3103 °K	1511 °K

Ilmenite (FeTiO_3) is a member of the iron titanate family, which also includes ulvospinel (Fe_2TiO_4) and pseudo-brookite (Fe_2TiO_5). Ilmenite belongs to the $\text{R}_2\text{O}_3(8)$ hematite-ilmenite group. Its lattice is rhombohedral, $a_R = 5.52\text{A}$, $a_H = 54.81\text{A}$ (hexagonal, rhombohedrally-centered, with $a_H = 5.087\text{A}$, $c_H = 14.042\text{A}$). Space group: $\text{R}\bar{3}$. It is isostructural with geikelite, MgTiO_3 , $c_H = 13.898$, $a_H = 5.054\text{A}$; pyrophanite MnTiO_3 , $c(H)=14.290\text{A}$, $a_H = 5.1396\text{A}$. Ilmenite is grayish-white with a brownish tinge. Picot and Johan (9) state that the anisotropy colors are greenish-gray and this agrees with the calculated color which is greenish-blue-gray. In Table II we compare some of the properties of ilmenite to more common semiconductors.

RESEARCH ACCOMPLISHMENTS AND GOALS

Ilmenite has been grown in single crystalline form at the Center for Electronic Materials, Texas A&M University (TAMU) by Dr. Pandey and his Associates in collaboration with Dr. Kumar at Prairie View A&M University (PVAMU). One of the main problems in the growth of good quality single crystals of pure ilmenite is the fact that both iron and titanium can assume different states of valency during their synthesis at high temperatures. The solid state reaction follows the chemical reaction



Details of growth processes will be published elsewhere. In addition, basic electronic and magnetic properties have been measured. Some optical characterization has been carried out with colleagues at Sam Houston State University. The results of some of these measurements are in Table II.

Future plans call for exploration of growth and annealing parameters and complete characterization of the materials. This work will be shared between TAMU and PVAMU. The processing research will continue at TAMU, while modeling and characterization will be shared by both institutions. This will take advantage of surface science facilities at PVAMU which include scanning electron microscopy, an x-ray photoelectron spectrometer, and a newly acquired ultra-high vacuum scanning tunneling microscope.

CONCLUSIONS

The intent of this paper has been to provide an overview of research on ilmenite performed at PVAMU and TAMU with specific focus on its dual uses in space and terrestrial applications. This is an excellent example of a spin-off of the space technology. A similar example in this field occurred with respect to silicon solar cells, in particular with respect to their improved geometry. The U.S. space program, in particular NASA's achievements have demonstrated an excellent synergism between space and terrestrial research.

Ilmenite and other members of the iron titanate family appear to be attractive candidates for research and development leading to some important electronic technologies. The initial work on synthesizing and characterizing ilmenite for its

536-25
1/4526

**THE MATERIALS CHEMISTRY OF ATOMIC OXYGEN
WITH APPLICATIONS TO ANISOTROPIC ETCHING OF SUBMICRON STRUCTURES IN
MICROELECTRONICS AND
THE SURFACE CHEMISTRY ENGINEERING OF POROUS SOLIDS.**

Steve L. Koontz, Lubert J. Leger, Corina Wu, JSC/ES5, Houston, Texas
Jon B. Cross, Los Alamos National Laboratory, Los Alamos, New Mexico
Charles W. Jurgensen, AT&T Bell Laboratories, Murray Hill, New Jersey

ABSTRACT

Neutral atomic oxygen is the most abundant component of the ionospheric plasma in the low Earth orbit environment (LEO; 200 to 700 kilometers altitude) and can produce significant degradation of some spacecraft materials. In order to produce a more complete understanding of the materials chemistry of atomic oxygen, the chemistry and physics of O-atom interactions with materials were determined in three radically different environments: (1) The Space Shuttle cargo bay in low Earth orbit (the EOIM-III space flight experiment), (2) a high-velocity neutral atom beam system (HVAB) at Los Alamos National Laboratory (LANL), and (3) a microwave-plasma flowing-discharge system at JSC. The Space Shuttle and the high velocity atom beam systems produce atom-surface collision energies ranging from 0.1 to 7 eV (hyperthermal atoms) under high-vacuum conditions, while the flowing discharge system produces a 0.065 eV surface collision energy at a total pressure of 2 Torr. Data obtained in the three different O-atom environments referred to above show that the rate of O-atom reaction with polymeric materials is strongly dependent on atom kinetic energy, obeying a reactive scattering law which suggests that atom kinetic energy is directly available for overcoming activation barriers in the reaction. General relationships between polymer reactivity with O atoms and polymer composition and molecular structure have been determined. In addition, vacuum ultraviolet photochemical effects have been shown to dominate the reaction of O atoms with fluorocarbon polymers. Finally, studies of the materials chemistry of O atoms have produced results which may be of interest to technologists outside the aerospace industry. Atomic oxygen "spin-off" or "dual use" technologies in the areas of anisotropic etching in microelectronic materials and device processing, as well as surface chemistry engineering of porous solid materials are described.

INTRODUCTION

Oxygen atoms are the most abundant neutral constituents of the Earth's ionosphere at altitudes ranging from 200 to 700 km¹⁻³ (low Earth orbit or LEO) and have been shown to be one of the more important environmental factors involved in the degradation of several important classes of spacecraft materials.⁴⁻⁶ Atomic oxygen is expected to display a higher reactivity than molecular oxygen under most circumstances. In addition, the high velocity of orbiting spacecraft (much greater than gas kinetic speeds in the

ionosphere) produces a high flux of atoms, molecules and ions with a collision energy on the order of 5 eV with the forward facing spacecraft surfaces. The result is a unique chemical and physical environment in which the temperature of the reactive surface is not coupled directly to the surface collision energy of the gaseous species striking that surface. In this paper, we present an overview of the Lyndon B. Johnson Space Center (JSC) investigations into the materials chemistry of atomic oxygen.

Our approach to achieving our atomic oxygen materials chemistry objectives was based on measurements of sample properties before and after exposure to known O-atom fluences in three well-characterized environments: (1) The LEO environment during a Space Shuttle mission^{7,8} (2) the high velocity neutral-atom beam (HVAB) system at the Los Alamos National Laboratory (LANL)^{9,10} and (3) a flowing discharge (downstream plasma) system at JSC.^{11,12}

Observation of the changes in various materials resulting from exposure to oxygen atoms in the LEO environment suggested that directed, high-translational-energy oxygen atoms might effect the anisotropic etching of the photoresists used to pattern microcircuits. In addition, observation of the changes in the surface chemistry of porous solids exposed to atomic oxygen in LEO suggested that thermal energy oxygen atoms might be used to modify the surface chemistry of porous solids in a regiospecific way to produce distributed pore chemistry solids.

MATERIALS AND METHODS

Atomic oxygen materials chemistry studies were conducted in three different experimental systems: (1) an atomic-oxygen materials-chemistry Space Shuttle experiment (Evaluation of Oxygen Interactions with Materials III or EOIM-III)^{7,8} (2) a high velocity O-atom nozzle beam system^{9,10}, and (3) a flowing plasma discharge (remote plasma) apparatus.^{11,12} The extent of O-atom reaction with the polymer film or sheet specimens was determined by mass loss measurements and/or surface recession measured by profilometry. O-atom reactivity measurements made by weight loss and profilometry were compared, wherever possible, as a check on the accuracy of the two measurement methods. In most cases, the polymer film or sheet samples were used, as received, from various commercial suppliers (see table 1). Prior to atomic oxygen exposures, polymer film and sheet samples

semiconducting properties has been completed, but much work remains. For example, its structure and the nature of the oxidation states of the iron must be understood before modifications can be made in the system to suit the needs of technology. Another important area of research is the annealing of the material to stabilize its structure and enhance atomic ordering. Preliminary results show that the physical properties are strongly dependent on annealing conditions.

REFERENCES

1. Knoll, G., "Detection and Measurement", Radiation John Wiley and Sons, New York, 1979.
2. Ginley, D., and R. Baughm, Mat. Res. Bull. vol.12, 1803, 1957.
3. Pulfrey, D., "Photovoltaic Power Generation", Van Nostrand Reinhold Co., New York, 1978.
4. Johnson, W., "Solar Voltaic Cells", Marcel Dekker Inc., New York, 1980.
5. Ishakawa, Y., and S. Akimoto, J. Phys. Soc. Japan, vol 12, 1803, 1957.
6. Ishakawa, Y., J. Phys. Soc. Japan, vol 13, 37, 1958.
7. Senfelte, F.E., et al., "Antiferromagnetic (Neel) Transitions in Lunar Glass and Ilmenite," Fifth Lunar Science Conference, 1974.
8. Trew, Robert J., Jing-Bang Yan, Philip M. Mock, Proc. IEEE, vol 79, no. 5, 598, 1991.
9. Singh, Jasprit , "Physics of Semiconductors and Their Heterostructures" McGraw-Hill, Inc., New York, 1993, pgs. 836-842.
10. Flood, Dennis J., "Space Solor Cell Research", Chem. Engin. Prog., 85 (4), 62, 1989.
11. Morrison, S. Roy, "Semiconducting-Oxide Chemical Sensors", IEEE Aerospace Electronic Systems, March 1991, pg 32.
12. Dixon, G. J., "Compact Blue-Green Lasers Get Down To Business", IEEE Circuits and Devices, November 1993, pg. 18.
13. Picot, P and Z, Johan "Atlas of Ore Minerals" Co-published by Elsevier Publishing Co and BRGM. (1982).
14. Peckett, A., "The Colors of Opaque Minerals" Van Nostrand Reinhold, New York, 1992.
15. B.L. Gries, "Single Crystal growth and Characterization of Ilmenite for Electronic Applications," M.S.Thesis, Texas A & M University, 1988.

TABLE 1 POLYMER REACTION EFFICIENCIES

POLYMER	Re (EOIM-III) x 10 ²⁴	Re (STS-8) x 10 ²⁴	Re (STS-41) x 10 ²⁴	Re LDEF x 10 ²⁴
KAPTON (LeRC R.R.)	3.1	3.0	3.3	3.0
EYMYD-F (ETHYL CORP.)	2.7			
CR-39 POLYCARB.	6.1	6.0		
PEEK (ICI)	3.4		4.3	
XYDAR (AMOCO)	2.9			
LCP-4100 (DuPONT)	3.2			
MYLAR A (DuPONT)	3.8	3.9		
PE, (LeRC R.R.)	4.4	3.7		
HDPE (PHILLIPS, EMH6606)	3.7	3.7	3.5	
POLYMETHYLPENTENE (MITSUI)	5.3			
POLYPROPYLENE	5.5		4.4	
TEDLAR (DuPONT)	3.5	3.2		
TEFZEL (CLEAR, DuPONT)	0.9		0.2	
TEFZEL (BLUE, RAYCHEM)	1.1			
TEFZEL (WHITE, RAYCHEM)	0.9			
KYNAR (PENWALT)	1.2			
KEL-F (PCTFE)	0.9			
HALAR (ALLIED)	1.9			
ACLAR 33C ALLIED)	1.0			
FEP TEFLON (LeRC R.R.)	0.05	<0.03	<0.03	0.3
TFE TEFLON	0.06	<0.03		0.5
EYPEL-F, (ETHYL CORP.)				
POLY(BISTRIFLUOROPR O-PYLPHOSPHAZENE)	<0.03			

were subjected to 48 hours of vacuum baking, at the maximum temperature expected during the exposure, to remove absorbed water and other volatile substances prior to pre exposure weighing. After weighing, the polymer samples were cleaned by brief rinsing with Q Cleanant Solvent (Thermo Analytical Inc., Monrovia Calif.) to remove surface contaminants, such as silicones, which could interfere with O-atom reactions. Polymers were tested for compatibility with the cleaning solvent before use. Polymer surfaces were analyzed before and after exposure by x-ray photoelectron spectroscopy (XPS) both to evaluate the extent of surface oxidation and to verify surface cleanliness.

The HVAB at LANL has been described.^{9,10} Briefly, a laser-sustained gaseous discharge (50 percent O₂/Ar or O₂/Ne at pressures on the order of 2000 Torr) undergoes supersonic nozzle expansion to form a seeded beam of oxygen atoms (O^{3P} electronic ground state) and inert gas atoms. Adjusting the location of the gaseous discharge relative to the expansion nozzle permits control of the kinetic energy distribution function in the final beam so that the average kinetic energy can be varied between 0.4 and 3 eV. Beam velocity distribution functions were determined directly using well-known velocity selector, i.e., time of flight (TOF), methods^{9,10} with mass spectrometric detection of atomic and molecular species. A schematic diagram of the HVAB system is shown in figure 1.

The O-atom flux in the beam was determined using both relative abundance from the TOF data and the pressure

rise in an accommodation chamber as was previously reported.¹⁰ Accommodation chamber pressure was measured using both a spinning rotor gauge (total pressure) and a residual gas analyzer (partial pressures at steady state) that was calibrated (for stable gas species such as Ar, Ne, O₂) using the same spinning rotor gauge. The pressure rise in an accommodation chamber was measured with the high velocity atom beam entering the chamber and the chamber turbo pump gated off. Under these conditions, the inert-gas effusive flux out of the chamber equals the inert gas beam flux into the chamber so that the inert gas flux in the high velocity atom beam can be immediately calculated. Given the inert-gas flux in the beam and the relative abundances, from phase-sensitive mass spectrometry, the fluxes of all other beam species, including atomic oxygen, were calculated using known values of electron impact cross sections and the mass spectrometric transmission function.

For O-atom reactivity studies, polymer film specimens were mounted on a heat sink to control heating by thermal radiation from the HVAB source. The typical Kapton film temperature during O-atom bombardment was 45 degrees Centigrade, and the samples could be heated to determine Arrhenius activation energies. Gaseous reaction products were detected using modulated beam techniques with phase-sensitive mass spectrometric detection to reject vacuum system background.^{9,10}

The atomic oxygen materials chemistry experiment (EOIM-III) was conducted during STS-46 at an altitude of

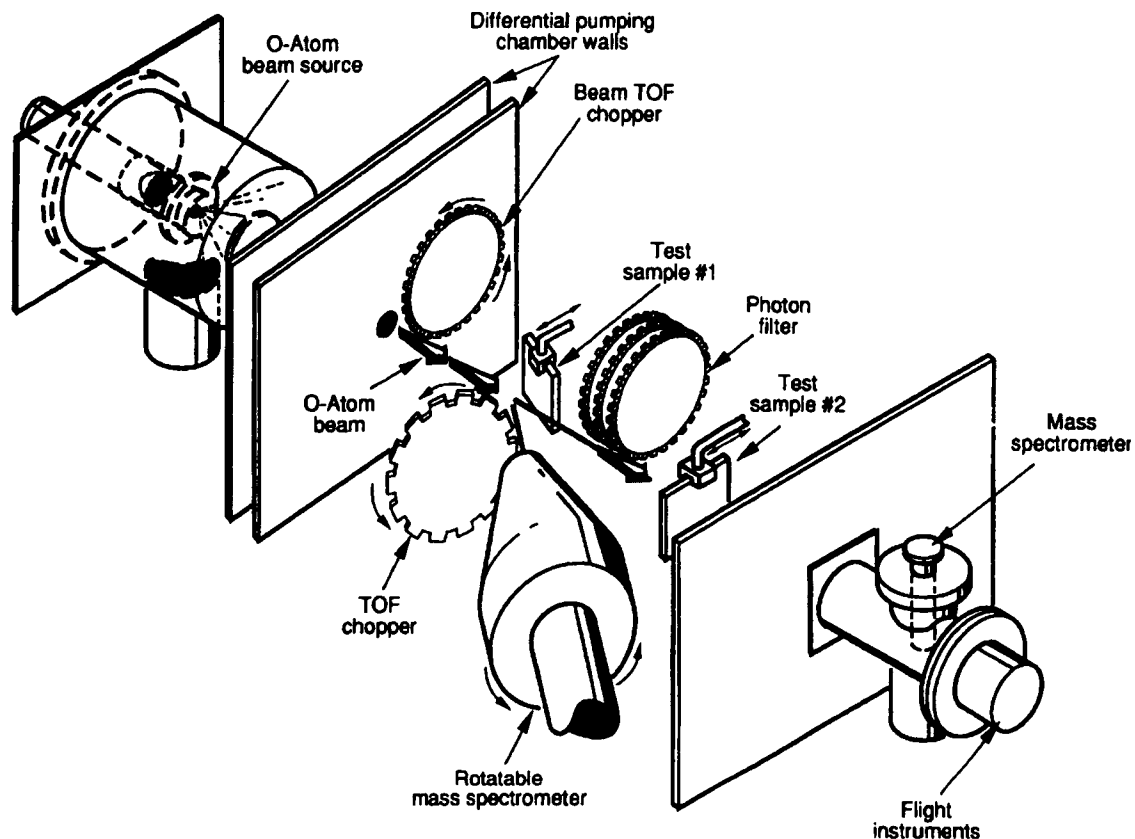
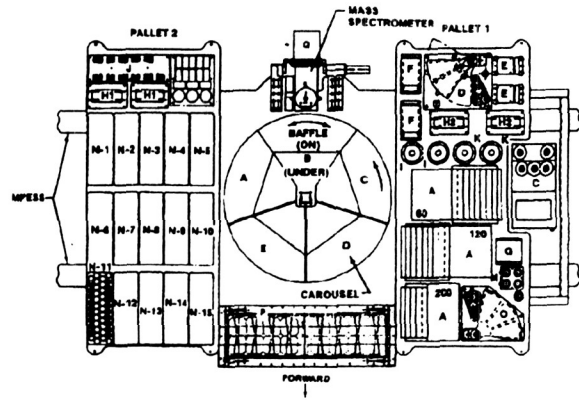
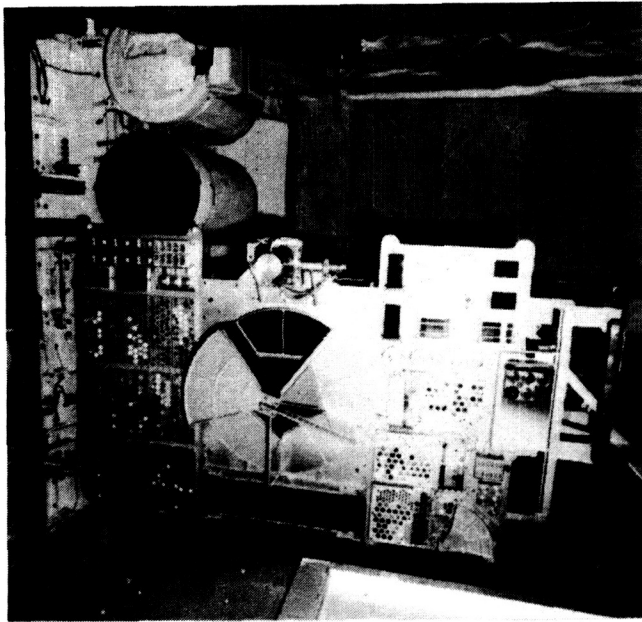


Figure 1. A schematic diagram of the high velocity neutral O-atom beam (HVAB) at Los Alamos National Laboratory. The HVAB was developed and operated under contract to JSC.

223 kilometers.^{7,8} The total O-atom fluence was $2.3 \pm 0.3 \times 10^{20}$ atoms/cm² as determined by calculation using a well established mathematical model of the thermosphere⁸ and by direct measurement using a mass spectrometer calibrated for O-atom measurements before and after flight in the HVAB referred to above, as well as Kapton film dosimeters also calibrated for O-atom fluence measurements in the HVAB. The materials samples and related instruments were oriented so that the angle between the surface plane and the incident atom flux was 90 degrees (normal incidence). In addition to the passive sample carriers, three sample trays were thermostated at 60, 120 and 200 degrees Centigrade to permit determination of Arrhenius activation energy parameters of the atomic oxygen reaction with various materials. The average temperature of the passive sample carriers during the EOIM-III O-atom exposure period was 19 degrees Centigrade. A schematic diagram of the EOIM-III payload is shown in figure 2. A key component of the EOIM-III flight experiment package was the mass spectrometer shown in the payload schematic. The mass spectrometer was calibrated for high velocity O-atom flux measurements both before and after STS-46 in the HVAB at LANL.^{7,8,10} The MSIS-86 model of the thermosphere³ was used to calculate O-atom fluence for the EOIM-III exposure using as-flown orbit data for STS-46 and daily

average values of the solar geomagnetic parameters. Daily average values were used instead of monthly averages because significant magnetic sub-storm activity was observed during the STS-46 mission.^{7,8}

The methods and apparatus used to determine the O-atom reactivities of polymers in the flowing discharge (remote plasma) apparatus have been described.¹¹⁻¹³ Briefly, a working gas (10 percent O₂/Ar or 10 percent N₂O/Ar), at total pressures on the order of 2 Torr, was passed through a 2.45 GHz Evenson discharge cell and flowed downstream of the discharge before coming into contact with the polymer samples so that the gas had cooled to room temperature but still contained O-atoms in the O^{3P} electronic ground state. The O-atom concentration was determined by chemiluminescent titration using NO₂.^{14,15} and the atom flux on sample surfaces was determined using well-known methods for modeling flowing reaction-diffusion systems.¹³ The flowing discharge approach was selected because a substantial database on the gas phase chemical kinetics of atomic oxygen has been produced in flowing discharge type equipment in the past 30 years^{14,15} so that the O-atom environment is well characterized. Both samples and reactive gas could be heated to determine Arrhenius activation energies. In addition, VUV discharge lamps could be installed to provide known



- ATOMIC INTERACTION EXPERIMENTS:
- A - HEATED PLATE (HPC), 2 EA
 - B - ATOM SCATTERING EXPERIMENT (ASE), 1 EA
 - C - ENVIRONMENT MONITOR PACKAGE (EMPC), 1 EA
 - D - SOLAR UV EXPERIMENT (SUE), 1 EA
 - E - STATIC STRESS FIXTURE (SSFC), 2 EA
 - F - UNIFORM STRESS FIXTURE (USFC), 2 EA
 - G - ATOMIC OXYGEN MONITOR (AOM), 1 EA
 - H1 - COMPOSITE STRESS FIXTURE (LSC), 2 EA
 - H2 - COMPOSITE STRESS FIXTURE (LSC), 2 EA
 - I - SCATTEROMETER (LPL), 2 EA
 - J - MECHANICAL STRESS FIXTURE (LSC), 11 EA
 - K - REFLECTOMETER (LSC), 2 EA
 - L - PINHOLE CAMERA (LSC), 1 EA
 - M - SCATTEROMETER/AEROSPACE CORP.), 1 EA
 - N - PASSIVE SAMPLE CARRIERS, 15 EA
 - O - VARIABLE EXPOSURE TRAY, 1 EA
 - P - FREEDOM ARRAY MATERIALS EXPOSURE EXPERIMENT (LSC), 1 EA
 - Q - QUADRUPLE MASS SPECTROMETER, 1 EA

Figure 2. The EOIM-III flight experiment: (a) a photograph of the payload in the Space Shuttle cargo bay prior to launch, and (b) a schematic diagram of the payload components.

doses of VUV radiation simultaneously with O-atom exposure. A schematic diagram of the flowing discharge apparatus is shown in figure 3. The flowing discharge environment is different from both the HVAB and on-orbit environments in three important respects: 1) O-atoms arrive at the polymer sample surface by diffusion from an isotropic gas, not as a well defined directed flux or beam, 2) the pressure is on the order of 2 torr, not 10-5 Torr, so that gaseous reactive intermediates can return to react with the polymer surface, and 3) the flux of molecular oxygen is much higher in the flowing discharge than in the LEO or HVAB environments described above, though large variations in O₂ partial pressure have revealed no effects on the reactivity of Kapton polyimide in this sys-

tem.¹² The dominant process is still O-atom reaction with the polymer surface so that, with some caveats, the flowing discharge system can still provide useful measurements of the reactivity of low kinetic energy (0.065 eV) O-atoms.

RESULTS AND DISCUSSION: POLYMER REACTIVITY

Polymer reactivity with O-atoms is reported as the reaction efficiency, R_e , which is easily calculated by dividing the surface recession in centimeters by the total O-atom fluence in atoms/cm². R_e has the units of cm³/atom. R_e allows a meaningful comparison of the O-atom reactivity of polymers independent of variations in atom dose or

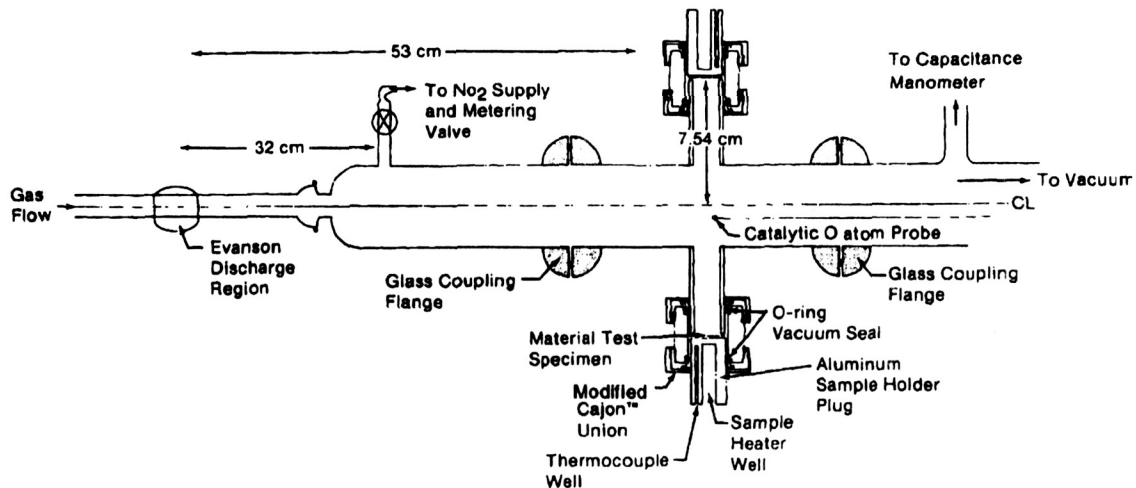


Figure 3. The JSC flowing discharge apparatus.

polymer composition and density as long as the polymer is in the form of a full-density solid. Polymer Re values (cm³ of material removed per incident O atom) determined following exposure on orbit in the EOIM-III passive trays, are shown in table 1, where EOIM-III measurements are also compared with those made following previous flight experiments. The reaction efficiencies reported in table 1 were determined by weight loss only; the repeatability of the measurement is indicated as the difference between the largest and smallest measurement, if more than one sample was exposed on orbit. Comparison of the EOIM-III reaction efficiency data in table 1 with those of previous missions demonstrates that the polymer reaction efficiency data base has been both enlarged and verified.

The reliability of the weight loss measurements was checked by direct comparison with profilometry measurements. The results of replicate measurements of the reaction Re of Kapton polyimide demonstrate excellent agreement between the profilometry and weight loss measurements. Weight loss on four Kapton samples produced a reaction efficiency of $3.05 \pm 0.1 \times 10^{-24}$ cm³/atom while profilometry produced a reaction efficiency of $3.16 \pm 0.1 \times 10^{-24}$ cm³/atom, using 2.3×10^{20} atoms/cm² as the fluence estimate. Both numbers are in excellent agreement with the Kapton reaction efficiencies produced by other on-orbit materials experiments such as STS-8¹⁶ (3×10^{-24}), the Long-Duration Exposure Facility¹⁷ (LDEF) (3.0×10^{-24}), and the Intelsat Solar Array Coupon (ISAC) experiment flown on STS-41¹⁸ (3.1×10^{-24}). However the STS-8, ISAC, and LDEF reaction efficiency measurements were all dependent on the MSIS-86 model of the thermosphere for O-atom fluence. An independent measure of Kapton reaction efficiency was made in the LANL HVAB as part of the kinetic energy dependence study reported below, and was found to be 3.3×10^{-24} cm³/atom, within 10 percent of the value based on the space flight exposures.

Several general trends in the relationship between O-atom reactivity and molecular structure are visible in table 2. For example, polyethylene, Tedlar, Tefzel, Kynar and Teflon are all linear carbon chain polymers with increasing fluorine content and decreasing hydrogen content as we move along the series from polyethylene, (CH₂-CH₂)_n, or polypropylene, to Teflon, (CF₂-CF₂)_n, or FEP Teflon. As can be seen in table 2, increasing fluorine content results in decreasing O-atom reaction efficiency, as we would expect if hydrogen atom abstraction is a rate-limiting process and fluorine atom abstraction occurs to a very limited extent, if at all. The EOIM-III reaction efficiency for Teflon is between that reported from STS-8¹⁶ and LDEF.¹⁹ We attribute the observed range of reaction efficiency values to different net doses of solar UV/VUV radiation in the different mission environments. Laboratory studies in the JSC flowing discharge system have shown that vacuum ultraviolet photochemistry is the controlling factor in the O-atom chemistry of Teflon and Kel-F.²⁰ The EOIM-III payload received a larger VUV radiation dose than STS-8 as a result of the solar inertial hold period following deployment of the European Retrieval Carrier satellite during an earlier portion of the STS-46 mission.

In contrast, incorporating two CF₃ groups into a polyimide structure results in little or no change in reaction efficiency, as can be seen by comparing the reaction efficiencies of Kapton polyimide and Eymyd-F. In general, the aromatic polymers displayed significantly lower reaction efficiencies than the linear straight-chain hydrocarbons, with the notable exception of the polycarbonate. The very low reaction efficiency of the poly(bistrifluoropropylphosphazene) based polymers X-221, X-222, and Eypel-F all showed little or no evidence of reaction, confirming earlier work in ground-based test facilities.²¹ Eypel-F is a durable, high-temperature elastomer which should find application in spacecraft atomic oxygen environments.

TABLE 2 THE EFFECTS OF ATOM-SURFACE COLLISION ENERGY ON THE REACTION EFFICIENCY, RE, AND THE PARAMETERS OF THE EMPIRICAL ARRHENIUS EQUATION, $RE = A \times \exp(-EA/KTS)$, WHERE TS IS POLYMER SURFACE TEMPERATURE AND EA IS THE ACTIVATION ENERGY IN EV. RE IS IN CM³/ATOM.

POLYMER	Re, LEO	Ea, LEO	Re,HVAB	Ea,HVAB	Re, FDS	Ea, FDS
KAPTON	3.1E-24	0.02 eV	3.3E-24	0.01 eV	2 E-28	0.3 eV
MYLAR	3.8E-24	0.05 eV	_____	_____	3 E-28	0.4 eV
D4PE	3.8E-24	0.0 eV	_____	_____	2 E-27	0.2 eV
PE	3.7E-24	0.0 eV	_____	_____	4 E-27	0.2 eV
KYNAR	1.2E-24	0.0 eV	_____	_____	3 E-29	0.4 eV
TEFZEL	0.9E-24	0.04 eV	_____	_____	3 E-29	0.5 eV
LCP-4100	3.2E-24	0.04 eV	_____	_____	_____	_____
XYDAR	2.9E-24	0.05 eV	_____	_____	_____	_____
CR-39	6.1E-24	0.04 eV	_____	_____	_____	_____
EYMYD-F	2.7E-24	0.03 eV	_____	_____	_____	_____
PEEK	3.4E-24	0.03 eV	_____	_____	_____	_____

Table 2 shows the temperature dependence of the polymer reaction efficiencies determined following exposure to a known O-atom fluence on the EOIM-III heated trays, in the flowing discharge apparatus, and in the HVAB. The temperature dependence of the O-atom reaction efficiency is shown as an empirical Arrhenius activation energy, i.e., the natural logarithm of the reaction efficiency is plotted against the reciprocal of the polymer sample temperature in degrees Kelvin, and the activation energy is determined from the slope. For all the cases examined to date, straight-line Arrhenius plots have been obtained with correlation coefficients between 0.95 and 0.99. Inspection of table 2 shows that a large decrease in the Arrhenius activation energy is obtained when moving from the flowing discharge to the HVAB or orbital environments. The large decrease in activation energy is accompanied by the large increase in polymer reaction efficiency.

O-atom kinetic energy does not appear in the Arrhenius equation. As a result, the activation energy calculated by this method can vary with O-atom kinetic energy, if atom kinetic energy is available to overcome energetic barriers to reaction as has been previously proposed.^{11,22} Alternately, the mechanism of reaction could change as atom-kinetic energy approaches a threshold value. Simple, empirical power laws or exponential functions have been

shown to produce reasonable agreement with the limited data available at that time in the 0.065 to 5.0 eV kinetic energy range which suggests that a single reaction mechanism as well as a single energetic barrier to reaction, may determine the reaction efficiency in the O-atom kinetic energy domain of interest. The question cannot be resolved without reaction efficiency data taken at several kinetic energies between 0.1 and 1.0 eV.

The LANL HVAB was used to obtain reaction efficiency data on Kapton polyimide at average atom kinetic energies (first moments of the kinetic energy distribution function) of 0.44, 0.72, 0.79, and 2.1 eV. Velocity distribution functions and HVAB composition measured were determined as described in the apparatus and methods section above. A typical HVAB O-atom kinetic energy distribution function (average kinetic energy = 0.79 eV) is shown in figure 4. For comparison purposes the kinetic energy distribution functions for ram-incident O atoms in LEO (average kinetic energy = 5.6 eV) and for O atoms striking a surface immersed in flowing discharge gas (average kinetic energy = 0.065 eV) are shown in figure 5. The flowing-discharge kinetic-energy distribution function is calculated as for an effusion beam, not for the normal component of the collision energy. The normal component is probably not appropriate in this case because the polymer surfaces are not highly oriented or highly crystalline.

The measured reaction efficiency of Kapton polyimide is plotted against the first moment (average value) of the kinetic distributions described in the previous paragraph in figure 6. A rapid increase in reaction efficiency is seen between 0.065 and 1.0 eV followed by relatively little change between 1.0 and 5.6 eV. The data shown in figure 6 suggest that a simple line-of-centers or Beckerle reactive collision model can be used to describe the data. Such models have proven highly successful in describing the kinetic energy dependence of a number of gas phase reactive scattering processes.²³⁻²⁶ An empirical inverse exponential model used to describe the collision-induced chemisorption of methane on nickel²⁷ should also provide a useful description of the kinetic energy dependence of O-atom reactions with polymers. A simple direct fit of the data plotted in figure 6 to such models is a gross oversimplification given the width of the kinetic energy distribution functions. A more accurate test of the kinetic energy dependence hypothesis is needed.

To test the hypothesis that the simple line-of-centers and Beckerle-Ceyer collision models provide a reasonably accurate description of the reaction dynamics of O atoms with polymers, we form the convolution integrals of the models with the normalized kinetic energy distribution function, $f(E_t)$, as shown in the equations below and then determine if the Re vs. Et data can be fit to the resulting functions. Finally, we ask if the Re equations, with parameters determined by least squares curve fitting to the

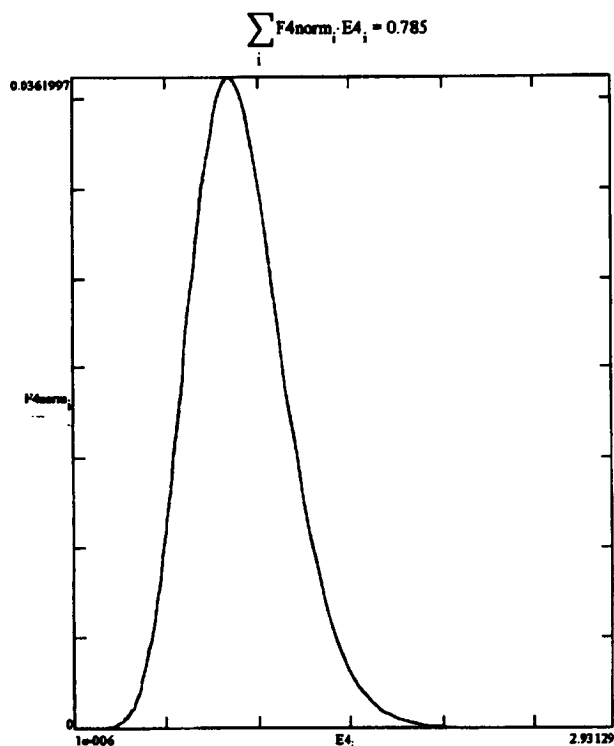


Figure 4. The kinetic energy distribution function of HVAB O atoms in the lab frame. HVAB operating parameters were set to product an average kinetic energy of 0.79 eV.

HVAB data, can predict values of Re for the flowing discharge and on-orbit environments.

LINE-OF-CENTERS MODEL

$$Re = \int_0^{\infty} A \left(1 - \frac{Ea}{Et}\right) f(Et) d(Et),$$

$$A = 5.3 \times 10^{-24} \text{ cm}^3/\text{atom} \quad Ea = 0.64 \text{ eV} \quad \Delta = 0.016 \quad (1)$$

BECKERLE MODEL

$$Re = \int_0^{\infty} \frac{A}{1 + \exp(-n(Et - Ea))} \times f(Et) d(Et)$$

$$A = 3.7 \times 10^{-24} \quad n = 10 \quad Ea = 0.98 \quad \Delta = 0.008 \quad (2)$$

Re , as defined by the Re equation above, is the average of a large number of reaction efficiencies, one for each kinetic energy interval in the kinetic energy distribution function of interest. The Re equation allows us to calculate the reaction efficiency given the normalized kinetic energy distribution function, $f(Et)$, and values for the pa-

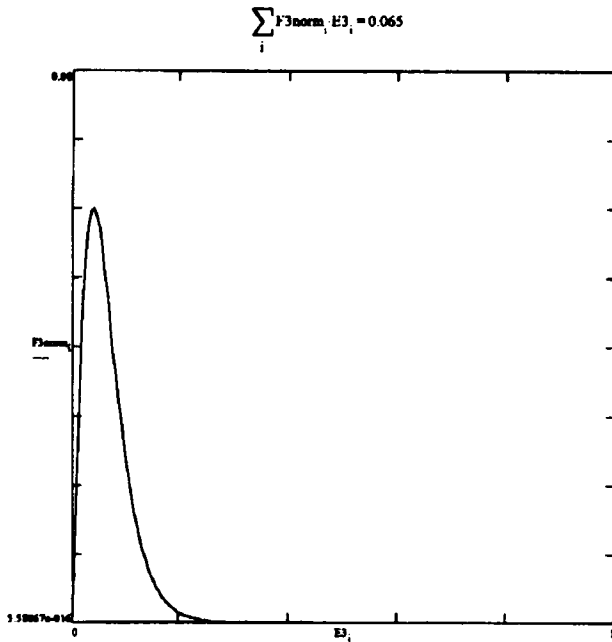
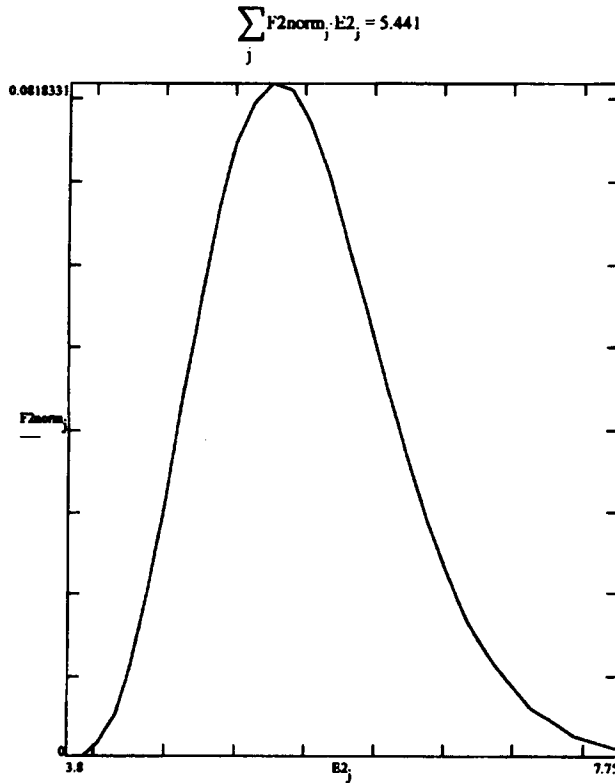


Figure 5. The kinetic energy distribution functions for: (a) O atoms striking a ram-oriented surface in LEO, and (b) O atoms striking a surface in the flowing discharge apparatus.

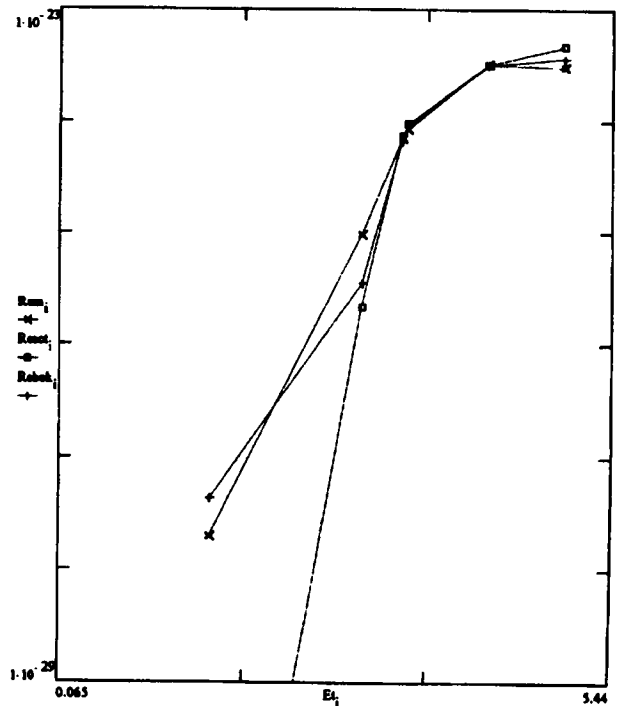


Figure 6. Measured values of Re of Kapton polyimide and Re values calculated with the Beckerle and line-of centers models which are both plotted against the first moment (average value) of the corresponding O-atom kinetic energy distribution functions.

rameters E_a , the magnitude of the energetic barrier to reaction, and A , the limiting reaction efficiency at large kinetic energies. The term is the residual sum of squares error at the conclusion of the curve-fitting process.

We test the hypotheses represented by the R_e equations as follows. First, because a priori values for A and E_a are not available, a gaussian least squares curve-fitting process is used with A and E_a as adjustable parameters. A and E_a are adjusted so that the R_e equation gives the best fit (minimum sum of squares error between measured and calculated R_e) to the R_e data produced by exposing Kapton samples in the four different atom beam kinetic energy distribution functions produced by the HVAB. The success of the curve-fitting operation both in terms of the reasonableness of the A and E_a values obtained and the magnitude of the sum of squares error at the end of the curve-fitting process is one test of the validity of the model. A second test involves asking how accurately the R_e equations, with A and E_a values determined as described above, can predict R_e values for kinetic energy distributions well outside the range of values used in the least squares process. Specifically, can the R_e equations, with A and E_a determined with HVAB data, predict R_e values obtained from the flowing discharge and EOIM-III experiments?

The predictions of the R_e equations are plotted with the measured R_e values in figure 6. Clearly, the Beckerle equation provides a reasonably accurate description of the kinetic energy dependence of the Kapton R_e over a four order-of-magnitude range in reaction efficiency and a two order-of-magnitude range in atom kinetic energy. The agreement between theory and experiment is excellent given that the EOIM-III and flowing discharge R_e values were not included in the original data-fitting process and in light of the significant differences between the three O-atom environments. In contrast, the line-of-centers model provides a good fit to the HVAB data and a reasonably accurate prediction of the on-orbit reaction efficiency, but underestimates the thermal atom reaction efficiency by six orders of magnitude. The line-of-centers model assumes a collision diameter and potential energy surface for the reactive event which is independent of the collision kinetic energy. Further work in this area will focus on developing improved line-of-centers models which accounts for changes in the shape of the potential energy surface of the reaction with changing O-atom kinetic energy.

RESULTS AND DISCUSSION: DUAL USE TECHNOLOGY

Investigating the materials chemistry of atomic oxygen in the LEO environment has led directly to two new technologies with commercial potential. First, the anisotropic etching of photoresists or other microelectronic materials using high-velocity O-atom beams results directly from the strong dependence of reactivity on the kinetic energy

of the atomic projectile coupled with very efficient momentum accommodation in nonreactive collisions. (Scattered atoms are much less reactive than primary projectiles in some circumstances.) Second, the reaction-diffusion kinetics of low-velocity (thermal) atomic oxygen in porous solids can be exploited to engineer a new class of porous solid materials with different pore surface chemistries in different parts of the solid. Application areas for the resulting distributed pore chemistry solids (DPC solids) include supports for cell and tissue culture in the biotechnology industry, membranes for separation and catalysis, chromatographic separation media and supported reagents for organic synthesis.

Various anisotropic etching processes are used to pattern microcircuits and microdevices on integrated circuit chips. As silicon device technology is pushed to ever smaller sizes to produce microcircuit chips with ever higher device densities, the limitations of state-of-the-art microcircuit fabrication technologies become apparent. Smaller devices mean thinner oxide gates in MOS devices, ion implantation, or diffusional doping to smaller depths, and smaller volumes for gates and junctions resulting in a greater sensitivity to device damage during fabrication. Reactive ion etching and plasma resist stripping can cause damage to small devices by a variety of mechanisms including dielectric breakdown of thin dielectric layers, undesirable ion implantation damage and heating of the chip so as to cause diffusion of dopants. The limitations of existing dry etching technologies as applied to future device generations is the principal motivation for developing neutral beam anisotropic etching technologies (see reference 30 and references therein).

Anisotropic etching of polymer films by high-velocity neutral O atoms was first suggested by electron photomicrographs of polymer films exposed to ram O-atom flux in LEO.¹⁶⁻¹⁸ Small inert particles on the polymer film surface have clearly acted as adventitious etch masks, protecting the underlying polymer from attack and resulting in the formation of the tall tower-like features visible in figure 7. Trilayer-photoresist anisotropic-etching targets were obtained from AT&T Bell Laboratories and were subjected to O-atom beam etching in the HVAB at LANL, and the results are shown in figure 8. Anisotropic etching with high-velocity O-atom beams is capable of sub-micron resolution. Because the plasma processes that produce device damage are not available in the HVAB, submicron device structures may be fabricated without plasma-induced device damage. Additional work is in progress at LANL, funded by the Department of Energy (DOE), exploring other uses of high-velocity O-atom beams in the fabrication of microelectronic materials. The principal objective of the DOE work is low-temperature formation of pure gallium arsenate films on gallium arsenide.²⁸ LANL and JSC participate in the SEMATECH Neutral Beam Working Group. SEMATECH is a government-industry consortium char-



Figure 7. Scanning electron photomicrograph of Kapton polyimide polymer film samples: The smooth portion of the film was exposed to ram O-atoms during the EOIM-III exposure and serves as a control. Anisotropic etching around unreactive film inclusions is visible as tower like structures.

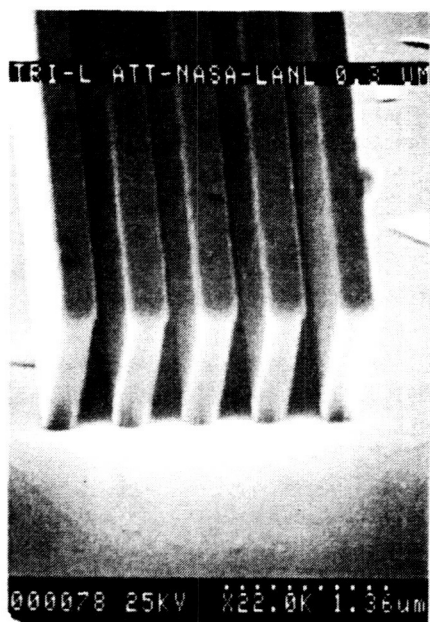


Figure 8. Results of anisotropic etching of a trilayer photoresist etching target provided by AT&T Bell Labs. The HVAB at LANL was the high kinetic energy O-atom source in this case.

tered to maintain American leadership in semiconductor fabrication technology.

The application of thermal energy O atoms to surface chemistry engineering of porous solids was suggested by the peculiar patterns in the loss of waterproofing from porous silica tiles used in the Space Shuttle thermal protection system which could not be explained by reentry

heating.²⁹ Waterproofing consists of blocking the hydrophilic silanol (Si-OH) groups on all pore surfaces throughout the tile with hydrophobic trimethyl silyl groups (Si-O-Si(CH₃)₃). Atomic oxygen removes waterproofing by oxidizing the methyl groups to carbon oxides and water leaving a new layer of Si-OH lining the pores. When a fully waterproofed porous solid is placed in the oxygen plasma asher or flowing discharge environment, the oxidation of methyl groups proceeds as a sharply defined reaction-diffusion front that moves into the porous solid from the exterior. As a result, a well defined boundary exists between the hydrophilic and hydrophobic regions in the porous solid. The observation that thermal O atoms change the surface chemistry of porous solids from the outside in, creating well defined regions with different surface chemistry rather than uniform alteration of the whole porous solid, is the basis of DPC solids technology.³¹

By controlling total pressure, O-atom concentration, and exposure time, we have produced DPC solids from trimethyl silylated 80/100 mesh size Porasil C as well as 10-micron diameter porous silica beads used in the chromatographic separation of peptides and proteins (300 Angstrom nominal pore size, C8 or C18 bonded stationary phase). In addition, porous membranes made of a variety of polymeric materials have been prepared with opposing hydrophilic and hydrophobic surfaces. The hydrophilic surfaces produced by O-atom attack on polymers contain functional groups which can be used to covalently attach a wide range of molecules to fine tune surface properties for specific applications. In addition, porous polyethylene materials (Porex) have been rendered hydrophilic to controlled depths³¹ as shown in figures 9 and 10. In contrast with surface modification, which proceeds rapidly inward from the exposed surface, net mass loss and pore enlargement were observed to occur uniformly throughout the Porex sheets up to 3 millimeters thick. Under conditions of constant O-atom flux, mass loss rates are linear for both high density polyethylene and the three Porex porosity grades as shown in figure 11. Scanning electron photomicrographs of Porex polyethylene controls are compared with samples exposed to the flowing discharge long enough to produce: (1) an 8 percent mass loss in the coarse porosity sample, (2) a 19 percent mass loss in the medium porosity sample, and (3) an 18 percent mass loss in the fine porosity sample as shown in figures 12-14. Inspection of figures 12-14 shows that mass loss is occurring uniformly throughout the cross section of the porous solids. The application of DPC materials to biotechnology problems is the subject of continuing work at JSC. Engineered porous solids with high surface areas for growth of surface attachment dependent cells could be incorporated in new type culture vessels with the same footprint as the conventional T-flasks. Porous polymeric microcarriers for the culture of fragile animal cells in stirred tank bioreactors is a second potentially fruitful application area.

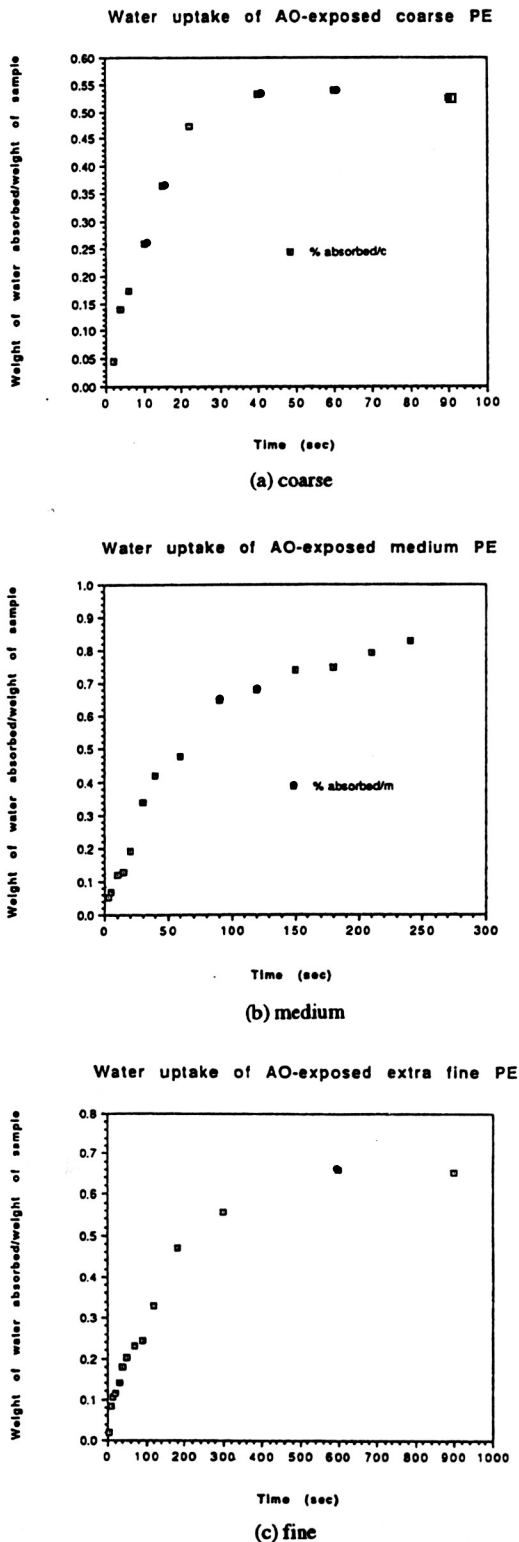


Figure 9. The development of hydrophilic properties in Porex polyethylene following exposure to thermal energy O-atoms in the JSC flowing discharge. The weight of water taken up by the treated Porex is plotted against exposure time for: (a) coarse, (b) medium, and (c) fine porosity Porex.

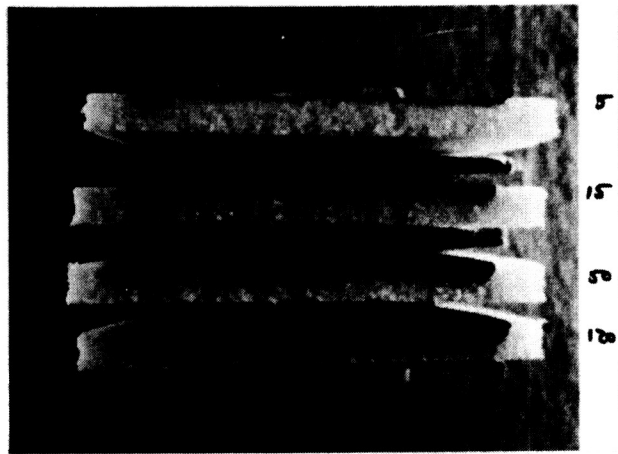


Figure 10. The movement of the hydrophilic region into fine porosity Porex as a function of time. The hydrophilic region, marked by the blue dye, moves from the surface exposed to the O-atom containing gas into the interior of the solid as a function of time. Top to bottom; (a) 5 seconds, (b) 15 seconds, (c) 50 seconds, and (d) 120 seconds.

SUMMARY AND CONCLUSIONS

Characterization of the materials chemistry of atomic oxygen in the LEO environment has provided valuable data for the NASA Space Station Program, the Ballistic Missile Defense Program and other space programs operating in the LEO environment. The role of O-atom kinetic energy on collision with the reacting surface in determining the rate of polymer reactions has been demonstrated unequivocally as has the role of VUV radiation in the reactivity of fluorocarbon polymers. An extensive data base of polymer reaction efficiency data has been produced and verified and is available to the technical community. Improved understanding of the mechanism and dynamics of O-atom reactions leads directly to more reliable ground-based materials testing protocols and, ultimately, to more reliable spacecraft. Investigating the dynamics and mechanisms of O-atom reactions has also resulted in some significant contributions to the applied chemical and materials sciences as well as several promising "spin-off" or "dual use" technologies in the areas of microelectronic circuit fabrication technology and materials for applied chemistry and biotechnology.

REFERENCES

1. Jursa, A. S., editor; Handbook of Geophysics and the Space Environment, Air Force Geophysics Laboratory, United States Air Force, 1985 (National Technical Information Service).
2. Hedin, A. E., "A Revised Thermospheric Model Based on Mass Spectrometer and Incoherent Scatter Data: MSIS-83," J. Geophys. Res., Vol. 88, No. 10, pp 170-188, 1983.

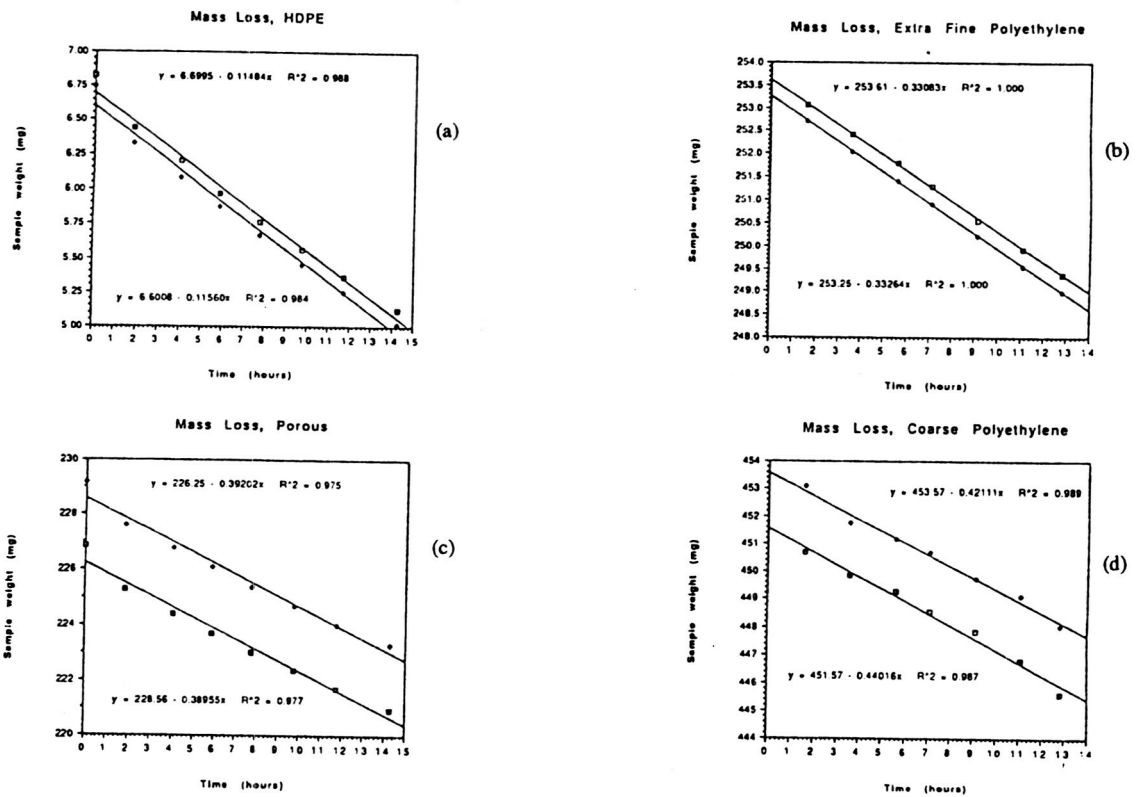


Figure 11. Mass loss rates for the three Porex polyethylene porosity types and high density polyethylene (HDPE): (a) HDPE, (b) fine, (c) medium, (d) coarse. The two lines on each plot are the mass loss results for duplicate samples. The thermal O-atom flux was 5×10^{18} O-atoms/cm²sec in all cases.

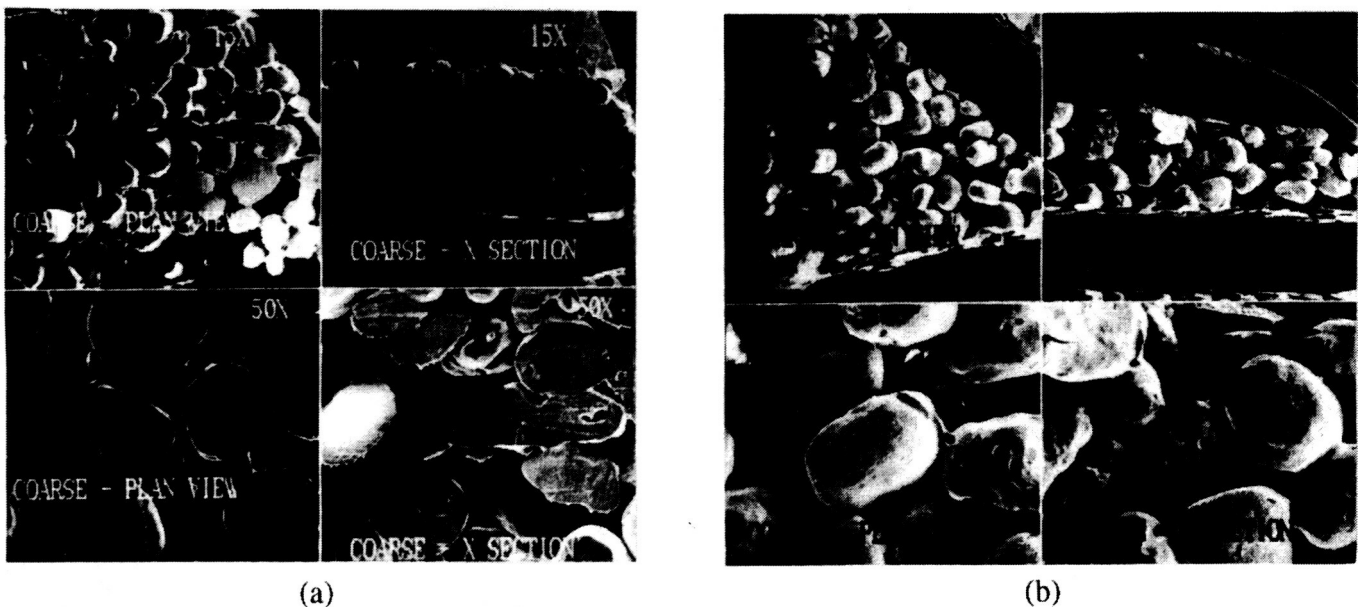
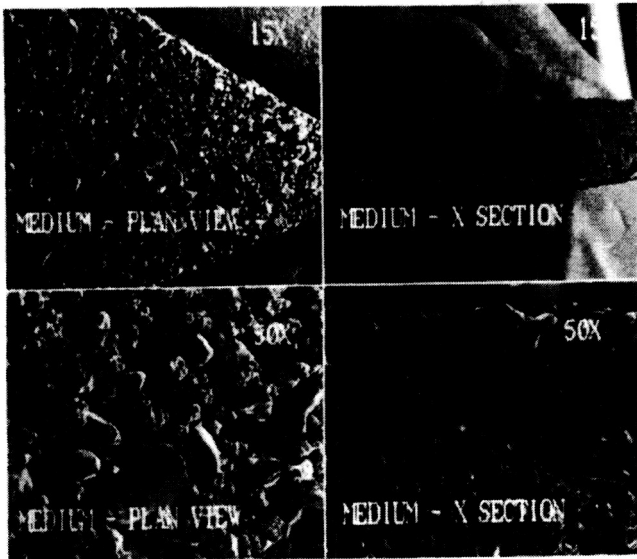
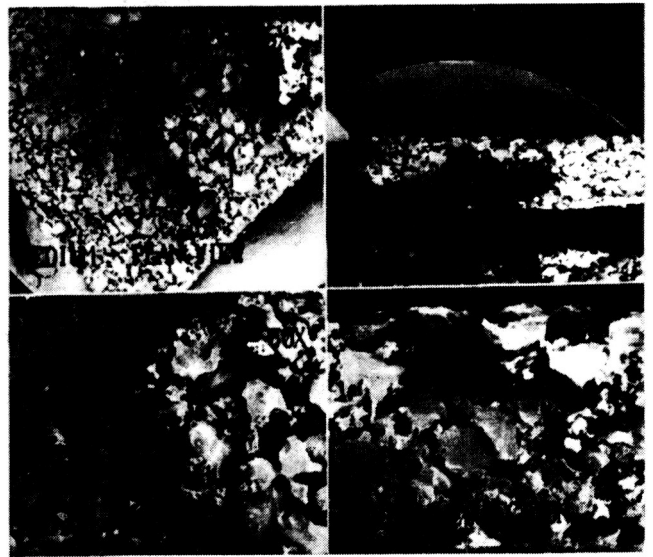


Figure 12. Face-on (plan view) and cross sections of coarse porosity Porex polyethylene before (a) and after (b) 165 hours exposure to the O-atom flux of figure 11.

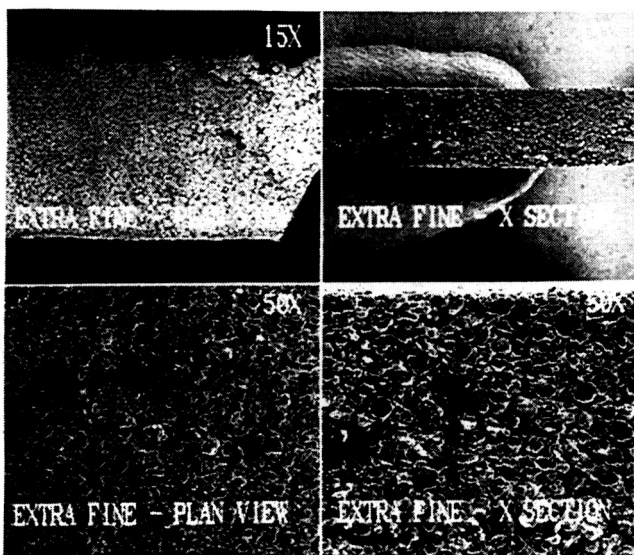


(a)

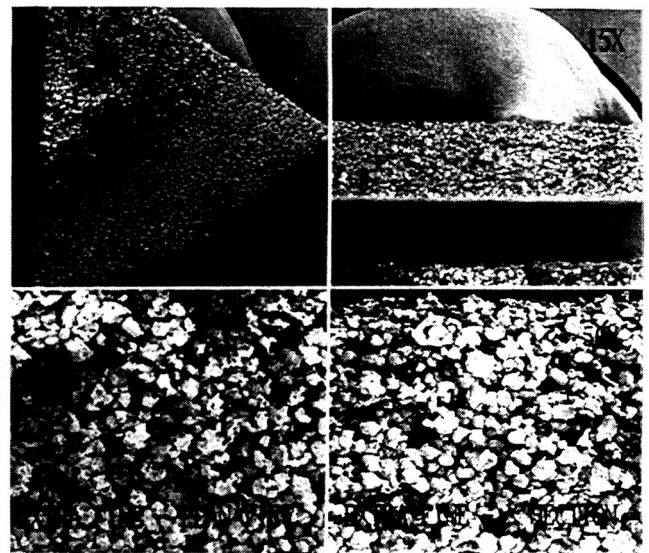


(b)

Figure 13. Face-on (plan view) and cross sections of medium porosity Porex polyethylene before (a) and after (b) 165 hours exposure to the O-atom flux of figure 11.



(a)



(b)

Figure 14. Face-on (plan view) and cross sections of fine porosity Porex polyethylene before (a) and after (b) 165 hours exposure to the O-atom flux of figure 11.

3. Chamberline, J. W., *Theory of Planetary Atmospheres*, Academic Press, New York, 1978, p. 305.
4. Leger, L. J., Visentine, J. T., Santos-Mason, B.; "Selected Materials Issues Associated with Space Station Freedom," in the SAMPE Quarterly, Vol. 18, No. 2, pp 48-54, Jan. 1987.
5. Visentine, J. T., comp.; NASA Technical Memorandum 100459, "Atomic Oxygen Effects Measurements for Shuttle Missions STS-8 and STS-41G: Vol. I-III." Sept. 1988.
6. Visentine, J. T., Leger, L. J.; "Materials Interactions with the Low-Earth Orbit Environment: Accurate Reaction Rate Measurements," AIAA paper AIAA 85-7019, AIAA Shuttle Environments and Operations II Conference, Nov. 1985, p 168.
7. Koontz, S. L., Leger, L. J., Visentine, J. T., Hunton, D. E., Cross, J. B., Hakes, C. L.; "An Overview of the Oxygen Interactions with Materials III Experiment: Space Shuttle Mission 46; July-August 1992," in Proceedings of the Third LDEF Post-Retrieval Symposium, A. Levine, Ed., Nov. 8-12, 1993, Williamsburg Va., A NASA Conference Publication.
8. Koontz, S. L., Rickman, S. L., Leger, L. J., Cross, J. B., Hakes, C. L., Bui, D. T.; "EOIM-III Mission and Induced Environments," in Proceedings of the Third LDEF Post-Retrieval Symposium, A. Levine, Ed., Nov. 8-12, 1993, Williamsburg Va., A NASA Conference Publication.
9. Cross, J. B., Blais, N. C.; "High Energy/Intensity CW Atomic Oxygen Beam Source," Progress in Aeronautics and Astronautics, Vol. 116, pp 143-155, American Institute of Aeronautics and Astronautics, Washington, DC, 1989.
10. Koontz, S. L., Cross, J. B., Lan, E.; "Characterization and Calibration of the EOIM-III Flight Mass Spectrometer in a High Velocity Atom Beam," in Materials Degradation in Low-Earth Orbit (LEO), Srinivasan, V., Banks, B. A., eds., The Materials, Metals and Minerals Society, 1990.
11. Koontz, S. L., Albyn, K., Leger, L. J.; "Atomic Oxygen Testing with Thermal Atom Systems: A Critical Evaluation," J. Spacecraft, Vol. 28, No. 3, pp 315-323, May-June, 1991.
12. Koontz, S. L., Albyn, K., Leger, L.; "Materials Selection for Long Life in Low Earth Orbit," Journal of the IES, March/April 1990, pp 50-59.
13. Koontz, S. L., Nordine, P.; "The Reaction Efficiency of Thermal Energy Oxygen Atoms with Polymeric Materials," in Materials Degradation in Low Earth Orbit (LEO), Banks, B. A., Srinivasan, V., Eds. pp 189-205, The Minerals, Metals and Materials Society, Warrendale Penn., 1990.
14. Huie, R. E., Herron, J. T.; "Reactions of Atomic Oxygen (O³P) with Organic Compounds," Progress in Reaction Kinetics, Vol.8, No. 1, pp1-80, 1975.
15. Fergeson, E. E., Feshenfeld, F. C., Schmeltenkoph, A. L.; "Flowing Afterglow Measurements," in Advances in Atomic and Molecular Physics, Vol.5, pp 12-13, Academic Press, New York, 1969.
16. Visentine, J. T., Leger, L. J., Kuminecz, J. F., Spiker, I. K.; "STS-8 Atomic Oxygen Effects Experiment," AIAA paper AIAA-85-0415, 23rd Aerospace Sciences Meeting, Jan. 14-17, 1985, Reno, Nevada.
17. Personal communication, Professor John Gregory, Chemistry Department, University of Alabama, Huntsville, Alabama.
18. Koontz, S., King, G., Dunnet, A., Kirkendahl, T., Linton, R., Vaughn, J.; "The ISAC Atomic Oxygen Flight Experiment," J. Spacecraft, in press, 1994.
19. Steigman, A. E., Brinza, D. E., Anderson, M. S., Minton, T. K., Laue, E. G., Liang, R. H.; "An Investigation of the Degradation of Fluorinated Ethylene Propylene (FEP) Copolymer Thermal Blanketing Materials Aboard LDEF and in the Laboratory," JPL publication 91-10, May 15, 1991, Jet Propulsion Laboratory, Pasadena, California
20. Koontz, S. L., Leger, L. J., Albyn, K. A., and Cross, J.; "Ultraviolet Radiation/ Atomic Oxygen Synergism in Materials Reactivity," J. Spacecraft, Vol. 27, No. 5, pp 346-348, May-June 1990.
21. Fewell, L., Fenney, L.; Polymer Comm., Vol. 32, No. 13, pp 393-396, 1991.
22. Arnold, G. S.; Peplinski, D. R.; Cascarano, F. M.; "Translational Energy Dependence of the Reaction of atomic Oxygen with Polyimide Films," J. Spacecraft, 24(5), pp 454-458, 1987.
23. Gonzalez Urena, A, "Influence of Translational Energy upon Reactive Scattering Cross Section: Neutral-Neutral Collisions," in Advances in Chemical Physics, Prigogine, I and Rice S. A. Eds. Vol. LXVI, John Wiley and Sons, New York, 1987, pp 231-335.
24. Stienfeld, J. I.; Francisco, J. S.; Hase, W. L. Chemical Kinetics and Dynamics, Chapter 8, "Dynamics of Bimolecular Collisions," Prentice Hall, New Jersey, 1989.
25. Zeiri, Y.; Lucchese, R. R. "Collision of hyperthermal atoms with an adsorbate covered solid surface II: Collision induced desorption," Surface Science, 264, pp 197-206, 1992.
26. Levine, R. D., Bernstein, R. B.; "Post Threshold Energy Dependence of the Cross Section for Endoergic Processes: Vibrational Excitation and Reactive Scattering," J. Chem. Phys., 56(5), March 1, 1972.
27. Beckerle, J.D.; Johnson, A.D.; Yang Q. Y.; Ceyer, S. T. "Collision Induced Dissociative Chemisorption of CH₄ on Ni: The Mechanism for Chemistry With a Hammer," J. Chem. Phys., 91(9), pp 5756-5777, 1989.
28. Cross, J. B., Hoffbauer, M. A., Bermudez, B. M.; "Growth of Oxide Layers on Gallium Arsenide with a High Kinetic Energy Atomic Oxygen Beam," Appl. Phys Letters, 57 (21), page 2193-2195, Nov. 1990.
29. Stevens, E. G., "Shuttle Thermal Protection System Atomic Oxygen Exposure Study," Laboratory Test

- Report 4081-4546, Rockwell International, Space Transportation System Division, February 1985.
30. Koontz, S. L., Cross, J. B.; United States Patent 5, 271,800, "Method for Anisotropic Etching in the Manufacture of Semiconductor Devices," Dec. 12, 1993.
31. Koontz, S. L.; United States Patent 5,141,806, "Microporous Structure with Layered Interstitial Surface Treatment, and Method and Apparatus for Preparation Thereof," Aug. 25, 1993.

S37-27

447-9

APPLICATION OF CHLORINE-ASSISTED CHEMICAL VAPOR DEPOSITION OF DIAMOND AT LOW TEMPERATURES

Chenyu Pan*, David A. Altemir[†], John L. Margrave*, and Robert H. Hauge*

*Department of Chemistry, Rice University, P.O. Box 1892, Houston, Texas 77251

[†]NASA/Johnson Space Center, Mail Code ES5, Houston, Texas 77058

ABSTRACT

Low temperature deposition of diamond has been achieved by a chlorine-assisted diamond chemical vapor deposition (CA-CVD) process. This method begins with the thermal dissociation of molecular chlorine into atomic chlorine in a resistively heated graphite furnace at temperatures between 1300 and 1500 °C. The atomic chlorine, upon mixing, subsequently reacts with molecular hydrogen and hydrocarbons. The rapid exchange reactions between the atomic chlorine, molecular hydrogen, and hydrocarbons give rise to the atomic hydrogen and carbon precursors required for diamond deposition. Homoepitaxial diamond growth on diamond substrates has been studied over the substrate temperature range of 100-950 °C. It was found that the diamond growth rates are approximately 0.2 $\mu\text{m/hr}$ in the temperature range between 102 and 300 °C and that the growth rates do not decrease significantly with a decrease in substrate temperature. This is unique because the traditional diamond deposition using H_2/CH_4 systems usually disappears at substrate temperatures below ~ 500 °C. This opens up a possible route to the deposition of diamond on low-melting point materials such as aluminum and its alloys.

INTRODUCTION

It is generally accepted that the presence of atomic hydrogen and carbon precursors such as methyl radicals are critical in the chemical vapor deposition (CVD) of diamond. Atomic hydrogen is generated largely by the high temperature dissociation of molecular hydrogen using hot filament (> 2000 °C) and plasma (> 4000 °C) methods. Recently we have developed a chemical method— chlorine-assisted CVD (CA-CVD) [1] — to produce atomic hydrogen and carbon precursors by first dissociating relatively weakly bonded molecular chlorine into atomic chlorine and subsequently mixing the atomic

chlorine with molecular hydrogen and hydrocarbons. The dissociation of molecular chlorine is accomplished by feeding the gas through a resistively heated graphite furnace at temperatures ≤ 1500 °C. The rapid exchange reactions between atomic chlorine, molecular hydrogen, and hydrocarbons give rise to atomic hydrogen and carbon precursors. Because of the nature of all carbon materials in a CA-CVD reactor and the low process temperature, it is possible to deposit electronic grade diamond films that are free of nitrogen and metal impurities which are usually associated with the presence of high process temperatures and hot filament materials.

EXPERIMENTAL WORK

Diamond films have been successfully deposited by the CA-CVD method on a variety of substrate materials such as platinum, tungsten, molybdenum, and silicon nitride over a broad temperature range. Qualitative studies of diamond deposition on heterogeneous substrates indicate that the quality of the diamond deposition depends on H_2/Cl_2 ratios, substrate temperatures, and methane concentrations. By controlling process parameters, transparent diamond films which are free of graphite and amorphous carbon have been deposited.

Quantitative studies of homoepitaxial diamond deposition on diamond substrates have also been carried out. The *in situ* measurement of the diamond growth rate is made possible by the employment of an optical interferometer that has a sensitivity of 1.3 nm (thickness change) using a 633 nm HeNe laser and a 2.4 refractive index of diamond. The diamond growth rate is found to increase with the increase of substrate temperatures and methane and chlorine concentrations. Under optimized conditions, diamond growth rates can easily reach 5-10 $\mu\text{m/hr}$. About 1 kW of power is used in the furnace to achieve these growth rates.

Homoepitaxial diamond deposition has been accomplished over a temperature range of 100-950 °C [2]. Our previous work has demonstrated that the addition of chlorine in the form of HCl to hot filament H_2/CH_4 systems can lower the substrate temperature without affecting the rate of deposition. CA-CVD is even more promising because the diamond growth rate in CA-CVD has been found to have less temperature dependence in the temperature range of 100-300 °C. This means that the diamond deposition rate does not decrease with substrate temperatures as rapidly as is the case in a pure H_2/CH_4 system. This is evidenced by ~ 0.2 $\mu\text{m/hr}$ diamond growth rate in the 100-300 °C range in CA-

CVD compared to the disappearance of diamond growth deposition in a pure H₂/CH₄ system at substrate temperatures below 500 °C.

The preliminary results from CA-CVD scaleup studies indicate that larger deposition areas can be achieved by the design of a linear reactor. Diamond deposition in this manner has been found to be uniform.

POTENTIAL APPLICATIONS

Composites

The mechanical properties of synthetic diamond make it an attractive candidate for use in composite materials. Although continuous-fiber composites are conceivable, the idea of producing continuous diamond fibers is somewhat problematic due to the limitations of present-day diamond growth technologies. It is, however, more practical to consider diamond or diamond-coated particulates dispersed within a metallic matrix. In this scenario, the diamond dispersoids would serve as obstacles to crack propagation and would impede dislocation motion. Therefore, one would expect to improve the fracture toughness and yield strength of the matrix material by the addition of diamond dispersoids.

A theoretical prediction of the elastic modulus of a diamond particulate/Aluminum 6061-T6 composite is possible using the micromechanical approach of Hashin [3]. Results obtained by this method are shown in Figure 1. For comparison, the theoretical modulus of a hypothetical diamond continuous-fiber composite is also shown.

A prediction of composite strength can also be made by assuming that the diamond dispersoids impede dislocation motion in a manner known as dispersion strengthening. The dispersion strengthening theoretically predicted for a diamond/aluminum particulate composite (shown in Figure 2) is approximately the same magnitude as that which has been experimentally observed for an analogous silicon carbide/Al 6061-T6 system [4]. However, the degree to which strengthening (or weakening) will actually occur depends on the character of the particle/matrix interface. Therefore, a program of mechanical testing and particle/matrix interface characterization is planned.

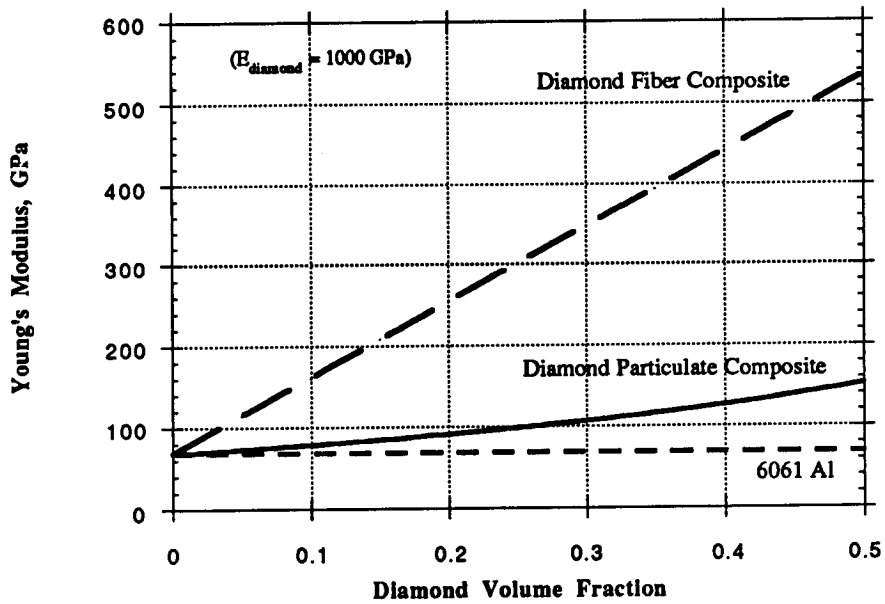


Figure 1. Theoretical Assessment of Elastic Modulus for a Diamond/Al 6061-T6 Particulate Composite

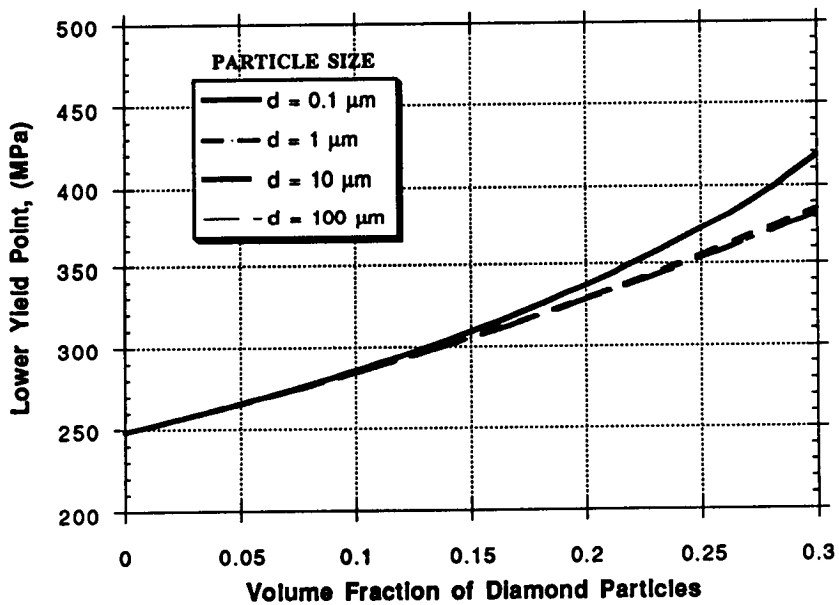


Figure 2. Theoretical Dispersion Strengthening Assessment of a Diamond/Al 6061-T6 Particulate Composite.

Fluidized Bed Growth

Related to the development of the new composite materials just discussed is the ability to coat very small ($\sim 100 \mu\text{m}$) particles with diamond. Conventional plasma diamond growth methods are limited because the active region in which deposition occurs is very small (typically $5\text{-}8 \text{ cm}^2$) and the temperatures at which deposition occurs are very high ($>2000^\circ\text{C}$). Therefore, it is compelling to adapt the lower temperature chlorine-activated process to a fluidized bed process.

In this concept, a bed of particles is suspended or "fluidized" in a low-pressure reactive gas flow. The turbulence associated with such a process produces a uniform distribution of the active gas species on the particle surfaces and, therefore, uniform diamond growth. However, the partial pressures of reactive and inert fluidizing gases must be optimized to insure that an adequate population of active diamond-growth species exists within the bed. Future work is planned for the development of this concept.

Diamond Deposition on High Temperature Polymers and Low Melting Point Metals

The ability of the low temperature diamond deposition by CA-CVD opens up possible routes to diamond deposition on low melting point metals such as aluminum and its alloys and on high temperature polymers provided that problems such as high heat flux and deleterious surface reactions can be dealt with. These problems can be resolved by providing sufficient cooling to the substrate (and/or distancing the substrate from the furnace) and providing a protective coating on the substrate material. In the case of metals, a metal nitride coating appears attractive as a substrate surface preparation.

The largest foreseen application for this technology is the production of scratch-resistant coatings for light-weight optics and metal/diamond composite heat sinks. So far, most of the experiments that have been tried to date have failed due to the effects of atomic chlorine etching. However, resistance to etching has been observed for polytetrafluoroethylene (PTFE or Teflon). It is anticipated that diamond will be nucleated on this surface as this work continues to mature.

Other Potential Uses

Other potential commercial applications of chlorine-activated diamond growth may include the production of chromatographic separation media, and, perhaps, the use of turbulent flows to coat larger objects such as gears, watch crystals, and mechanical joints.

REFERENCES

1. Pan C., Chu C.J., Margrave J.L., and Hauge R.H., "Chlorine-Activated Diamond CVD", *Journal of the Electrochemical Society*, submitted 1993.
2. Pan C., Margrave J.L., and Hauge R.H., "Temperature Investigation of Diamond Growth Rates in a Chlorine-activated Diamond CVD Reactor", *Journal of Applied Physics Letters*, submitted 1994.
3. Hashin Z., "The Elastic Moduli of Heterogenous Materials", *ASME Journal of Applied Mechanics*, Vol. 29, 1962, pp. 143-150.
4. Schmidt K., Zweben C., and Arsenault R., "Mechanical and Thermal Properties of Silicon-Carbide Particle-Reinforced Aluminum", *THERMAL AND MECHANICAL PROPERTIES OF COMPOSITES*, ASTM STP 1080, J.M. Kennedy, H.H. Moeller, and W.S. Johnson, Eds., American Society for Testing and Materials, Philadelphia, 1990, pp. 155-164.

529-27

H-1532

ADAPTATION OF AN AMMONIA DETECTING PAINT TO COMMERCIAL APPLICATIONS

by

Renee C. Graef and William A. Mallow
Southwest Research Institute
P.O. Drawer 28510
San Antonio, Texas 78228-0510

BACKGROUND

An ammonia-sensitive coating was developed in feasibility studies sponsored by Lockheed Engineering. The paint comprised a cellulosic binder, a dye, a family of solvents, and plasticizer. The resulting coating formulation proved to be very sensitive to trace levels of Lewis-bases and/or nucleophiles, but was limited to concentrations of approximately 1 ppm and higher. In the months following completion of this Lockheed project, efforts to perfect and optimize the paint formulation continued with internal funding. A unique sensitivity enhancing additive was identified and patented which when applied at concentrations of approximately 10% by weight, transcended the 1 ppm sensitivity barrier to fractional ppb. The paint composition was modified by the addition of certain proprietary agents, which promoted a much more intense yellow color in the coating and enhanced its durability in high-vacuum. Paint preparation is a simple one day procedure that can be completed using conventional equipment. Most of the materials are easily obtained.

The NASA-sponsored project (Contract No. NAS 9-18051) which followed, examined the effect of preflight, flight, and recovery temperatures and conditions on the performance of the detector paint. The existing recipe at the initiation of this program was derived by an a-priori approximation of grade, viscosity, and class of binder which was deemed appropriate at the time. Shelf life and UV durability were evaluated, and general adhesiveness of the coating was enhanced.

The primary objectives of the project were:

- (1) Enhance the durability and reproducibility of the coating's sensitivity to ammonia under high vacuum at elevated temperatures.
- (2) Identify ambient aging effects on sensitivity and reproducibility of the coating's ammonia detection capability.
- (3) Establish ultraviolet (UV) aging impact on coating performance.
- (4) Correct shortcomings in the composition of the paint or resulting coating.

EXPERIMENTAL PROCEDURE

Three paints, with formulations representing a spread of the primary ingredients, were applied to aluminum coupons using an airless spray unit pressurized with nitrogen. One of the formulas failed to retain its original color and integrity after coating the coupons and was dropped from further study. The other two formulations were tested for their responsiveness to ammonia shortly after coating. Coupons from each remaining formula were subjected to short durations of rough vacuum, high vacuum, ultraviolet radiation, and contaminant exposure (CO_2 and CH_4) tests. After completion of each test condition, the coupons were retested for their ammonia sensitivity under inert conditions and ambient pressure conditions. These tests were sufficient to establish the potential use of the material after exposure to all environments except ultraviolet radiation. A test method was then developed to verify that the material would respond to ammonia under high vacuum conditions. Long-term aging under high vacuum and the limits of exposure to ambient conditions and ultraviolet light were evaluated. Finally, compatibility of the coating with a variety of materials was tested.

MEASUREMENT OF COLOR CHANGE

The color response of each coating was evaluated using a photo spectrometric system developed for a medical application which analyzes green and infrared light reflectances characteristics of the coating. The test chamber accepts 1 x 1-inch coupons (Figure 1). Green and infrared light omitters are positioned over the coated coupon, along with a photo diode. Pulses of green, infrared, and ambient light are directed at the coating (Figure 2). Reflected light from the coating is picked up by the photo diode and the resulting signal separated by signal processing so that the change in green reflectance can be monitored.

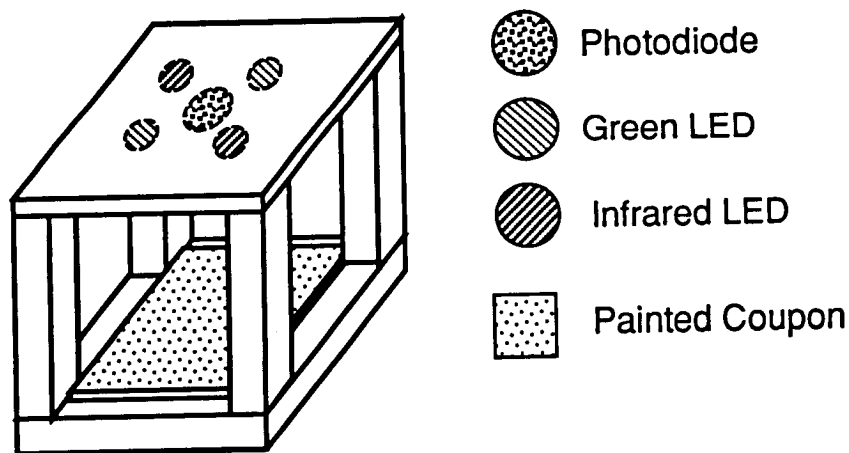


FIGURE 1. Opto-Electric Test Chamber.

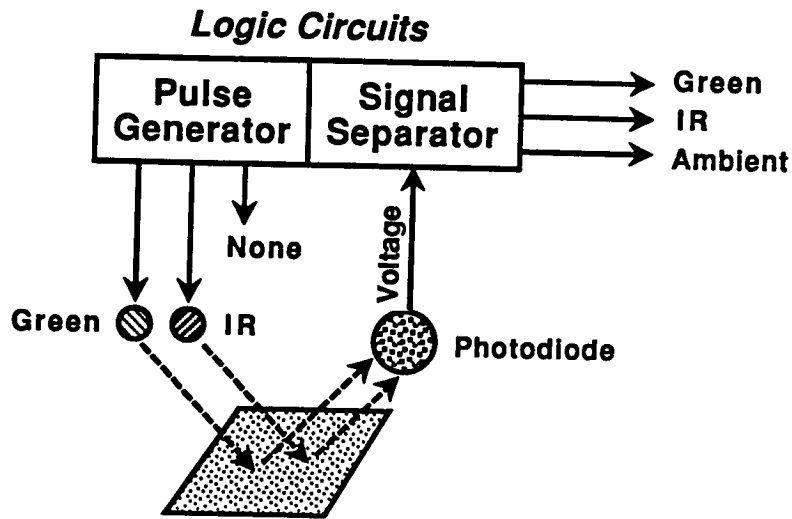


FIGURE 2. Schematic of Opto-Electric Response System.

Ammonia exposure was carried out in a closed system consisting of a bell jar, a reservoir for gas dispersion, a commercial nitrogen tank, and a peristaltic pump for gas circulation (Figure 3). After a coupon was attached to the aluminum base and placed under the bell jar, nitrogen gas was circulated through the system for 10 minutes to flush out the atmospheric air. The system was then closed and ammonia was injected through a tubing injection port. Based on a volume of 20,000 cc, the concentration was computed in advance to the injection. Ammonia concentrations were incrementally increased from 1 ppm to 100 ppm, with time allowed for each increment to disperse before measuring the coating response photoelectrically.

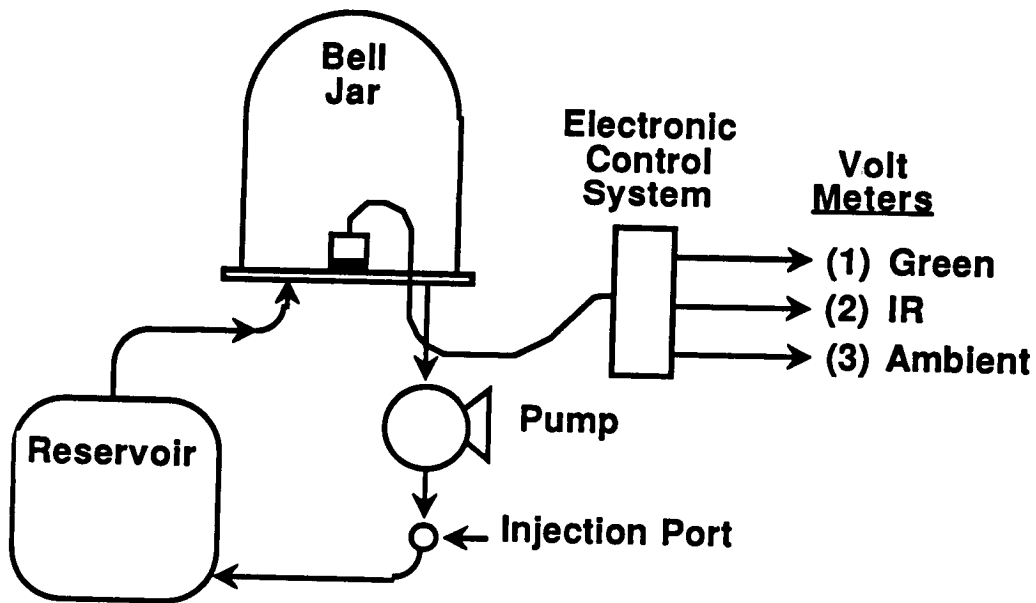


FIGURE 3. Testing Configuration for Phase One.

ENVIRONMENTAL AGING

Aging mechanisms evaluated include ultraviolet radiation, vacuum exposure, ambient conditions, and testing atmospheres of carbon dioxide and methane. Each coupon was tested for its ammonia sensitivity shortly after coating, exposed to the aging mechanism, and then retested as before. Ultraviolet exposure for seventy-five 12-hour days in a Q UV(B) chamber with an ambient air atmosphere at 40°C destroyed the visual and photo-voltaic response of the coating to ammonia. Subsequent studies showed that exposure to UV in the absence of air (via UV(B) bulbs or ambient sunlight) did not significantly diminish ammonia sensitivity. Hence, oxidation in combination with light, rather than light alone, appears to be the important limiting factor.

Vacuum exposure proved detrimental to the coating sensitivity, although, seemingly conflicting results were obtained from rough and high vacuum tests. The coupons exposed to rough vacuum retained only 10% of their original responsiveness, while those exposed to high vacuum retained 43% and 29% (formulas A and B, respectively) of their initial response level. This anomalous behavior is probably due to the manner in which the coupons were stored after testing although damaged surfaces, variations in film thickness, or surface properties, may have contributed to the conflict. A downward trend in coating sensitivity over time was also noted during the ambient aging and contaminant exposure tests. It was discovered later during the high vacuum tests that the amount of ionic media in the coating (in this case water) greatly affects the sensitivity of the dye to ammonia. For these tests, the coupons were stored in a desiccator while waiting for sensitivity evaluation, thus eliminating the absorbed water necessary for the coating to function.

SIMULATED SPACE STATION LEAK (SSSL) TESTING

An SSSL test was designed to simulate leak conditions on the space station. All previous testing of the paint's ammonia sensitivity, done at atmospheric pressure, showed the coating to be a likely candidate for early detection of ammonia leaks. Additional testing, representative of flight condition responses, was now needed to verify the applicability of this material for space use. Figure 4 outlines the test developed for high vacuum testing of coating response to ammonia. The coatings were pretreated at 10^{-6} Torr for four hours at ambient temperatures before undergoing this test.

- | |
|--|
| <ol style="list-style-type: none">1) Leak ammonia into chamber under dynamic vacuum2) Hold vacuum at consecutively lower half-orders of magnitude<ul style="list-style-type: none">• 15 minutes each• until color change occurs3) Return to 10^{-6} Torr4) Monitor color retention time |
|--|

FIGURE 4. Simulated Space Station Leak Test Procedure

When the previously tested formulas were exposed to SSSL conditions they exhibited no color response. Three modifications were made to the coating system to rectify the situation. First was the addition of an elastomeric primer coating to enhance the coating integrity and permit a higher solids content. Formulations developed much earlier, consisting of much higher total dye and solids content, were found to be too friable for rugged service. These recipes, however, when tested after aging in vacuum at elevated temperatures, were found to maintain excellent sensitivity to ammonia. By priming the surface of the test coupon with an elastomeric coating, we were able to incorporate these high concentrations of both dye and activator into the formula. The dye and coating is semi-encapsulated and imbibed within the primer coating, resulting in a coating more resistant to impact and abrasion damage. Thirdly, a humectant was added to the detector paint formula to facilitate color change at high vacuum. The dye works in an ionic environment which is typically provided by absorbed moisture at atmospheric conditions. Under high vacuum conditions, the coating is quickly desiccated and the dye is unable to perform unless other ionic media (i.e., the humectant) is available.

The response of the modified formula to SSSL testing was very encouraging. The coating turned a medium blue-green after five minutes at 1×10^{-4} Torr. The color change was maintained for two and one-half minutes after addition of ammonia was stopped and vacuum returned to 10^{-6} Torr. Additional refinements to the formula ratios brought the response level up to 5×10^{-5} Torr.

LONG TERM AGING

Two, long-term static tests were performed: a twelve-week and a two-week test. The two-week test was used as a control for the thermal cycling. A total of 24 panels were aged for twelve weeks at 150°F and greater than 10^{-6} Torr of vacuum. Four SSSL tests were done throughout the twelve weeks (1-day, 5-day, 6-week, and 12-week tests) to monitor the progress of the paint. The visual response data shows a slow degradation of sensitivity over time (Figure 5).

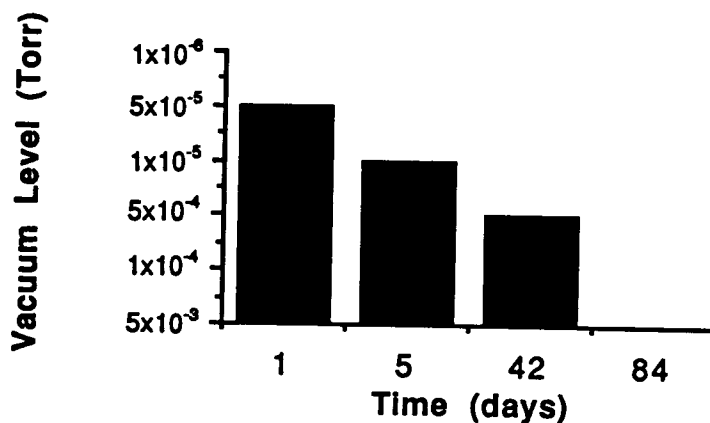


FIGURE 5. SSSL Vacuum Response Levels for Paint Aged at 150°F .

The paint lost one order of magnitude sensitivity to ammonia in vacuum during the first six weeks. It was noted that the paint turned from a bright yellow to a dull yellow during that period of time and then developed brown spots during the seventh week. These brown spots are believed to have been caused by the crystallization of one of the paint components and were a contributing factor to the paint failure at twelve weeks.

Although the paint lost sensitivity, it gained color durability over time. This is most likely because the dye slowly distills out under the high vacuum conditions leaving the paint surface devoid of indicator. The ammonia vapor has to diffuse deeper into the paint before encountering the dye and causing the color change. The reverse process takes longer and, therefore, color durability is increased.

THERMAL CYCLING

The temperature cycle representing the expected extremes on the space station was selected for a two week test. The temperature range chosen was 80°C to -60°C over a 30-minute cycle. The temperatures were regulated by a liquid nitrogen stream and heating elements. The test results show no significant change in the sensitivity due to the thermal cycling. Due to the effect of the liquid nitrogen on the system seals, a pressure fluctuation of 5×10^{-6} to 1×10^{-3} Torr (Figure 6) was also noted during the cycle.

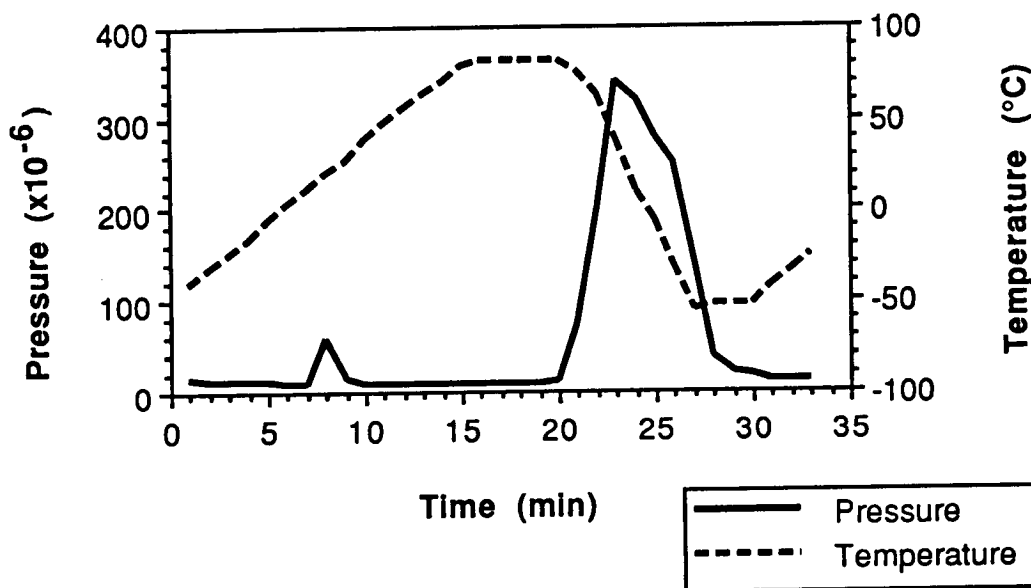


FIGURE 6. Thermal Cycling: Pressure and Temperature History.

SSSL testing was done after one week and again after two weeks were completed. In order to separate the effect of the thermal cycling from simple aging of the paint another two-week control test was done holding the panels static at the average temperature and pressure of the thermal cycling test. The test results show no significant change in sensitivity due to the thermal cycling (Figure 7).

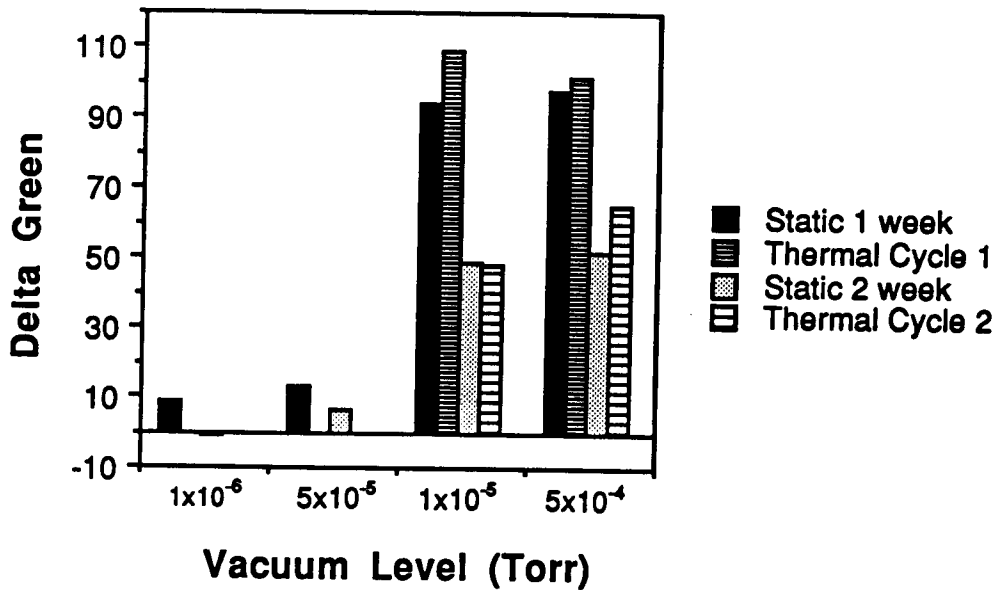


FIGURE 7. Effects of Thermal Cycling on Sensitivity.

PAINT SHELF LIFE

Tests were carried out on several panels painted at approximately 7 months apart from the same batch of paint. The paint was stored in a closed, glass container during the interim, under standard laboratory light. Both tests show the paint to be sensitive to ammonia at the 5×10^{-5} Torr vacuum level, therefore, indicating no need for concern regarding storage time.

MATERIAL COMPATIBILITY

The compatibility of ammonia detector paint was measured against ten substrate materials and ten identical materials primed with an elastomeric coating in advance. The primer was a blend of Dow Corning's silastic, with a 15% by weight of International Filler Corp.'s cellulose fiber. The coated samples were stored in a dark, paper board container for seven weeks. Examination at the end of seven weeks revealed the following qualities:

- When the substrate was unprimed and comprised of polycarbonates, standard glass, brass, mullite, a high linen content paper, silicon carbide, and some grades of aluminum, the paint altered and became ineffective as a detection media. However, when the surfaces were primed with room temperature vulcanizable elastomer, it proved to be compatible with all substrate materials.
- Polar minerals such as mullite, silicon carbide, and aluminum, together with certain metals such as brass, may contribute to discoloration, darkening or bluing of the coating depending upon the nature of the cation, the pH, and the electrophilicity of the substrate. The silastic primer, however, serves both as an aid to durable adhesion and as a buffer or barrier to the otherwise deleterious substrate surface chemistry.

The apparent implication is that any substrate material offers a viable choice for such a coating providing it is first coated with the silastic primer, whose design function was to assure a durable bond with minimal friability and susceptibility to erosion or abrasive loss. In addition, it serves as a protective barrier to isolate the surface chemistry of otherwise incompatible substrate materials and prevent interaction with the ammonia detector paint.

CONCLUSIONS OF THE NASA PROGRAM

Testing performed in a methane/carbon dioxide environment showed no interference in ammonia detection from these compounds. The effects of high vacuum combined with heat identified a need for a less volatile dye, or for measures to constrain the dye from sublimation or evaporation. A humectant was added to retain the polarity of the detector coating in absolute vacuum assuring an ionic media for response to ammonia. In addition, priming the surface of the test coupon with an elastomeric coating allowed us to incorporate high concentrations of both dye and activator into the formula. When applied to the elastomeric substrate, the dye coating became semi-encapsulated and imbibed within that primer coating. This composite coating is as sensitive to ammonia as the unprimed systems tested in this program and eliminates damage to the coating resulting from casual handling.

Testing of the detector coating under end-use conditions explored long-term durability under high vacuum, thermal cycling effects, and material compatibility with a variety of substrates. The paint recipe was altered to balance good tactile properties with reasonable sensitivity to ammonia under long-term high-vacuum conditions. No degradation due to thermal cycling was observed and the paint survived for seven weeks at 150°F at 10⁻⁷ Torr. The paint is compatible with nearly all substrate materials provided that an RTV primer coating is applied first. The largest single design consideration that must be accounted for is the pressure condition required for the paint to operate. The paint can withstand 10⁻⁷ Torr; however, it needs to be at 10⁻⁵ Torr or higher pressures before it can detect ammonia. This means it must be positioned in a semi-closed environment near the source of an ammonia leak, so that a partial pressure can build up at the paint surface during a leak.

ALTERNATE NON-AEROSPACE APPLICATIONS OF DETECTION PAINT

The simplicity, instantaneous response, and low cost characteristics of this paint technology suggest a variety of non-aerospace applications, such as:

1. early detection of industrial toxins;
2. chemical warfare agents detection;
3. detection of food spoilage (protein decomposition products);
4. leak detection in pressure vessels, aircraft fuselages, storage tanks, etc.;
5. detection of biological contaminants in air;
6. detection of urinary tract problems detection; and
7. work area environment monitoring.

In each of the above applications, coated coupons, roll out tapes, or painted components could provide the method or media for deployment of the indicator. A pressure vessel could be leak tested by simply filling the vessel with air containing a few parts per million of ammonia (8 ppm is the threshold for average olfactory response), and a painted wand, coupon, or probe used to inspect welds, fittings, enclosures, flanges, etc.. Alternatively, suspect areas could be spot painted and visually monitored for tell-tale traces of blue, green, or purple color.

Air could be monitored for toxins and pollutants by a columnated stream of air pumped over the surface of a painted coupon, or a plenum established to collect and direct a stream of air over a coated surface.

Frozen or refrigerated foods in transit, storage, or retail display could be passively monitored by packaging with a small, visible, coated button in a wrapped or boxed item. Meat and seafood emit amino compounds as they decompose. A nominal, acceptable color change could be determined to represent the acceptable rate or degree of deterioration. This would provide a litmus-like scale of freshness. The paint's color change is proportional to concentration. The indicator button could simply be a small paper, plastic, metal, or cardboard substrate coated with the silicon rubber primed paint, and would cost fractions of a cent to produce.

Cats, dogs, and other pets are frequently prone to bladder and kidney infections, which quickly manifest by the enzyme decomposition of urea, releasing an amino compound that is very aggressive toward the detector paint. Coupons, Q-tips, and other hand-held monitoring devices could be used to survey the litterbox or the collection area of the animal's excrement and determine the health of the animal. This would be particularly useful in veterinarian hospitals and kennels, zoos, etc.. Since all urine discharges ultimately do hydrolyze by enzymatic action, this technique would only be useful if the facilities are frequently washed or flushed and the aging excrement periodically and frequently removed.

RECOMMENDATIONS

There is no known, ultra-sensitive, passive technique available which can equal the response time and semi-quantitative nature of this detection system. It may be deployed as a simple coating that provides visual indications of the presence of hazardous chemicals or a leaking vessel, or it may be instrumented to provide continuous recorded monitoring of the targeted environment, whether it be municipal, military, or industrial.

This low cost, easily prepared, and simple to use detection paint technology offers numerous options to industry, government, and municipalities for applications relevant to safety, health, and the environment. SwRI staff members are prepared to adapt this material to the clients' needs and to tailor the detection monitoring, and control systems to their specifications.

AUTHOR: Evelyne Orndoff

TITLE: DEVELOPMENT AND TESTING OF TEXTILE FIBERS
IN THE JSC CREW AND THERMAL SYSTEMS DIVISION

ABSTRACT

New textile fibers have been developed or modified to meet the complex and constraining criteria of space applications. The most common of these criteria are light weight, nonflammability or flame retardancy, and high strength and durability in both deep space environment and the oxygen enriched crew bay area of the spacecraft. The fibers which successfully pass the tests of flammability and toxicity, and display the desired mechanical properties are selected for space applications. Such advanced fibers developed for the Crew and Thermal Systems Division (CTSD) at the Johnson Space Center include "Beta" fiber, heat stabilized polybenzimidazole and polyimide, as well as modified aramid Durette™, multifibrous Ortho™ fabric, and flame resistant cotton. The physical, mechanical, and chemical properties of these fibers are briefly discussed. The testing capabilities in the CTSD laboratory to ascertain some of the properties of these and other fibrous materials are also discussed. Most of these materials developed for spacecraft, space suit, and flight equipment applications have found other commercial applications. These advanced textile fibers are used mostly for aircraft, transportation, public buildings, hospitals, and protective clothing applications.

DEVELOPMENT AND TESTING OF TEXTILE FIBERS IN THE JSC CREW AND THERMAL SYSTEMS DIVISION

One of the major challenges of manned space flight applications is the selection of protective clothing for the astronauts. Since the beginning of manned space flights, various textile materials have been used in protective clothing for the astronauts. In the Mercury program, the astronauts' suits were made of two layers of fabrics: a layer of aluminized nylon and a bladder of neoprene impregnated nylon. In the Gemini program, the astronauts' suits were also made of two layers. The inner layer was a bladder made of nylon, the same as used for the Mercury program's suit. In addition, a new material was used for the outer layer to serve as a thermal and micrometeoroid garment: Nomex, an aromatic polyamide. Then, the need for materials with better resistance to heat and flame came with the Apollo program in which extravehicular activities (EVA's) had to be performed on the lunar surface. The newly developed Beta glass fiber was then introduced in the space program and used as the outer layer of the EVA spacesuit. The Beta fabric was used throughout the Apollo, Skylab, and Apollo-Soyuz Test programs. It was replaced by the Ortho composite fabric in the Space Shuttle programs. Beta and Ortho were both developed for space applications. The intensive research to improve the properties of textiles for space flight applications also led to the modification of existing fibers as in the case of Durette, stabilized polybenzimidazole, and treated cotton.

The development of Beta fiber was initiated because an extra-fine non-flammable fiber was needed for use in a 100 percent oxygen atmosphere. Beta is characterized by its diameter of 3.8 microns, its low relative bending stiffness and high tenacity (Table 1), and its low flammability and low offgassing (Table 2). Because of its inorganic nature Beta fiber can withstand temperatures at which no known organic material can exist. Beta has a service temperature range between -300°F and 900°F (-184°C and 482°C) and is degraded at 1550°F (843°C).

Table 1.- Comparison of Fibers Properties (1)

Fiber	Weight (size) denier (microns)	Tenacity, gpd	Relative Bending Stiffness*
Beta	1/4 (3.8)	15.3	1
Acetate		1.5	5
Viscose	1-1/2 (16)	1.0- 5.3	7
Silk	1.2 (14x8)	5.1	7
Cotton	1-1/2 (16x6)	4.9	8
Nylon 6,6	1-1/2 (14)	2.3 -6	11
Polyester	1-1/2 (12)	2.0 -9.0	14
Glass Fiber (G)	1-1/2 (9)	6.5	36

Table 2.- Flammability of Beta Fiber (1)

- COMPARISON OF FLAMMABILITY AND OFFGASSING
PROPERTIES OF HIGH-TEMPERATURE AND FLAME-RESISTANT TEXTILE
FIBERS

[Silicone ignitor, 100 percent oxygen]

Material	Flame spread rate, in/sec					Offgassing		
	Top ignition		Bottom ignition			Total organics, µg/g	Carbon monoxide, µg/g	Odor
	16.5 psia	6.2 psia	16.5 psia	16.5 ^a psia	6.2 psia			
Beta	NI ^b	NI	NI	NI	NI	0	.6	1.3
Teflon	SE	SE	.55	.30	.30	34	.7	.9
PBI	.20	.16	.41	.35	.30	.4	1.7	1.5
Asbeston	SE ^c	SE	SE	SE	SE	1.3	1	1.7
Nomex	.33	.16	1.00	.60	.60	1	.4	.7

^a60 percent oxygen, 40 percent nitrogen.

^bNI = no ignition

^cSE = self-extinguishing

Beta fibrous structures were used in the Apollo EVA spacesuit, the Apollo intravehicular flight suit, and also as a fire barrier for different flight items. Beta is used now in the Space Shuttle payload bay as contamination control and thermal protection systems covers. Its commercial applications include public structures such as hospitals and nursing homes; Beta is used as roofing material for stadiums, airports, shopping malls, and convention centers. It is also used as curtains and draperies in aircraft.

In the Space Shuttle Program, Beta was replaced by the Ortho composite structure for two reasons: the 100 percent oxygen environment was no longer needed, and the same suits would be used for many flights. These two new circumstances implied that the material selection would be driven less by flammability and more by resistance to wear. Consequently, in order to withstand the 30 percent oxygen environment limit imposed for Space Shuttle operation, the material could be made of organic fibers offering the advantage of reduced brittleness and better resistance to wear. Ortho, a composite multifibrous structure, offers strength, excellent abrasion and wear resistance, high temperature and chemical resistance, as well as low temperature flexibility (appendix A). Ortho is actually made of three fibers: Nomex, Kevlar, and Gore-Tex. These three fibers are assembled in a complex weave pattern which gives two distinct faces to the Ortho fabric; there is Gore-Tex on one face forming a surface with low coefficient of friction, and Nomex and Kevlar on the other face for strength and durability. Ortho is currently used only for space applications in the extravehicular mobility unit. Ortho is a good example of NASA development for the unique requirements of space applications.

Aside from using known fibers in a unique combination, as demonstrated with the Ortho composite structure, known textile fibers have also been modified for space applications. This is the case with Durette which was produced by Monsanto under CTSD contract in the 1960's. Durette was Nomex treated with chlorine and bromine gas in inert atmosphere at approximately 300 °C for 15 minutes (2). Such a treatment increased the flame retardancy of the material so that it could meet space applications requirements. Other treatments of fibers have been used to increase flame retardancy as well. In the case of cotton, a tetrakis phosphonium hydroxide/ammonia basic method is used. This particular treatment was chosen because it did not adversely affect other requirements such as minimum alteration of the fabric's properties and durability. A second finishing process using diammonium phosphate/urea would increase the flame resistance further up to a 30 percent oxygen atmosphere at 10.2 psia (3). Stabilized polybenzimidazole (PBI) fiber is obtained with a 3 percent phosphoric acid treatment at high temperature. The result is a

raise of the limiting oxygen index of PBI to 45 percent. The role of agents like halogens, phosphorus, or phosphoric acid is to inhibit solid/liquid state oxidation reactions, retard the formation of volatile condensable materials, and promote self-extinguishment of the fibers.

All these fibers, depending on their other properties, have different space applications. Durette (properties in appendix A) was used in the Skylab flight suit, inflight boots and gloves, and in the Apollo-Soyuz inflight suit. It is also used in various tests and ground support equipment applications as a felt or a woven or knitted fabric. The current flight suit and the training suit are made of treated cotton (appendix A). The astronauts enjoy the comfort of cotton, its good hand and drape, and its high moisture regain. Stabilized PBI (appendix A) which has had many applications since the Apollo program is also used in the Space Shuttle suit, mostly as cord, webbing, and tape. Stabilized PBI has the highest limiting oxygen index of the materials discussed and the highest moisture regain after cotton. This made it a material of choice for multiple applications (appendix B). In spite of the properties which made PBI an excellent material for space applications, costly processing methods to produce the monofilament and a limited market contributed to the decision made by the management of Celanese to close their production plant. Currently, polyimide monofilament has been chosen to replace PBI monofilament in many Space Shuttle applications. PBI and polyimide have comparable physical and mechanical properties (appendices A & C). Polyimide fibers already have many commercial applications for protective apparels and will have more in future station and advanced space programs.

Most of these materials which also have commercial applications for protective apparel are accepted for space applications only after extensive testing in the Advanced Materials Laboratory (AML) in the Crew and Thermal Systems Division and at White Sands Testing Facility. Physical and mechanical properties are usually evaluated in the AML. Essentially standard ASTM, FTMS, and AATCC methods are applied for testing textiles as shown in the test data sheets in appendix A. Tensile and bursting strength, abrasion resistance in various conditions, flexing and folding endurance, stiffness, moisture content, and limiting oxygen index are included in the systematic evaluation of textiles for space applications.

The selection of textile fibers for space applications is a challenging task. It has led to the development of the inorganic Beta fiber, the development of a composite structure like Ortho, and the modification of existing fibers like Durette, treated cotton, and stabilized PBI.

References

1. Dawn, Frederic, "Development and Application of Nonflammable, High-Temperature Beta Fibers," NASA TECHNICAL MEMORANDUM 102158, National Aeronautics and Space Administration, December 1989
2. Hathaway, C. E., Early, C. L., "New Fire Resistant Fabrics for Critical Applications," J. APPL. POLYM. SCI., Applied Polymer Symposia, 21, 101
3. Dawn, Frederic S., Morton, Glenn P., " Cotton Protective Apparel for Space Shuttle," Natural Fibers Textile Conference, Atlanta, Georgia, September 26-28, 1978

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE May/94	3. REPORT TYPE AND DATES COVERED Conference Publication - February 1-3, 1994
----------------------------------	--------------------------	---

4. TITLE AND SUBTITLE Dual-Use Space Technology Transfer Conference and Exhibition - Volumes I and II	5. FUNDING NUMBERS
--	--------------------

6. AUTHOR(S) Kumar Krishen, Compiler	
---	--

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Lyndon B. Johnson Space Center Houston, Texas 77058	8. PERFORMING ORGANIZATION REPORT NUMBERS S-765
--	---

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546	10. SPONSORING/MONITORING AGENCY REPORT NUMBER NASA CP-3263
--	---

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified/Unlimited Available from the NASA Center for AeroSpace Information 800 Elkridge Landing Road Linthicum Heights, MD 21090 (301) 621-0390	12b. DISTRIBUTION CODE
--	------------------------

Subject Category: 99

13. ABSTRACT <i>(Maximum 200 words)</i> This document contains papers presented at the Dual-Use Space Technology Transfer Conference and Exhibition held at the Johnson Space Center February 1-3, 1994. Possible technology transfers covered during the conference were in the areas of information access; innovative microwave and optical applications; materials and structures; marketing and barriers; intelligent systems; human factors and habitation; communications and data systems; business process and technology transfer; software engineering; biotechnology and advanced bioinstrumentation; communications signal processing and analysis; new ways of doing business; medical care; applications derived from control center data systems; human performance evaluation; technology transfer methods; mathematics, modeling, and simulation; propulsion; software analysis and decision tools; systems/processes in human support technology; networks, control centers, and distributed systems; power; rapid development; perception and vision technologies; integrated vehicle health management; automation technologies; advanced avionics; and robotics technologies. More than 77 papers, 20 presentations, and 20 exhibits covering various disciplines were presented by experts from NASA, universities, and industry.

14. SUBJECT TERMS microwave, biotechnology, intelligent systems, software analysis, automation, bioinstrumentation, software engineering, advanced avionics and robotics	15. NUMBER OF PAGES 921
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited
---	--	---	---